Climate Change, Demand Volatility, and Corporate Investment Decisions

Chen Lin
University of Hong Kong

Thomas Schmid
University of Hong Kong

and

Michael S. Weisbach
Ohio State University and NBER

March 15, 2019

Abstract

How does climate change affect firms' decisions? We investigate how more volatile electricity demand due to more extreme weather affects investment decisions of energy utilities. Using a global sample of early-stage power plant projects, we find that electricity-producing firms invest more in regions in which climate change is more severe. This increase is concentrated in plants using flexible technologies, for which firms can adjust output at relatively low cost. Thus, firms seem to increase their investments to adjust their production assets to changes in demand volatility. Overall, these results are consistent with the view that climate change has far reaching effects on firms' behavior, affecting their investment decisions and undoubtedly many others..

JEL classification: G30, G31

Key words: Climate change, firm investment, electricity generation, power plants, operational flexibility

Contact information: Chen Lin, Faculty of Business and Economics, University of Hong Kong, Hong Kong, email: chenlin1@hku.hk; Thomas Schmid, Faculty of Business and Economics, University of Hong Kong, Hong Kong, email: schmid@hku.hk; Michael S. Weisbach, Department of Finance, Fisher College of Business, Ohio State University, Columbus, OH 43210, email: weisbach@fisher.osu.edu. We are very grateful to seminar participants at University of Chicago for helpful suggestions.

1. Introduction

The G20 recently called climate change “one of the greatest challenges of our time”. Increasing temperatures and more extreme weather events in most regions of the world will lead to dramatic changes in society and the economy. While the most commonly discussed effects of global warming are potentially disastrous flooding and changes in the food supply, there are also a plethora of other changes that will occur in the world economy because of climate change. We are only beginning to understand the nature of these changes.

The U.S. Department of Commerce estimates that 70 percent of companies are directly affected by the weather. More federal money was spent on the consequences of the volatile weather during 2012 than on education or transportation; globally, the economic losses related to weather events amounted to about $150 billion.[1] Increasingly extreme weather conditions occurring because of global warming potentially have a major impact on many sectors of the economy. Changing conditions lead to prolonged periods of drought, heat waves, cold snaps, and more frequent natural hazards like tornados. Because changes in climate conditions affect firms’ profitability, stock markets as well as regulators become increasingly concerned about the issue, and firms’ decisions can be affected by changing climatic conditions. In a survey conducted for CME Group and Storm Exchange, eight out of ten senior finance and risk managers state “that the emergence of global climate change and accompanying volatile weather patterns will require changes to their business models in the decades ahead” (Myers, 2008, p. 4).

This paper investigates how firms respond to climate risk exposure through investments and changes to their asset structure. We focus on one industry for which detailed asset-level data on investment decisions is available: the electricity producing industry. This industry is meaningfully affected by climate change: more extreme and less predictable weather leads to larger fluctuations in the wholesale price of electricity, which in turn affects the optimal production process for firms to use.[2]

[1] Cf. Allianz (2013), p. 7.

[2] Electricity producers continually turn on and off their power plants as a function of the prices they face, and the cost of adjusting production varies substantially by method of production (see Reinartz and Schmid, 2016, or Lin, Schmid, and Weisbach, 2017).

Furthermore, this industry is very concerned about weather risk and has made a systematic attempt to quantify its impact (Myers (2008)).

In this paper, we consider a sample of 384 publicly traded electricity-generating firms from 67 countries between 2000 and 2015, constituting the vast majority of publicly traded electricity-generating firms in the world. These firms operate about 60,000 power plants, and make investments in 13,867 new, early-stage plant construction projects during our sample period. We evaluate the extent to which these investment decisions are affected climate conditions. In particular, we measure the impact of changing weather conditions on the quantity of firms' investments in new power plants, as well as the type of power plants that these firms build.

Using asset-level data from worldwide energy utilities to analyze investment decisions has at least three advantages relative to accounting-based measures like capex. First, our data includes detailed information on early-stage power plant projects, so we know exactly when each firm makes its investment. Second, we know the exact location of a particular plant. We can use this locational information to identify the effect of climate change on investment, since we can measure the change in weather over time in any particular location. Empirically, this information on differences in weather changes across the globe allows us to exploit within firm-year variations in investment decisions across different regions of the world. Finally, and most importantly, we can observe the type of power plant that is built. The *Platts World Electric Power Plant* database provides detailed information on the planned power plants, including their production technology. This information allows us to distinguish between flexible power plants such as gas-fired plants and inflexible generation assets like nuclear or coal power plants.

To measure climate change, we follow the climate literature (e.g., Hansen et al., 2012) and use measures based on the change in the yearly average temperature in a region. As our main measure of climate change, we use the yearly abnormal temperature index which measures the perceived change in temperatures in a region.

Theoretically, climate change could potentially either decrease or increase firms' investments. The first possible effect of climate change, which would lead to fewer investments, is an increase in uncertainty facing the firm. This increased uncertainty could come from a number of factors, including demand for their products, the production processes they should use, or potential new regulations. For instance, climate change could have substantially affect environmental regulations, but if it is difficult to forecast the exact changes, firms could defer investments until the regulations are enacted. Alternatively, electricity producing firms could invest more to adjust their power plant portfolio to the changing conditions, which likely include more extreme weather events and less predictable and more volatile demand for electricity. For example, climate change likely will lead to more events like the 2015 heat wave in Texas, during which there was a record high demand for electricity, suggesting that an increase in capacity could be valuable to electricity producers.[3]

If firms invest more in new power plants in response to climate change, an increase in demand volatility should lead to a change in not only the capacity of power plants, but also their type. First, firms could desire a more flexible supply to avoid system breakdowns ("blackouts") because electricity is not storable.[4] In regions with wholesale markets for electricity, higher demand volatility likely also leads to more volatile prices, which makes flexible power plants more attractive for firms can adjust output at a low cost. For this reason, when weather becomes less predictable, energy utilities will tend to adjust their power plant portfolio towards plants with more operational flexibility, such as gas-fired power plants that can start and stop within minutes. Second, firms could want to increase their investments in renewable power plant types like hydro, wind, or solar to reduce their CO2 emissions in regions in which climate change is more severe. In these regions, public pressure is likely to be especially high on firms to act in an environmentally friendly manner. These arguments imply that we should observe more new power plants in regions in which climate change is more pronounced, especially in flexible and/or renewable generation assets.

[3] http://www.reuters.com/article/us-texas-electricity-idUSKCN0QG1H320150811
[4] An overview on the different approaches to balance electricity supply and demand is provided by Hunt (2002), Chapter 7.

Our estimates indicate that utilities tend to increase their investments in new power plants in regions in which climate change is more pronounced. The estimates imply that a one-standard deviation increase in the abnormal temperature index leads to an increase in the capacity of planned plants relative to existing ones of about two percentage points, which equals 12.5% of what investment would be absent climate change. This increase in investment does not appear to be driven by region-specific factors like economic development. These results imply that electricity producing firms increase their investments in new power plants as response to climate change.

To understand the way in which firms change their investment decisions in response to climate change, we examine the type of incremental power plants that they build. Our estimates suggest that changing weather leads utilities to invest more in flexible production assets (e.g. gas-fired power plants) in regions in which climate change is more severe, measuring new investment both in absolute terms and relative to the firm's existing investments. These new plants will tend to increase firms' overall operating flexibility in these regions.

In addition, we examine whether climate change leads to an increase in other types of power plants, especially renewables such as solar or wind power. Our estimates indicate that climate change does not affect investments in other types of power plants such as renewables or inflexible generation assets.

A possible reason why firms adjust their production portfolio in this manner is that climate change could affect uncertainty about electricity demand, which in turn could lead firms to favor production techniques for which quantities can be adjusted at low cost. The logic of this argument depends on larger temperature changes being associated with more extreme weather events, which cause volatility in the demand for electricity. We document that in our sample, changes in average temperatures do lead to more extreme weather events. In addition, we show that for regions with competitive wholesale markets, wholesale electricity price volatility increases with the region's abnormal average temperatures. These results are consistent with the view that flexible power plants, which can start and stop within short periods of time, have become relatively more attractive for energy utilities because of the increase in electricity demand uncertainty, itself a consequence of climate change.

Our results suggest that energy companies adjust their production portfolios to changing weather conditions by increasing the quantity and flexibility of their production facilities. As has been suggested by Al Gore, climate change does create investment opportunities for firms.[5] If we accept that on net, CO_2 emissions have negative social consequences, then an interesting question concerns the way in which the new power plant investments induced by more extreme weather affect emissions. Since coal-fired plants are perhaps the worst offender in terms of CO_2 and also happen to be a relatively inflexible production technique, the incentives produced by climate change toward more flexible production has the social benefit of discouraging coal-fired plans. However, our results do not imply that climate change leads to higher use of renewables. Instead, firms appear to prefer more flexible techniques such as gas-fired plants, which do produce CO_2, but not to as high a degree as coal-fired plants.

Although this study focuses exclusively on energy utilities, our results potentially have implications for other industries as well. Weather-sensitive industries also include the food production industry, which needs to adjust its production process to changing climatic conditions,[6] the tourism industry which needs to adjust its "assets" because consumers' demand shifts to different regions,[7] and other industries like steel that will undoubted face more stringent environmental regulations in the future.[8] While the nature of the impact of climate change is likely to vary by industry and is undoubtedly different from the electricity generating industry we study here because of the unique relation between energy demand and weather, the real effects of climate change on businesses in many industries are likely to be consequential, and are not well understood at this point.

Our analysis extends the literature in a number of ways. First, we provide asset-level evidence on firms' investment decisions. In contrast to traditional measures like capex, this asset-level data allows us

---

[5] Stanford GSB Conradin von Gugelberg Memorial Event 2016 (https://www.gsb.stanford.edu/insights/al-gore-business-will-drive-progress-climate-change).
[6] https://www.theguardian.com/environment/2012/sep/19/climate-change-affect-food-production
[7] http://www.ktoo.org/2014/08/04/report-alaska-tourists-may-shift-new-areas-climate-change/
[8] http://www.reuters.com/article/china-steel-environment-idUSL4N0VE3R820150204

to measure details about the investment project, such as production technology it uses and its exact location. Furthermore, this approach enables us to approximate the time when the decision to invest was made better because we can observe early-stage projects. This novel measure allows us to contribute to the investment literature. In particular, this study is related to prior work which investigates how investment is linked to uncertainty (e.g., Abel, 1983; Dixit and Pindyck, 1994; Bloom, 2009; Julio and Yook, 2012; Gulen and Ion, 2016) or product market characteristics (e.g., Dixit, 1980; Akdogu and MacKay, 2008; Frésard and Valta, 2016).[9]

Second, we are – to the best of our knowledge – the first to investigate how climate change affects firms' investments. Determents for firms' investment decisions that are discussed in the previous literature are, among others, debt capacity and collateral (e.g., Gan, 2007a), bond market access (e.g., Harford and Uysal, 2014), access to bank lending (e.g., Gan, 2007b; Amiti and Weinstein, 2017), corporate governance (e.g., Billett, Garfinkel, and Jiang, 2011), or market timing considerations (e.g., Bolton, Chen, and Wang, 2013). We extend this literature by showing that climate change as a macro factors has a strong explanatory power for firm-level investments.

Third, this paper also contributes to a long line of research about the economics of energy utilities. Fabrizio, Rose and Wolfram (2007), for instance, analyze how deregulation affects the efficiency of energy utilities. Becher, Mulherin, and Walkling (2012) investigate corporate mergers in the energy utilities industry, Perez-Gonzalez and Yun (2013) use energy utilities to measure the value of risk management with derivatives, and Rettl, Stomper, and Zechner (2016) evaluate the importance of competitor inflexibility in this industry. Reinartz and Schmid (2016) analyze the impact of production flexibility on the financial leverage in the electricity-generating industry, and Lin, Schmid, and Weisbach (2017) investigate how price risk due to electricity price volatility and production inflexibility affects firms' cash holdings.

[9] Other papers using asset-level data in an investment context include, among others, Gilje and Taillard (2016) who investigate how listing status affects investments in gas drilling projects, Kellog (2014) who analyzes how firms' drilling activities react to changes in uncertainty, and Greenwood and Hanson (2014) who study investment cycles in the shipping industry.

2. Data description

2.1. Sample of energy utilities

To construct a global sample of energy utilities, we start by combining lists of active and inactive utility companies from *Thomson Reuters*. We focus on stock market listed utilities because reliable data for unlisted firms is often not available. The sample covers the years 2000 to 2015, which is the period for which we can obtain the necessary annual data on firms' production assets. We perform several steps to clean the sample. First, we eliminate all firms without a primary security classified as equity. Second, we wish to consider only companies that focus on the generation of electricity. To ensure that other companies are not included, we rely on firms' SIC and ICB codes, the business description obtained from *Capital IQ*, and additionally conduct manual research on the companies' business lines. For this sample, we collect accounting data like total assets or total debt are obtained from *Worldscope*. After applying these selection criteria, we end up with a sample of 384 energy utilities. These firms operate about 60,000 unique power plants in 67 countries and they conduct 13,867 early-stage power plant construction projects.

Because of the nature of the available weather data (see below), we group the power plants into "regions", which a region is a state or province for all plants in the U.S., Canada, and Australia, and a country for all other plants. There are 135 different regions represented in our sample.

2.2. Measuring power plant investments

Data on individual power plants is obtained from the annual versions of the *Platts World Electric Power Plant* database, which provides information on power plants and their technologies around the globe. It includes information on single power plant units, including their production technologies, capacities, geographic locations, start dates of commercial operation, and their owners/operators.[10] We

[10] A detailed description of the database is provided by *Platts' Data Base Description and Research Methodology*.

obtain the annual version of this database for all years between 2000 and 2015 and manually match each power plant in this database to the energy utilities sample.[11] About 50% of the existing plants match to our sample firms; the remainder are owned by large utilities that are not publicly listed and are excluded from our sample for this reason.

Most important for our purposes, this database does not cover only completed power plants, but also contains information on early stage construction projects.[12] These data on existing power plants as well as early-stage power plant construction projects allow us to construct our main investment variables. The variable "TOTAL INVESTMENT" is defined as early-stage power plant construction projects (in megawatt, MW) of firm *i* in region *j* and year *t*, scaled by the capacity of existing power plants (in MW) of the same firm *i* in the same region *j* and year *t*. The variable is set to one for values above one and to zero if firm *i* in region *j* and year *t* has existing production capacity of at least one megawatt, but no planned investment.

For robustness, we also use alternative variables to measure total investment. First, we apply a dummy variable that equals one if firm *i* has any early-stage power plant projects in region *j* and year *t.* This variable is set to zero if total investments is zero. Second, we use the unscaled logarithm of one plus the total capacity of all early-stage power plant projects of firm *i* in region *j* and year *t*. Again, this variable is set to zero if firm *i* in region *j* and year *t* has existing production capacity of at least one megawatt, but no planned investment. Third, we consider the number of planned power plants, rather than their capacity. For this variable, we scale the number of planned projects of firm *i* in region *j* and year *t* by the number of existing plants of the same firm in the same region. Fourth, we use the logarithm of the total number of plants instead of their capacity.

(www.platts.com/IM.Platts.Content/downloads/udi/wepp/descmeth.pdf). Reinartz and Schmid (2016) contain additional information about it as well as other information about electricity markets.

[11] We use the yearly version of the database because they only include the current owner/operator.

[12] Platts states that "the decision to include new power projects in the WEPP Data Base is […] made on a case-by-case basis. Key determinants in approximate order of importance are: 1) order placement for generating equipment or engineering, procurement, and construction (EPC) services, 2) the status of licensing or permitting activities, 3) funding, and 4) the availability of fuel or transmission access. Projects may also be included even if such data are lacking if there are generalized national or regional policies that are driving power plant development." (Platts Data Base Description and Research Methodology, p.19).

An advantage of the *Platts* database is that it provides detail on the technology a particular plant uses to produce electricity. This information is important because the flexibility to adjust production levels a low cost is a potentially important advantage when faced with volatile demand for electricity. For this reason, we classify new power plants as flexible, inflexible, or renewable. We classify gas, gas-combined cycle, and oil power plants to be flexible generation assets, coal-fired plants and nuclear plants to be inflexible, and hydro, wind, and solar plants to be renewable generation assets. We also construct firm-level measures which indicate a firm's fraction of existing flexible power plants overall ($\text{FLEXIBILITY}_{\text{overall}}$) or in a region ($\text{FLEXIBILITY}_{\text{region}}$). As a measure of the overall flexibility of a firm's plants, we calculate the average "run-up time" for its plants, which measure how long on average a firm's plants take to start up. (see Reinartz and Schmid, 2016, for more details about the construction of run-up time).

Table 1 provides an overview on the early-stage power plant projects, separately for flexible, inflexible, and renewable plants as well as the single technologies therein. There are 4,551 early-stage projects to construct flexible power plants, which account for a total of 1,048 GW. The average capacity of planned flexible plants is about 200 MW, and the projects are located in 96 regions. The majority of planned flexible plants are gas combined-cycle plants. For inflexible plants, there are 1,974 projects with a capacity of 1,167 GW. The average planned inflexible plant has a capacity of about 800 MW, four times the average capacity of planned flexible plants. For renewables, we identify about 6,000 early-stage projects, mostly in wind and hydro plants. Their combined capacity is 670 GW, and the average size of planned renewable plants is 93 MW. The "others" group, which includes for example pump-storage or waste plants, is not explicitly considered when we distinguish the power plant types, but these plants are included in our analysis focusing on total investments. In total, there are 13,867 unique early-stage investment projects in our sample, which have a combined capacity of 3,057 GW.

2.3. Measuring climate change

Our weather data come from the Global Historical Climatology Network (GHCN). We use the daily average temperatures (GHCN-DAILY) to construct our climate change proxies. Based on approximately 100 million individual observations, we start by calculating the average temperature in degrees Celsius in each region and year from 1951 to 2015.[13] We use the average temperature during the base period 1951 to 1980 for each region as the "base period," and define the ABNORMAL TEMPERATURE in a year and region as the difference between the actual temperature and the average temperature during the base period.[14]

To account for the fact that temperatures can be more volatile in some regions than in others, we also construct the ABNORMAL TEMPERATURE INDEX. For its calculation, we follow Hansen et al. (1998) and divide the abnormal temperature in a region and year by the inter-annual standard deviation during the base period 1951 to 1980 in the same region. Thus, a value of one means that the temperature is one standard deviation higher than the average temperature during the base period.[15]

The average abnormal temperature across all sample countries over time is presented in Figure 2(a). This figure documents that there is a strong temperature increase starting around 2000: from 2000 to 2015, the average abnormal temperature is about 0.6 degrees Celsius. A similar time trend can be observed for the abnormal temperature index in Figure 2(b). The average abnormal temperature index

---

[13] Temperatures are calculated as the average across all weather stations in a region. We only include weather stations with at least 25 years of data to avoid having newly added stations bias the time trend of the temperature in a region. In Austria, for example, has xx weather stations, so the temperature for Austria on a given day is the average of the temperatures reported by the weather stations on that day. This approach is potentially problematic for larger countries like Russia, which has yyy weather stations. The results are similar if we exclude Russia and China from the sample, which are sufficiently large that the average of the weather stations is not potentially reflective of the weather at any particular location.

[14] Hansen et al. 2012 explain they "choose 1951–1980 as the base period for most of our illustrations, for several reasons. First, it was a time of relatively stable global temperature, prior to rapid global warming in recent decades. Second, it is recent enough for older people, especially the "baby boom" generation, to remember. Third, global temperature in 1951–1980 was within the Holocene range, and thus it is a climate that the natural world and civilization are adapted to. We require at least 20 years of non-missing data on the average temperature during this base period.

[15] Hansen et al. (1998) state that "[a] value +1 (or -1) is great enough to be noticeable, because a value that large or larger would normally (that is, in the period 1951-1980) occur only about 15% of the time. For example, if the summer is warm enough to yield an index of +1 or greater at a given place, most people who had been living at that location for a long time would tend to agree that it was a "hot" summer." (p. 4114)

during our sample period is slightly above one. Overall, these statistics show that average temperatures have increased substantially during our sample period.

Next, we consider whether there are differences in temperature increases across regions. Figures 3 and 4 present the average abnormal temperature and the average abnormal temperature index for each sample country during sample period. Given the overall increase in temperatures, it is not surprising that most regions experienced higher temperatures during our sample period if compared to the base period 1951 to 1980. Of our 135 regions, 117 experienced a temperature increase and 18 a temperature decrease. The strongest increase in temperature can be found in Europe, Equatorial Africa, and Central America. For several U.S. states close to the West Coast and a handful of countries we find a moderate decrease in the average temperature. The overall picture looks quite similar if we analyze the abnormal temperature index in Figure 4. Overall, this evidence indicates that temperatures on average have risen considerably around the globe, and that these increases are heterogeneous across different regions.

2.4. Financial variables

Our source for financial variables is *Worldscope*. The control variables which we use are size (measured as the logarithm of total assets in $US), profitability (EBITDA scaled by total assets), Tobin's Q (market capitalization plus total debt scaled by the sum of book value of equity plus total debt), leverage (total debt scaled by the sum of total debt and book value of equity), cash holdings (cash and short-term equivalents scaled by total assets), and the logarithm of GDP per capita in the headquarter region of the firm. Fiscal years that end between January and June are allocated to the previous year; only complete fiscal years are considered. All financial variables are winsorized at the 1% and 99% levels. A detailed description of all variables can be found in Appendix A.

2.5. Descriptive statistics

Table 2 presents descriptive statistics for our sample firms, averaged for the whole sample period. On average, energy utilities have early-stage investment projects that account for about 16 percent of their existing generation in a region and year. Of these 15 percent, about five percent are related to investments in flexible generation, about three percent to investments in inflexible generation, and renewables account for about eight percent of existing capacity. If we only consider regions in which a firm is investing, we find that investments in flexible generation assets account for about one-third of all investments. The corresponding number for inflexible and renewable investments are 15 percent and 42 percent, respectively.[16] The new projects cause the average run-up time as a proxy for operating inflexibility to decrease by about 0.2. If we compare the old portfolio to the hypothetical new portfolio including the planned power plants, we find that run-up time remains approximately constant.

3. Estimating the impact of climate change on utilities' investments

3.1. Empirical specification

To estimate the effect of the changing climate on utilities' investments, we wish to have a specification that exploits variation in weather, which, at least from the point of view of any particular firm, is exogenous. As is evident from Figures 2 – 5, there is both variation in average temperatures over time and variation across different locations in the magnitude. Our goal is to estimate a specification that takes advantage of this variation in weather changes while at the same time controlling for firm-specific and regional factors.

The fact that many energy utilities operate in more than just one region allows us to observe multiple investment decisions of the same firm in the same year. The example of Vattenfall AB is illustrated in Figure 1. As of 2014, the Swedish power company owns production assets in Sweden, Denmark, Netherlands, Germany, the U.K. and, to a small extent, Finland. It is also worth noting that the production assets of Vattenfall are quite heterogeneous across its different regions. For instance, hydro power accounts for about one-third of the generation in Sweden, but less than 10% in Germany. Due to

[16] These numbers do not sum up to one because the "other" category, which is also considered for total investment, is not explicitly reported.

these multiple regions per energy utility, we can observe multiple climate change–investment combinations for the same firm in the same year because climate change differs across regions.[17]

We estimate an equation predicting investment for a particular firm in each region in which it operates. We present alternative specifications that vary in their use of firm-level and regional level controls. We first present a specification with just year effects, and one with year effects that contain firm level controls for firm-level factors such as size, profitability, etc. These specifications have the advantage of utilizing both time series and cross-sectional variation in weather. However, they do not control for firm or country level variables that are not captured by characteristics included in the equations. We next estimate a specification that contains year x firm-fixed effects. Because many firms in our sample, like Vattenfall, operate in more than one region, this specification nets out all firm-level effects while still taking advantage of variation in weather across the regions in which the firm operates. However, for firms that only operate in one region, it only utilizes time series variation in weather, and since multi region firms tend to operate in nearby regions with similar weather, the impact of cross-sectional weather variation is minimized. Finally, we present models with firm-region fixed effects, which come with the disadvantage of only exploiting time-series variation in climate change.

3.2. Estimates of the impact of climate change on investment levels

We present estimates of these specifications in Table 3. All columns include year-fixed effects; Columns 3 and 4 additionally include firm-fixed effects and Column 5 includes year x firm-fixed effects.[18] In each specification, the estimated coefficient for the abnormal temperature index (ATI) is positive and statistically significantly different from zero. This positive coefficient indicates that firms invest more in regions in which temperature increases are larger. The magnitude of the coefficient for ATI is between 0.01 and 0.02, which implies that a one-standard deviation change of ATI leads to a 1.5 to

[17] To avoid having tiny markets affecting our findings too much, we exclude regions in which a firms' production assets are less than 1 megawatt of capacity. The results are robust if we include those regions as well.

[18] The firm specific drop out if firm x year fixed-effects are included because these models only exploit variation within a firm and year. For this reason, only firms that have operations in different regions can be considered for this test, which also explains the smaller number of firms in column 5.

3 percentage point increase in the total investment (which is defined as the capacity of early-stage investment projects relative to existing power plants). Because the average total investment is 16 percent, these estimates imply that a one standard deviation increase in temperature corresponds to a 10 to 20 percent increase. Thus, this effect is not only statistically significant, but also large enough to be economically relevant. The effect of the control variables are in line with expectations: large firms, firms with higher market to book, and those with more cash on the balance sheet invest more on average. The coefficients for leverage and the GDP of a firm's headquarter country are negative, but the effect is statistically not significant in the specification with firm-fixed effects.

3.3. Robustness

Tables 4, 5, and 6 contain estimates of alternative specifications, which we include to ensure that that the relation between temperature changes and investment in new power plants are robust to alternative assumptions. The first robustness test in Table 4 focuses on the measurement of climate change. In the analysis presented in Table 3, we measure changes in climate using the abnormal temperature index (ATI). In Columns 1 and 2 of Table 4, we replace this with the average value of ATI over the past three years. If climate change is a slow process a longer-term trend rather than the one year temperature is likely to be informative about future weather conditions, and thus be a better predictor of firms' investment choices. In addition, we reestimate the equation using the abnormal temperature (not standardized) and its 3-year average in Columns 3-6 of Table 4. Both the magnitude and statistical significance of the coefficients on the variables measuring temperature in these alternative specifications are similar to those reported in Table 3.

As a second series of robustness tests, we measure investment in several different ways. Our main measure of investment scales the total capacity of early-stage plant construction projects by the capacity of the existing generation assets. In Column 1 of Table 5 we instead use a dummy variable that equals one if a firm has any investments in a region and year and zero otherwise. In Column 2, we use the non-scaled natural logarithm of the total capacity of all early-stage plant projects (plus one). Columns 3 and 4 focus on the number of plants instead of their capacity. In these specifications, we scale the number of all plant projects by either the number of existing plants, or the natural logarithm of all plant projects (plus one). The estimates in each specification are similar and all suggest that climate change leads to more total investments in new power plants.

The third series of robustness tests includes additional controls for regional characteristics. If these characteristics are correlated with climate change, excluding them from the specification could meaningfully affect our estimates. Although climate change, measured by the *abnormal* temperature in a region, is likely to be exogenous to other country-level factors, we nevertheless add regional macroeconomics and firm-level variables in Columns 1 to 6. The macroeconomic controls are regional GDP (instead of headquarter GDP), GDP growth in a region, and the level of inflation in a region. The firm-level controls are a firm's overall level of flexibility (i.e., capacity of flexible plants to total plants), its flexibility in a region, its total production capacity in a region, and the firm's capacity in this region to its total capacity. The estimates using each of these specifications are similar to the ones in Table 3, and suggest that abnormal temperature indices meaningfully affect firms' investment decisions.

In the Column 7 of Table 6, we add firm-region fixed effects to the specification. These fixed effects effectively control for all time-constant regional characteristics, such as legal origin or anti-director rights. We do not use this model as our main model because it has an important disadvantage: it only exploits time-series variation of our climate change proxy *within* a region. However, climate change is a slow process, so differences *between* regions should be more relevant than changes over short time periods. Nonetheless, the results from this model are similar to our prior findings, although the economic magnitude is smaller.

4. The Channel through Which Climate Change Affects Investment

4.1. Types of power plants

To understand why firms increase investment in response to climate change, we next turn to the issue of the types of power plants in which they invest. As discussed above, climate change is likely to lead to less predictable weather and to more extreme weather events, and thus less predictable and more volatile demand for electricity. Consequently, energy utilities potentially have incentives to adjust their power plant portfolios towards more operational flexibility. One way to increase operational flexibility is by building more gas-fired power plants which can start and stop within minutes. Alternatively, energy utilities could increase their investment in renewable generation assets because of public pressure or stricter environmental regulations to reduce CO2 emissions in regions in which climate change is more severe.

In Table 7, we estimate the way in which changes in weather affect investment in different types of power plants. We construct variables entitled FLEXIBLE INVESTMENT, INFLEXIBLE INVESTMENT, and RENEWABLE INVESTMENT in the same way as TOTAL INVESTMENT, except that that only power plant projects of a particular type are used to construct each variable. In Panel A, we present estimates of equations that predict each type of power plant investment using the same specification as in Table 3.

In the equations predicting investments in flexible power plants in Columns 1 and 2, the estimates indicate that climate change affects investment. The coefficient on the abnormal temperature index is about .01, and is statistically significantly different from zero. This coefficient is of a similar magnitude to the coefficients from the corresponding models for total investment in columns 4 and 5 of Table 2. This fact that the implied increase in flexible investments from a given change weather conditions is about the same magnitude as the overall change suggests that the increase in total investments is driven by investments in flexible production assets, such as gas-fired plants.

In Columns 3 to 6 we estimate the impact of weather changes on investments in inflexible power plants and renewables. In contrast to the estimates for flexible power plants, these estimates indicate that climate change has no effect on investment in inflexible or renewable production assets. The impact of changing weather on investment appears to come from a higher demand for flexible power plants whose output can be varied at low cost.

As an additional way of examining the types of power plants affected by changes in weather, we consider an alternative specification that predicts the relative investments in different kinds of plants. For this specification, we restrict our sample to firm-years in which firms make some investments, and use the fraction of a firm's total investment of a particular type as our dependent variable. We present estimates of equations predicting this fraction in Panel B of Table 7.

The estimates in Panel B of Table 7 are consistent with those in Panel A, and also suggest that increases in temperature lead to investments in flexible power plants, but not inflexible or renewable ones. In the equations in Columns 1 and 2 predicting the fraction of investments in flexible power plants, the coefficient on Absolute Temperature Index is positive and statistically significantly different from zero. In contrast, in the equations in Columns 3 – 6 predicting the fraction of investments in inflexible power plants and renewables, the coefficients on Absolute Temperature Index are small and not statistically significantly different from zero.

Finally, we analyze the extent to which changes in weather lead to overall changes in the production flexibility of a particular energy utility. To evaluate the change in a firm's production flexibility, we rely on a firm's "run-up time" as a firm-level flexibility measure. This variable is defined as the capacity-weighted average time which is necessary to start-up the power plants in hours.[19] Higher values of run-up time go along with less production flexibility because it takes longer to start and stop the plants.

We present estimates of the impact of weather changes in a firm's run up time in Panel C of Table 7. We first analyze the relative difference between the run-up time of early-stage plant construction projects and the existing plants. The results in Columns 1 and 2 indicate that when temperatures increase, the run-up time of the early stage projects tends to be lower than that of firms' existing plants. Consequently, once investments of firms faced with increasing temperatures are completed, their production processes will become more flexible. In Columns 3 and 4, we estimate the effect of

[19] It is based on the production technologies of the firms' power plants. See Reinartz and Schmid (2016) for technology-specific values.

temperature increases on the difference between the average run-up time of the firms' new power plant portfolio (consisting of their early-stage projects together with their existing plants) to that of the existing plants alone. The results again indicate that temperature increase lead the new production portfolios to become be more flexible than firms' existing plants, although the coefficient on Absolute Temperature Index is not statistically significantly different from zero in the specification presented in Column 3.

4.2. How does Climate Change Affect the Volatility of Electricity Demand?

We have documented that changes in climate appear to have a meaningful impact on the investments of electric utilities in new power plants. In particular, higher abnormal temperatures lead firms to increase their construction of new power plants, with the increase coming from plants that rely on relatively flexible production technologies. One potential reason for this pattern is that electricity demand fluctuates with the weather, so that climate change should lead to more volatile demand for electricity.

To evaluate the extent to which increased volatility of electricity demand is the reason why higher temperatures lead to increased construction of flexible power plants, we examine the underlying hypothesis of this argument: whether temperature increases do in fact increase the volatility of electricity demand. To do so, we focus on two implications of this argument that can be empirically verified. First, we investigate whether climate change leads to more extreme weather events. If the increase in average temperature is associated with more extraordinarily hot or cold days, demand for electricity is likely to deviate from normal patterns, making flexible generation assets valuable. Second, we utilize the fact in much of the world, electricity is sold on a wholesale market, whose prices reflect the short-term demand for electricity (see Lin, Schmid, and Weisbach (2017)).[20] If changing weather conditions lead to increased demand volatility, this higher demand volatility should be reflected in higher wholesale price volatility.

To define extreme weather events, we first estimate the average temperature in each month during the 1951 to 1980 period and calculate its standard deviation. Then we calculate the fraction of days

[20] For other regions without wholesale markets for electricity, we also expect that climate change leads to higher demand fluctuations. However, we cannot test the hypothesis in these regions because demand is not observable.

in each year during our sample period for which the temperature was above the historical average plus 2.5 times the standard deviation or below the historical average minus 2.5 times the standard deviation. If we assume that temperatures follow the normal distribution, we would expect that about 1.2 percent of all days are outside this range. The actual fraction of days outside this range, averaged across all sample regions, is shown in Figure 5. We see a clear trend that the fraction of extreme days increases, which indicates that the weather became more extreme over time. While only around one percent of all days are classified as extreme until the 1980s, this fraction increases to about two percent in the 1990s and above three percent in the 2000s.

However, while there is a trend toward increasing temperatures and also a trend to more extreme weather events, it is not clear if the two are related, and if the places in the world with higher increases in temperatures also tend to have more extreme events. In Panel A of Table 8, we estimate equations predicting how abnormal average temperatures in a specified region in a particular year affect the likelihood of extreme weather days. The estimates in this panel imply that a one-standard deviation increase of the abnormal temperature index leads to an increase in of the fraction of extreme days of slightly less than one percentage point. This result is consistent with the notion that climate change leads to more extreme weather.

To measure wholesale electiricity price volatility, we rely on a database of wholesale electricity prices. For each region with a wholesale market, we collect data on electricity prices from different sources (e.g., directly from the websites of the exchanges). We then calculate electricity price volatility as standard deviation of returns of hourly electricity prices in market and year. Returns are calculated as differences between hourly prices in U.S. dollar and standardized by the average price in a market.

In Panel B of Table 8, we estimate the extent to which abnormal temperatures lead to higher wholesale price volatility. Because neither electricity price volatility nor climate change is firm-specific, we estimate, at the region-year level, the extent to which weather volatility leads to wholesale price volatility. In these equations, we include electricity market fixed effects in all specifications because electricity price volatility depends heavily on the market design and is thus hardly comparable across markets. The estimates are consistent with the notion that climate change leads to more volatile electricity prices. They indicate that a one standard deviation increase in the Absolute Temperature Index leads to an increase of electricity price volatility of about 0.05. The mean value for electricity price volatility is about 0.4, so an increase of .05 represents a relative change of more than 10 percent. The results in Table 8 suggest that climate change has a strong impact on the occurrence of extreme weather and the volatility of electricity prices, which presumably comes from changes in the underlying demand for electricity.

5. Conclusion

The changing climate is potentially one of the most consequential phenomena in human history. Much attention has been focused on the way changing weather patterns affects ocean levels, the likelihood and violence of storms, and agricultural productivity. Yet, there are many other potential effects of climate change that could impact many aspects of the economy. Firms in a number of different industries will have to alter the way that they do business, sometimes in a substantial way. We study the effect of climate change on one industry that is likely to be considerably affected by it, the electricity producing industry.

A major factor in the demand for electricity is the weather. When there is a period of prolonged heat or cold, demand for electricity increases and electricity producers would like to increase their production. Much evidence has suggested that climate change has led to less predictable and more extreme weather events. The effect of climate change on investment is not clear. One possibility is that the uncertainty about future demand will lead firms to postpone or reduce investment. Alternatively, it is possible that electricity producers respond by increasing both their overall capacity for producing electricity, and the methods that they use, so that they can alter output at relatively low cost.

In this paper, we consider a sample of 384 electricity producing firms from 67 countries over the 2000-2015 period. These firms made investments in 13,867 new power plants during our sample period. We evaluate the extent to which changing weather conditions affected their investment decisions. Our analysis exploits both time series and cross-sectional variation in abnormal weather to identify this effect. The estimates indicate the quantity of new power plants that firms build increases in regions in which climate change is more severe. In addition, the estimates suggest that the incremental plants built because of weather are likely to be relatively flexible ones that can adjust output at low cost. These estimates imply that a one standard deviation increase in the abnormal temperature leads to a 10 to 20 percent relative increase in investment, suggesting that changes in weather have had a substantial impact on these firms' investment decisions.

These results are consistent with the view the effects of climate change have been incorporated into the investment decisions of electricity producing firms. Presumably, as the earth continues to warm and weather becomes even more extreme, firms will continue to favor flexible power plants for which output can be adjusted easily. Ironically, the type of plant most responsible for the CO_2 emissions that cause climate change is the coal-fired plant, which happens to be relatively inflexible, so adjusting output is relatively high cost. Consequently, because of climate change, firms appear to be shifting away from the coal fired plants, not because of their CO_2 emissions, but because of their inflexibility. Unfortunately, the weather induced shift has not be to renewable energy, although there has been an increase in renewables for other reasons. Instead, changing weather conditions have led firms to invest in relatively flexible plants such as gas fired ones, which do produce CO_2 emissions, although fewer than a coal fired one.

While investments in power plants are an important topic, we hope our paper makes a larger point: changing weather fundamentally changes the economics of many businesses. Our results suggest that it leads energy producing companies to increase investments to enhance their operating flexibility. It potentially leads firms to invest more in other industries as well to adjust to the changing conditions, but the impact of culminate change on the way firms in different industries invest is likely to vary substantially. Future research that characterizes the way in which climate change affects different industries is likely to be fruitful.

References

Abel, A. B. (1983) “Optimal investment under uncertainty”, *American Economic Review*, 73, 228-233/

Allianz (2013): The Weather Business – How Companies Can Protect Against Increasing Weather Volatility. *Industry report*.

Amiti, M. and Weinstein, D. E. (2017), „How Much do Idiosyncratic Bank Shocks Affect Investment? Evidence from Matched Bank-Firm Loan Data”, *Journal of Political Economy*, forthcoming.

Becher, D. A., J. H. Mulherin, and R. A. Walkling (2012) “Sources of Gains in Corporate Mergers: Refined Tests from a Neglected Industry”, *Journal of Financial and Quantitative Analysis* 47, 57–89.

Billet, M. T., J. A. Garfinkel and Y. Jiang (2011) „The influence of governance on investment: Evidence from a hazard model”, *Journal of Financial Economics*, 102, 643-670.

Bloom, N. (2009) “The impact of uncertainty shocks”, *Econometrica*, 77, 623–685.

Bolton, P., H. Chen, and N. Wang (2013) “Market timing, investment, and risk management”, *Journal of Financial Economics,* 109, 40-62.

Dixit, A. K. (1980) “The role of investment in entry-deterrence”, *Economic Journal*, 90, 95-106.

Dixit, A. K. and Pindyck, R. S. (1994) “Investment under Uncertainty”, *Princeton University Press*, Princeton

Fabrizio, K., N. Rose, and C. Wolfram (2007) “Do Markets Reduce Costs? Assessing the Impact of Regulatory Restructuring on US Electric Generation Efficiency,” *American Economic Review* 97, 1250–77.

Frésard, L. and Valta, P. (2016) “How Does Corporate Investment Respond to Increased Entry Threat?”, *Review of Corporate Finance Studies*, 1, 1-35.

Gan, J. (2007a) “Collateral, debt capacity, and corporate investment: Evidence from a natural experiment”, *Journal of Financial Economics*, 85, 709-734.

Gan, J. (2007b) “The Real Effects of Asset Market Bubbles: Loan- and Firm-Level Evidence of a Lending Channel”, *Review of Financial Studies*, 20, 1941–1973.

Gilje, E. P. and Taillard, J. (2016) “Do Private Firms Invest Differently than Public Firms? Taking Cues from the Natural Gas Industry”, *Journal of Finance*, 71, 1733-1778.

Greenwood, R. and Hanson, S. G. (2014) “Waves in Ship Prices and Investment”, *Quarterly Journal of Economics*, 130, 55-109.

Gulen, H. and Ion, M. (2015) “Policy Uncertainty and Corporate Investment”, *Review of Financial Studies*, 29, 523-564.

Hansen, J., Sato, M., Glascoe, J., and Ruedy, R. (1998) “A common-sense climate index: Is climate changing noticeably?” *Proceedings of the National Academy of Sciences* 95 (8), 4113–4120.

Hansen, J., Sato, M., and Ruedy, R. (2012) “Perception of climate change,” *Proceedings of the National Academy of Sciences* 109 (37), 241-2423.

Harford, J. and V.B. Uysal (2014) “Bond market access and investment”, *Journal of Financial Economics*, 112, 147-163.

Hunt, S. (2002) “Making competition work in electricity”, *John Wiley & Sons*, New York.

Julio, B. and Yook, Y. (2012) “Political Uncertainty and Corporate Investment Cycles”, *Journal of Finance,* 67, 45-83.

Kellog, R. (2014) “The Effect of Uncertainty on Investment: Evidence from Texas Oil Drilling”, American Economic Review, 104, 1698-1734.

Lin, Chen, Thomas Schmid, and Michael S. Weisbach (2017) “Price Risk, Production Flexibility, and Liquidity Management: Evidence from Electricity Generating Firms,” *NBER Working Paper No. 23434* .

Myers, Robert (2008) “What Every CFO Needs to Know About Weather Risk Management,” CME Group and Storm Exchange Inc.

Pérez González, Francisco and Hayong Yun (2013) “Risk Management and Firm Value: Evidence from Weather Derivatives,” *Journal of Finance*, 68, 2143-2176.

Reinartz, Sebastian and Thomas Schmid (2016) “Production Flexibility, Product Markets, and Capital Structure Decisions,” *Review of Financial Studies*, 29, 1501-1548.

Rettl, Daniel A., Alex Stomper, and Josef Zechner (2016) “The Stability of Dividends and Wages: Effects of Competitor Inflexibility,” *Working Paper*.

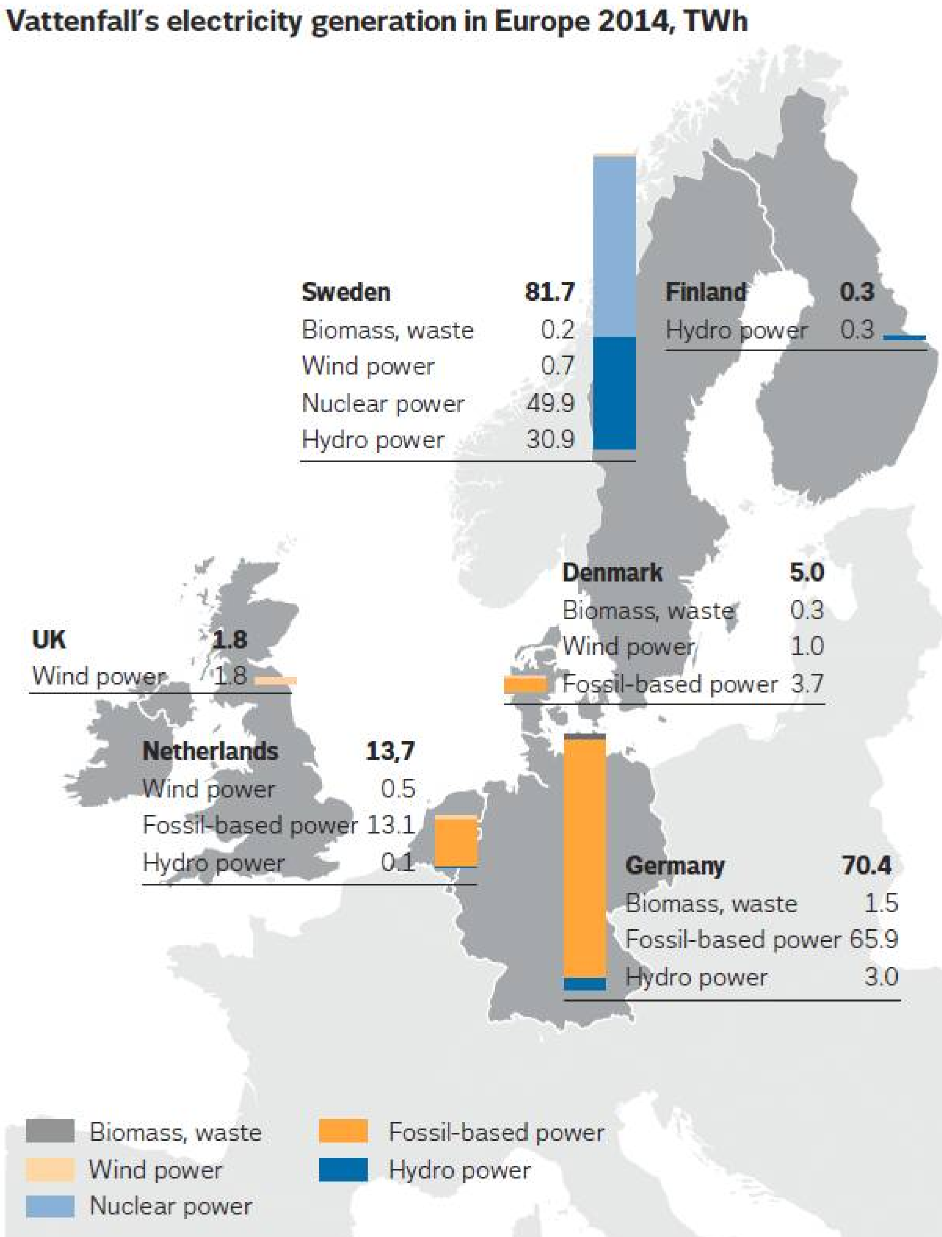
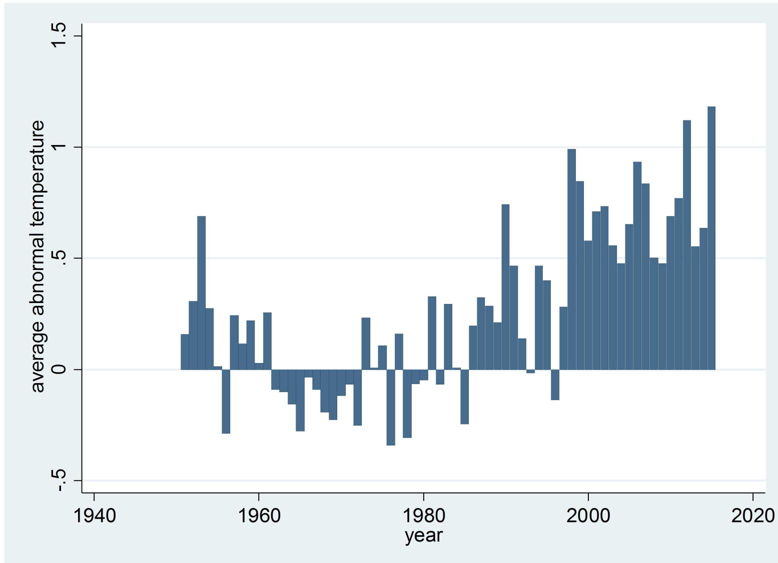
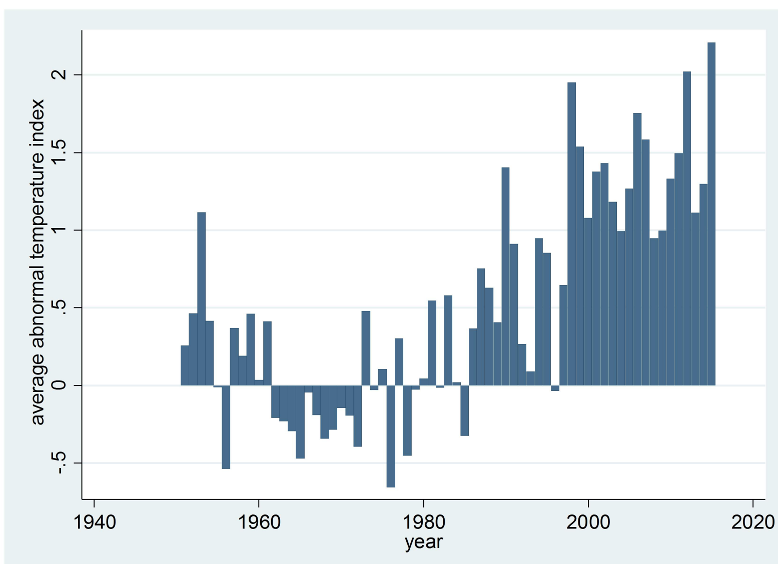


Figure 1: To illustrate our data structure, this figure shows the main countries in which Vattenfall AB owns production capacity (Source: Vattenfall annual report 2014). In our dataset, we would observe Vattenfall's existing production capacity as well as planned new power plants individually for each country (and year).



(a) Abnormal Temperature



(b) Abnormal Temperature Index

Figure 2: This figure shows the development of the average abnormal temperature and the average abnormal temperature index over time. Abnormal temperature for a region (i.e., country or state for U.S., Canada, and Australia) and year is measured relative to the base period 1951 to 1980. The abnormal temperature index is constructed as abnormal temperature divided by the interannual standard deviation during the base period 1951 to 1980 in the same region. More details about its construction can be found in Appendix A.

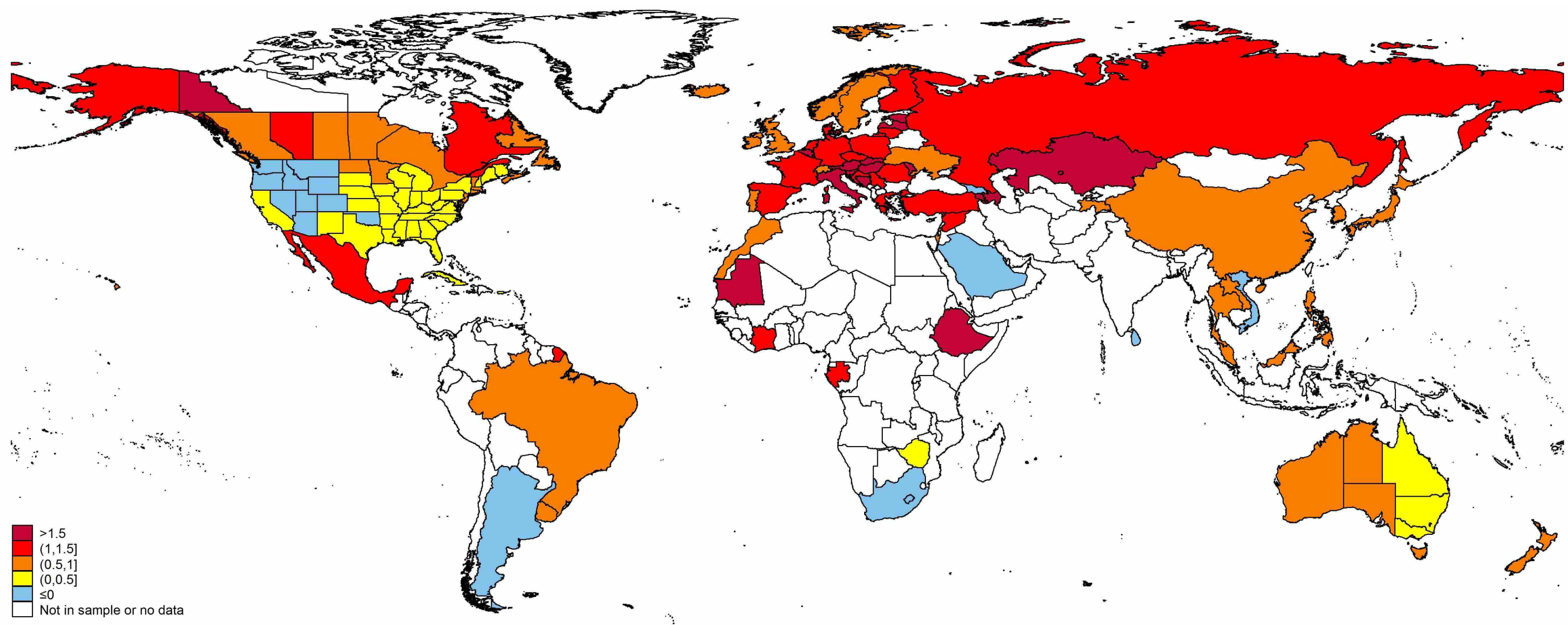


Figure 3: This figure shows the average abnormal temperature during our sample period 2000 to 2015. Abnormal temperature for a region (i.e., country or state for U.S., Canada, and Australia) and year is measured relative to the base period 1951 to 1980. More details about its construction can be found in Appendix A.

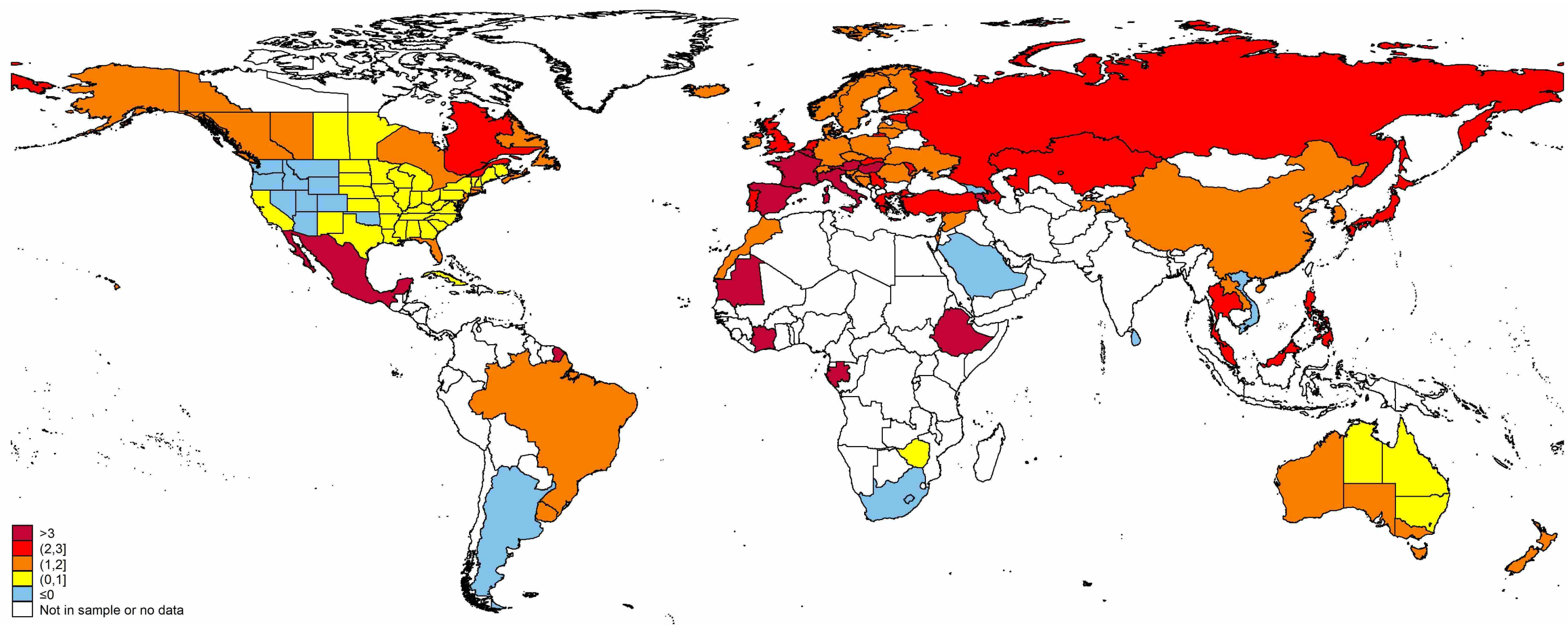


Figure 4: This figure shows the average abnormal temperature index during our sample period 2000 to 2015. The abnormal temperature index is constructed as abnormal temperature divided by the interannual standard deviation during the base period 1951 to 1980 in a region (i.e., country or state for U.S., Canada, and Australia). More details about its construction can be found in Appendix A.

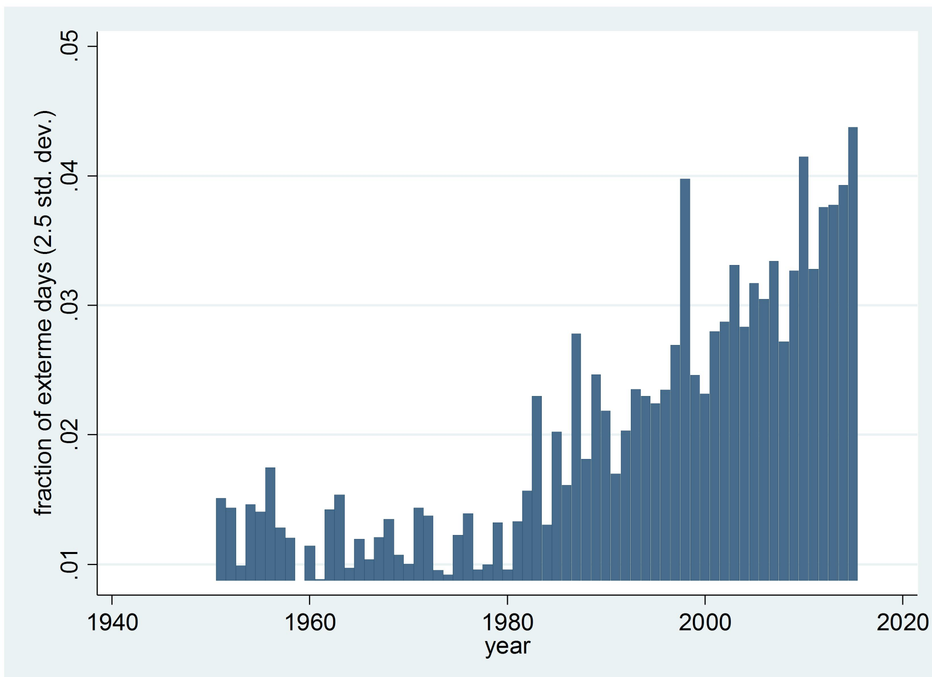


Figure 5: This figure shows the development of the yearly fraction of extreme days over time. A day is defined to be extreme if the average temperature on that day is higher (lower) than the average temperature in the corresponding months during the base period 1951 to 1980 plus (minus) 2.5 times the standard deviation of temperatures during that month in the base period. More details about its construction can be found in Appendix A.

Table 1: Descriptive statistics: Early-stage power plant projects

Technology	total number	total GW	capacity (MW) average	capacity (MW) median	countries number
Flexible	4,551	1,048	199	124	96
Oil	627	70	112	19	58
Gas	1,841	262	142	83	108
Gas comb. cycle	2,083	716	344	270	121
Inflexible	1,974	1,167	807	850	59
Coal	1,819	1,003	551	600	87
Nuclear	155	165	1,062	1,100	30
Renewable	5,999	670	93	37	92
Solar	635	25	40	13	69
Hydro	2,710	388	143	48	89
Wind	2,654	256	97	50	119
Other	1,352	172	127	70	64
Total	13,876	3,057	291	250	83

This table presents descriptive statistics for the early-stage investment projects. Each project is only considered in the first year in which it appears in the database. Reported are the total number of power plant projects, the total capacity of planned plants in gigawatt, the average and median capacity of planned plants, and the number of regions (i.e., countries or states for the U.S., Canada, and Australia) in which any project takes place.

Table 2: Descriptive statistics: Firms

Variable	**Obs**	**Mean**	**p25**	**p50**	**p75**	**SD**
Investment variables						
Total investment	12,767	0.16	0.00	0.00	0.14	0.31
Flexible investment	12,767	0.05	0.00	0.00	0.00	0.18
Inflexible investment	12,767	0.03	0.00	0.00	0.00	0.13
Renewable investment	12776	0.08	0.00	0.00	0.00	0.24
Flexible to total	4,587	0.34	0.00	0.00	0.81	0.42
Inflexible to total	4,587	0.15	0.00	0.00	0.00	0.32
Renewable to total	4587	0.42	0.00	0.08	1.00	0.47
Δ $\text{RuT}_{new-projects}$	3,400	-0.23	-0.92	-0.39	0.19	0.72
Δ $\text{RuT}_{new-portfolio}$	10,026	0.01	0.00	0.00	0.00	0.24
Weather variables						
Abnormal temp. index	12,767	1.20	0.09	1.12	2.21	1.67
Abn. temp. index (3y avg)	12,105	1.18	0.27	1.07	2.05	1.39
Abnormal temp.	12,767	0.60	0.05	0.61	1.13	0.86
Abnormal temp. (3y avg)	12,105	0.59	0.16	0.59	1.02	0.68
Extreme $\text{days}_{2.5std.dev.}$	12,776	0.024	0.005	0.014	0.027	0.037
Control variables						
Log(assets)	12,767	16.16	14.92	16.71	17.66	1.95
Profitability	12,713	0.05	0.04	0.05	0.07	0.07
Tobin's Q	12,121	1.17	1.02	1.11	1.26	0.40
Leverage	12,767	0.55	0.45	0.56	0.66	0.18
Cash holdings	12,767	0.07	0.02	0.04	0.10	0.07
$\text{Log(GDP/capita)}_{HQ}$	12,141	10.36	10.49	10.77	10.81	0.85

This table presents descriptive statistics. Reported are the number of observations (N), mean value, 25% percentile, median, 75% percentile, and standard deviation (SD). A detailed description of all variables can be found in Appendix A.

Table 3: Climate change and investments in power plants

Column	1	2	3	4	5
Abn. temp. index	**0.021***** **(6.22)**	**0.017***** **(3.88)**	**0.0088***** **(3.16)**	**0.0093***** **(3.11)**	**0.012***** **(2.95)**
Log(assets)		0.0051		0.033**	n/a
		(0.80)		(2.19)	
Profitability		-0.16		-0.0066	n/a
		(-1.75)		(-0.14)	
Tobin's Q		-0.0068		0.024*	n/a
		(-0.26)		(1.96)	
Leverage		-0.066		-0.053	n/a
		(-1.48)		(-1.13)	
Cash holdings		0.38***		0.23***	n/a
		(3.43)		(4.46)	
Log(GDP/capita)$_{HQ}$		-0.059***		-0.086	n/a
		(-3.15)		(-0.83)	
Year-FE	yes	yes	yes	yes	yes
Firm-FE	no	no	yes	yes	yes
Firm x Year FE	no	no	no	no	yes
Observations	11,637	10,897	11,613	10,876	9,474
Firms	384	357	360	336	163
R2	0.017	0.046	0.26	0.26	0.25

The dependent variable is TOTAL INVESTMENT, which is defined as early-stage power plant construction projects (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). The abnormal temperature index is constructed as abnormal temperature in a year and region, relative to the base period 1951 to 1980, divided by the interannual standard deviation during the same period. All variables are lagged by one year. T-statistics based on robust standard errors clustered by countries/firms/years are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 4: Robustness: measurement of climate change

Column	1	2	3	4	5	6
ATI_{3y}	**0.014*** **	**0.014*** **				
	(3.13)	**(3.51)**				
Temp. index			**0.023*** **	**0.031*** **		
			(3.63)	**(4.63)**		
Temp. index_{3y}					**0.033*** **	**0.034*** **
					(3.31)	**(4.01)**
Log(assets)	0.037**	n/a	0.032**	n/a	0.037**	n/a
	(2.38)		(2.15)		(2.38)	
Profitability	0.0039	n/a	-0.0070	n/a	0.0047	n/a
	(0.089)		(-0.16)		(0.10)	
Tobin's Q	0.027**	n/a	0.023*	n/a	0.027**	n/a
	(2.42)		(1.98)		(2.44)	
Leverage	-0.058	n/a	-0.051	n/a	-0.058	n/a
	(-1.25)		(-1.09)		(-1.27)	
Cash holdings	0.25***	n/a	0.22***	n/a	0.25***	n/a
	(4.95)		(4.54)		(5.00)	
Log(GDP/capita)$_{HQ}$	-0.049	n/a	-0.081	n/a	-0.045	n/a
	(-0.41)		(-0.79)		(-0.38)	
Year-FE	yes	yes	yes	yes	yes	yes
Firm-FE	yes	yes	yes	yes	yes	yes
Firm x Year FE	no	yes	no	yes	no	yes
Observations	10,363	10,038	10,922	10,500	10,363	10,038
Firms	311	163	337	170	311	163
R2	0.25	0.24	0.26	0.24	0.25	0.24

ATI stands for abnormal temperature index. The dependent variable is TOTAL INVESTMENT, which is defined as early-stage power plant construction projects (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). All variables are lagged by one year. T-statistics based on robust standard errors clustered by countries/firms/years are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 5: Robustness test: measurement of investment

Column	1 dummy	2 log(MW)	3 # to total	4 log(#)
Abn. temp. index	**0.10*** **	**0.15** **	**0.016*** **	**0.093*** **
	(5.29)	**(2.87)**	**(2.96)**	**(3.14)**
Log(assets)	0.061**	0.30**	0.0083	0.052
	(2.19)	(2.31)	(0.57)	(1.56)
Profitability	-0.54*	0.29	0.015	0.053
	(-1.69)	(0.69)	(0.52)	(0.81)
Tobin's Q	0.020	0.28**	0.025	0.058*
	(0.27)	(2.60)	(1.37)	(1.81)
Leverage	-0.43***	-0.47	-0.071	-0.15
	(-2.60)	(-1.59)	(-1.68)	(-1.35)
Cash holdings	0.58*	0.48	0.21***	0.45**
	(1.65)	(0.89)	(3.06)	(2.23)
Log(GDP/capita)$_{HQ}$	-0.26***	0.71	-0.16**	0.052
	(-3.57)	(1.44)	(-2.72)	(0.36)
Year-FE	yes	yes	yes	yes
Firm-FE	yes	yes	yes	yes
Firm x Year FE	no	yes	no	yes
Observations	11,747	11,728	10,876	10,876
Firms	363	344	336	336
R2	0.041	0.30	0.22	0.26

The dependent variable is indicated in each column. The dummy equals one if there is any early stage power plant project and zero otherwise. Ln MW is the natural logarithm of all early stage power plant projects (not scaled). # stands for number of power plants projects. In column 3, the number is scaled by the number of total plants of firm i in market j and year t. In column 4, the natural logarithm of this number is used (not scaled). All variables are lagged by one year. T-statistics based on robust standard errors clustered by countries/firms/years are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 6: Robustness: regional characteristics

Column	1	2	3	4	5	6	7
	macro controls		micro controls		both		+FE
ATI	**0.0084****	**0.011****	**0.010*****	**0.011****	**0.0094****	**0.011****	**0.0053****
	(2.71)	**(2.80)**	**(3.11)**	**(2.78)**	**(2.52)**	**(2.73)**	**(2.62)**
Log(assets)	0.030*	n/a	0.027*	n/a	0.032*	n/a	0.028
	(1.78)		(1.86)		(1.93)		(1.60)
Profitability	-0.014	n/a	-0.0014	n/a	-0.012	n/a	0.0086
	(-0.28)		(-0.036)		(-0.26)		(0.23)
Tobin's Q	0.024*	n/a	0.013	n/a	0.016	n/a	0.012
	(1.93)		(1.13)		(1.32)		(0.93)
Leverage	-0.043	n/a	-0.040	n/a	-0.037	n/a	-0.0025
	(-0.87)		(-0.88)		(-0.83)		(-0.050)
Cash holdings	0.21***	n/a	0.23***	n/a	0.21***	n/a	0.18***
	(3.76)		(5.16)		(4.27)		(3.37)
$\text{Log}(\frac{GDP}{capita})_{region}$	-0.059**	-0.051**			-0.054**	-0.045**	-0.028
	(-2.73)	(-2.65)			(-2.80)	(-2.43)	(-0.17)
Δ GDP_{region}	0.00073	0.0084			0.0014	0.0085	-0.0038
	(0.24)	(1.03)			(0.33)	(1.03)	(-1.14)
$\text{Inflation}_{region}$	-0.35	-0.42			-0.39	-0.50	-0.29
	(-0.88)	(-0.79)			(-0.99)	(-0.96)	(-1.03)
$\text{Flexibility}_{overall}$			-0.0065	n/a	-0.025	n/a	-0.020
			(-0.100)		(-0.40)		(-0.33)
$\text{Flexibility}_{region}$			-0.043	-0.045	-0.042	-0.046*	-0.018
			(-1.63)	(-1.76)	(-1.73)	(-1.87)	(-0.72)
Log(MW)_{region}			0.0043	0.0082*	0.0031	0.0057	-0.0048
			(0.86)	(1.87)	(0.62)	(1.22)	(-0.71)
$\frac{MW_{region}}{MW_{overall}}$			-0.016	-0.050*	-0.0022	-0.024	0.00018
			(-0.63)	(-1.79)	(-0.083)	(-0.81)	(0.0056)
Year-FE	yes	yes	yes	yes	yes	yes	yes
Firm-FE	yes	yes	yes	yes	yes	yes	yes
Firm x Year FE	no	yes	no	yes	no	yes	no
Firm x Market FE	no	no	no	no	no	no	yes
Observations	10,406	8,510	10,439	8,644	9,937	8,108	9,848
Firms	338	149	338	151	336	145	333
R2	0.27	0.26	0.26	0.24	0.27	0.25	0.58

continued on next page

Table 6 continued

The dependent variable is TOTAL INVESTMENT, which is defined as early-stage power plant construction projects (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). The abnormal temperature index is constructed as abnormal temperature in a year and region, relative to the base period 1951 to 1980, divided by the interannual standard deviation during the same period. All variables are lagged by one year. T-statistics based on robust standard errors clustered by countries/firms/years are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 7: In which types of power plants do firms invest?

Panel A: absolute investment levels						
Column	1	2	3	4	5	6
	flexible investment		inflexible investment		renewable investment	
Abn. temp. index	**0.0090****	**0.011*****	**0.00092**	**0.00052**	**0.0018**	**0.0029**
	(2.93)	**(3.14)**	**(0.80)**	**(0.35)**	**(0.58)**	**(0.74)**
Log(assets)	0.0039	n/a	0.014	n/a	0.023**	n/a
	(0.42)		(1.45)		(2.24)	
Profitability	-0.00089	n/a	0.043***	n/a	-0.031	n/a
	(-0.025)		(6.12)		(-0.92)	
Tobins Q	0.014	n/a	0.0096*	n/a	0.0058	n/a
	(1.60)		(1.76)		(1.04)	
Leverage	-0.019	n/a	-0.034	n/a	0.0017	n/a
	(-0.78)		(-0.99)		(0.054)	
Cash holdings	0.11**	n/a	0.032	n/a	0.14**	n/a
	(2.61)		(1.45)		(2.32)	
Log(GDP/capita)$_{HQ}$	0.081**	n/a	-0.052	n/a	-0.099**	n/a
	(2.70)		(-0.65)		(-2.34)	
Year-FE	yes	yes	yes	yes	yes	yes
Firm-FE	yes	yes	yes	yes	yes	yes
Firm x Year FE	no	yes	no	yes	no	yes
Observations	10,876	9,474	10,876	9,474	10,881	9,474
Firms	336	163	336	163	337	163
R2	0.16	0.21	0.37	0.28	0.29	0.25

continued on next page

Table 7 continued

Panel B: relative investment levels (only investing firms)						
Column	1	2	3	4	5	6
	flexible investment		inflexible investment		renewable investment	
Abn. temp. index	**0.021****	**0.027****	**-0.0038**	**-0.0048**	**-0.013**	**-0.015**
	(2.35)	**(2.62)**	**(-0.58)**	**(-0.64)**	**(-1.25)**	**(-1.15)**
Log(assets)	-0.015	n/a	-0.00039	n/a	0.032*	n/a
	(-0.52)		(-0.019)		(1.81)	
Profitability	-0.0022	n/a	-0.013	n/a	-0.064	n/a
	(-0.027)		(-0.17)		(-0.70)	
Tobins Q	0.011	n/a	0.00055	n/a	-0.016	n/a
	(0.52)		(0.033)		(-0.67)	
Leverage	-0.0070	n/a	0.0092	n/a	0.015	n/a
	(-0.10)		(0.17)		(0.21)	
Cash holdings	0.11	n/a	-0.048	n/a	0.029	n/a
	(0.60)		(-0.83)		(0.19)	
Log(GDP/capita)$_{HQ}$	0.42***	n/a	-0.11	n/a	-0.30***	n/a
	(4.62)		(-1.63)		(-3.01)	
Year-FE	yes	yes	yes	yes	yes	yes
Firm-FE	yes	yes	yes	yes	yes	yes
Firm x Year FE	no	yes	no	yes	no	yes
Observations	4,889	3,616	4,889	3,616	4,897	3,616
Firms	256	120	256	120	257	120
R2	0.40	0.31	0.40	0.097	0.51	0.51

continued on next page

Table 7 continued

Panel C: change of firms flexibility (run-up time, RuT)				
Column	1	2	3	4
	Δ RuT$_{new-projects}$		Δ RuT$_{new-portfolio}$	
Abn. temp. index	**-0.021****	**-0.058*****	**-0.0036**	**-0.0074***
	(-2.31)	**(-3.00)**	**(-0.87)**	**(-2.03)**
Log(assets)	0.0045	n/a	-0.011	n/a
	(0.057)		(-0.53)	
Profitability	-0.17	n/a	0.066*	n/a
	(-0.57)		(1.85)	
Tobin's Q	0.061	n/a	0.0073	n/a
	(1.46)		(0.73)	
Leverage	-0.11	n/a	-0.014	n/a
	(-0.79)		(-0.31)	
Cash holdings	-0.23	n/a	-0.021	n/a
	(-0.84)		(-0.32)	
Log(GDP/capita)$_{HQ}$	0.38*	n/a	0.089**	n/a
	(2.00)		(2.40)	
Year-FE	yes	yes	yes	yes
Firm-FE	yes	yes	yes	yes
Firm x Year FE	no	yes	no	yes
Observations	3,162	1,766	8,722	7,256
Firms	229	73	314	136
R2	0.33	0.28	0.21	0.18

continued on next page

Table 7 continued

Gas, gas combined-cycle, and oil power plants are considered as flexible plants. Coal and nuclear power plants are considered as inflexible plants. Run-up time combines the (in)flexibility of all types of plants into one measure; it's defined as average time which is necessary to start-up a power plant in hours. Higher values of run-up time go along with lower flexibility.

Panel A: The dependent variables indicate absolute investments in power plant construction projects. FLEXIBLE investment is defined as early-stage projects to construct flexible power plants (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). INFLEXIBLE and RENEWABLES are defined in the same way.

Panel B: The dependent variables indicate relative investments in power plant construction projects. They are calculated as flexible (inflexible, reneables) power plant projects (in MW) divided by all early-stage power plant construction projects (in MW).

Panel C: The dependent variables indicate the change of the run-up time as measure for inflexibility due to the early stage power plant projects. In columns 1 and 2, the run-up time of early stage plant projects is compared to all existing plants' run-up time in t-1. In columns 3 and 4, the hypothetical run-up time of the new power plant portfolio (including existing and early stage plants) is compared to the existing plants' run-up time in t-1.

Abnormal temperature index is constructed as abnormal temperature in a year and region, relative to the base period 1951 to 1980, divided by the interannual standard deviation during the same period. All variables are lagged by one year. T-statistics based on robust standard errors clustered by countries/firms/years are presented in parentheses. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Table 8: Channel: extreme weather, electricity price volatility, and investments

Panel A: Climate change and weather extremes (region-year analysis)				
Dependent variable: extreme days$_{2.5 std.dev.}$				
Column	1	2	3	4
Abn. temp. index	**0.0059*** **	**0.0052*** **	**0.0046*** **	**0.0046*** **
	(4.82)	**(4.31)**	**(3.32)**	**(3.32)**
Log(GDP/capita)$_{reg}$			0.018	0.020
			(1.03)	(1.24)
Δ GDP$_{region}$				-0.013
				(-0.60)
Inflation$_{region}$				-6.3e-06
				(-0.016)
Year-FE				
Market-FE				
Observations	2,524	2,524	1,996	1,994
Regions	174	174	120	120
R2	0.78	0.79	0.76	0.76
Panel B: Climate change and el. price volatility (region-year analysis)				
Dependent variable: electricity price volatility				
Column	1	2	3	4
Abn. temp. index	**0.025*** **	**0.032** **	**0.033** **	**0.031** **
	(2.77)	**(2.37)**	**(2.37)**	**(2.46)**
Log(GDP/capita)$_{reg}$			1.16*	0.85*
			(1.90)	(1.72)
Δ GDP$_{region}$				2.17*
				(1.96)
Inflation$_{region}$				0.027**
				(2.29)
Year-FE	no	yes	yes	yes
Market-FE	yes	yes	yes	yes
Observations	801	801	774	774
Regions	65	65	62	62
R2	0.76	0.77	0.78	0.78

continued on next page

Table 8 continued

The dependent variable is EXTREME DAYS$_{2.5 std.dev.}$ in Panel A and EL. PRICE VOLATILITY (which is defined as standard deviation of returns of hourly electricity prices in market m and year t) in Panel B. The analyses are done on the region-year level both Panels because the variables of interest do not vary across firms. T-statistics based on robust standard errors are presented in parentheses. The standard errors are clustered by regions in Panels A and B. ***, ** and * indicate significance on the 1%-, 5%- and 10%-levels, respectively. A detailed description of all variables can be found in Appendix A.

Appendix

Appendix A: Definition of variables

Variable	**Description**
General definitions	
Region$_j$	Country or state (for the U.S., Canada, and Australia).
Market$_m$	Wholesale market region for electricity. Typically the market regions equals a country, but the U.S., Canada, and Australia have multiple markets which cover only particular regions (i.e., states). Another exception is Nordpool, which is the common market region for several Northern European countries.
Flexible plant	Gas, gas combined-cycle, and oil power plants are considered as flexible plants.
Inflexible plant	Coal and nuclear power plants are considered as inflexible plants.
Investment planning variables (Source: Own calculations based on WEPP database)	
Total investment$_{i,j,t}$	Early-stage power plant construction projects (in megawatt, MW) of firm i in region j and year t, scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j and year t. The variable is set to one for values $>$ one and to zero if firm i in region j (and year t) has existing production capacity of at least 1 MW but no planned flexible investment.
Flexible investment$_{i,j,t}$	Early-stage flexible power plant construction projects (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). The same adjustments as for TOTAL INVESTMENT are made.
Inflexible inv.$_{i,j,t}$	Early-stage inflexible power plant construction projects (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). Inflexible power plants are coal and nuclear power plants. The same adjustments as for TOTAL INVESTMENT are made.
Renewable inv.$_{i,j,t}$	Early-stage inflexible power plant construction projects (in megawatt, MW) of firm i in region j (and year t), scaled by the capacity of existing power plants (in MW) of the same firm i in the same region j (and year t). Renewables are hydro, wind, and solar power plants. The same adjustments as for TOTAL INVESTMENT are made.
Dummy$_{i,j,t}$	Dummy variable which equals one if total investment is greater than zero and zero if total investment equals zero.
log(MW)$_{i,j,t}$	Natural logarithm of one plus the total capacity of power plant projects of firm i in region j (and year t). Set to zero if firm i in region j (and year t) has existing production capacity of at least 1 MW but no power plant projects. Source: Own calculations based on WEPP database.
# to total$_{i,j,t}$	Number of flexible power plant projects of firm i in region j (and year t), scaled by the number of existing power plants of the same firm i in the same region j (and year t). Set to zero if firm i in region j (and year t) has existing production capacity of at least 1 MW but no power plant projects.

Definition of Variables - continued

Variable	Description
$\log(\#)_{i,j,t}$	Natural logarithm of one plus the number of power plant projects of firm i in region j (and year t). This measure is not scaled. The variable is set to one for values > one and to zero if firm i in region j (and year t) has existing production capacity of at least 1 MW but no power plant projects.
Flexible to total$_{i,j,t}$	Flexible power plant projects (in MW) divided by all early-stage power plant construction projects (in MW) of firm i in region j (and year t).
Inflexible to total$_{i,j,t}$	Inflexible power plant projects (in MW) divided by all early-stage power plant construction projects (in MW) of firm i in region j (and year t).
Renewable to total$_{i,j,t}$	Renewable power plant projects (in MW) divided by all early-stage power plant construction projects (in MW) of firm i in region j (and year t).
Δ RuT$_{new-projects}$	Relative difference in run-up time between early-stage power plant projects of firm i in region j and year t to existing power plants of firm i in region j and year $t-1$. Run-up time is the capacity-weighted average time which is necessary to start-up the power plants in hours. It is based on the production technologies of the firms' power plants. See Reinartz and Schmid (2016) for technology-specific values.
Δ RuT$_{new-portfolio}$	Relative difference in run-up time between the hypothetical new power plant portfolios (consisting of existing plants and early-stage plant projects) of firm i in region j and year t to existing power plants of firm i in region j and year $t-1$. Run-up time is the capacity-weighted average time which is necessary to start-up the power plants in hours. It is based on the production technologies of the firms' power plants. See Reinartz and Schmid (2016) for technology-specific values.
Weather variables (Source: Own calculations based on GHCN data)	
Abnormal temp. index (ATI)$_{j,t}$	Main measure for climate change. The abnormal temperature index in region j and year t is defined as abnormal temperature$_{j,t}$ (see below) divided by the interannual standard deviation during the base period 1951 to 1980 in the same region (see Hansen et al. (1998)). Thus, a value of one indicates that the temperature in this region is one standard deviation higher in the specific year compared to the average temperature during the base period.
Abnormal temp$_{j,t}$	Average temperature in a region j and year t minus the expected temperature in the same region. The expected temperature is calculated as the average temperature in the base period 1951 to 1980 in the same region (see, for instance, Hansen et al. (2012)).
Extreme days$_{2.5 std.dev.}$	Fraction of days in region j and year t which are extreme. A day is defined to be extreme if the average temperature on that day is higher (lower) than the average temperature in the corresponding months during the base period 1951 to 1980 plus (minus) 2.5 times the standard deviation of temperatures during that month in the base period.
Other variables	
Log(assets)$_{i,t}$	Logarithm of total assets [wc02999] in U.S. dollar.
Profitabiltiy$_{i,t}$	Earnings before interest, taxes, depreciation, and amortization (EBITDA) [wc18198] / total assets [wc02999].

Definition of Variables - continued

Variable	Description
Tobin's Q $_{i,t}$	Market capitalization [wc08001] plus total debt [wc03255] divided by book value of common equity [wc03501] plus total debt [wc03255].
Leverage$_{i,t}$	Total debt [wc03255] / (Total debt [wc03255] + book value of common equity [wc03501]).
Cash$_{i,t}$	Cash & short term investments [wc02001] / total assets [wc02999].
Log(GDP/capita)$_{HQ\|i,t}$	Natural logarithm of GDP per capita (in 2010 U.S. dollar) in year t in the headquarter country of firm i. Source: Worldbank.
Log(GDP/capita)$_{j,t}$	Natural logarithm of GDP per capita (in 2010 U.S. dollar) in region j and year t. Source: Worldbank.
Δ GDP$_{j,t}$	Change of GDP per capita in region j between year $t-1$ and year t. Source: Worldbank.
Inflation$_{j,t}$	Inflation rate in region j and year t. Source: Worldbank.
Flexibility$_{overall}$	Fraction of flexible plants (in MW) of firm i in year t. Source: Own calculations based on WEPP database.
Flexibility$_{region}$	Fraction of flexible plants (in MW) of firm i in region j and year t. Source: Own calculations based on WEPP database.
Log(MW)$_{region}$	Natural logarithm of the total capacity of all power plants of firm i in region j and year t. Source: Own calculations based on WEPP database.
$\frac{MW_{region}}{MW_{overall}}$	Fraction of total capacity of power plants of firm i in region j and year t to the total capacity of all power plants of firm i in year t. Source: Own calculations based on WEPP database.
El. Price Volatility$_{m,t}$	Volatility of electricity prices. Defined as standard deviation of returns of hourly electricity prices in market m and year t. Returns are calculated as differences between hourly prices in U.S. dollar and standardized by the average price in a market.

(Debt) Overhang: Evidence from Resource Extraction*

Michael D. Wittry
University of Washington

May 16, 2019
Link to latest version

Abstract

I study the empirical importance of debt overhang using a unique dataset on resource extraction firms, which provides ex ante measures of investment opportunities and important variation in the terms of a firm's obligations. In particular, unsecured reclamation liabilities create overhang that is costly to resolve and induces firms to forgo and postpone positive NPV investments. Traditional debt, in contrast, imposes few overhang-related investment distortions. These results show that: (i) the overhang problem is potentially large and applies more broadly to a firm's non-debt liabilities; and (ii) overhang problems associated with traditional debt can be avoided through contracting and debt composition.

JEL classification: D22, G30, G32

Keywords: Debt overhang, underinvestment, reclamation liability, general liability overhang

*University of Washington, e-mail: mwittry@uw.edu. I am grateful to my dissertation committee: Ran Duchin, Jarrad Harford, Ed Rice, Jake Vigdor, and especially my chair, Jon Karpoff, for continued guidance and support. I thank Philip Bond, Jonathan Brogaard, Aaron Burt, Alvin Chen, Quentin Dupont, Laura Field, Shan He (discussant), Jennifer Koski, Tarun Patel, Stephan Siegel, and Mark Westerfield, and seminar participants at Harvard Business School, Dartmouth College, Rice University, University of California San Diego, University of Colorado Boulder, Ohio State University, Arizona State University, University of British Columbia, University of Notre Dame, University of Texas Dallas, University of Pittsburgh, University of Kansas, Tulane University, University of Washington, and the Colorado Finance Summit for helpful comments. I also thank Tera Thast and Christina Bowman at Mining Intelligence for answering my (many) inqueries on the the nature of the data. I gratefully acknowledge the financial support of the JFQA, Verus Financial, and Russell Investments. Peter Heglund-Lohman, Jiarun Zhang, Ruyue Feng, Devin Reimers and Nerissa Morrow provided excellence research assistance.

1 Introduction

Debt overhang is a pillar of corporate finance theory. Myers (1977) demonstrates that existing debt obligations have the potential to induce underinvestment in valuable growth options, as the benefit of investing in such projects primarily accrues to debtholders. Establishing the importance of debt overhang in capital structure decisions is difficult, however, because contracting and debt composition mechanisms endogenously arise to mitigate its effects. In addition, fully identifying the costs of overhang requires observing the firm's investment opportunity set. This paper exploits a novel setting that allows me to both disentangle the costs of overhang from its potential solutions, and observe the firm's ex ante investment opportunity set.

I focus on a sample of Canadian resource extraction firms that provides ex ante estimates of the net present value (NPV) of a firm's new mining projects. In December 2000, the Ontario Securities Commission passed regulation that significantly increased disclosure requirements for publicly listed mining firms, notably requiring firms to file technical reports on mining projects that include an estimate of the project's NPV. These feasibility reports enable me to see exactly when firms take positive NPV projects, and more importantly, when they forgo or postpone them.

The resource extraction setting also allows me to directly compare two types of liabilities, traditional debt and mine reclamation liabilities, which differ in the costs associated with avoiding overhang.[1] The overhang-related investment distortions induced by traditional debt can be mitigated at relatively low cost by firms avoiding debt altogether, issuing short maturity debt, renegotiating the contract ex post (Myers, 1977), or securing their debt issuances (Stulz and Johnson, 1985). In contrast, applying such solutions to mine reclamation liabilities is significantly more costly. In particular, resource extraction firms cannot employ the most obvious solution to the debt overhang problem—finance projects solely through

[1] A mine reclamation liability is the obligation of a mining operator to restore disturbed mining land to a natural or economically usable state after the productive life of a mine.

equity. Rather, the production function of these firms dictates the creation of reclamation liabilities in order to extract valuable minerals from the ground. Beyond this, it is very difficult for firms to shorten the maturity of reclamation liabilities or to renegotiate the terms of the obligation ex post.

The nature of mining regulations also enables me to observe the importance of secured liabilities. Over time most jurisdictions have implemented legislation that requires mining operators to financially guarantee, or bond, their reclamation liabilities. However, the accepted forms of these guarantees vary substantially across jurisdictions. Most can be generally classified into a group that requires explicit collateral (externally-bonded) or a group that does not (self-bonded). I exploit the differences in local bonding regulations around the world to identify plausibly exogenous variation in self-bonding, and to separate reclamation liabilities into a treatment group (self-bonded) that is comparable to unsecured debt, and a control group (externally-bonded) that is comparable to secured debt.

Using the data in Table 1, I define a self-bonded mine as any mine permitted during a time self-bonding is allowed. All other mines are classified as externally-bonded. Each year, I sum the estimated reclamation liabilities of a firm's self-bonded and externally-bonded producing mines to get a total dollar amount for each type of reclamation liability. I create two sets of empirical measures of firms' exposure to the overhang problem. The first set consists of market leverage ratios, or the dollar amount of debt, self-bonded and externally-bonded reclamation liabilities, respectively, divided by the market value of assets of the firm (Frank and Goyal (2009)). The second is a group of indicator variables equal to 1 if a firm's existing liabilities exceed the potential value added by a new mining project. These indicators identify growth options that go unfunded in Myers' baseline model.

These empirical measures of overhang allow me to distinguish the impact of each respective liability on a firm's propensity to invest in new mineral projects. Mineral projects are well suited for studying overhang as they require two pre-production investments, one up-front to acquire the mining rights and a second in infrastructure and capital directly before production begins. The initial investment in the mining rights can be viewed as a growth option that expires immediately if the investment is not made. If a firm acquires the rights to extract the mineral, however, the project represents a second real option that is exercised when the firm makes the secondary investment and begins production. Myers's (1977) model shows the costs of overhang arise from firms completely forgoing investment in growth options that immediately expire, while in both Mello and Parsons (1992) and Mauer and Ott (2000), the loss of value due to debt overhang stems from firms suboptimally delaying exercise of the real option. Mineral projects allow me to test the implications of both aspects of debt overhang theory.

Consistent with both, I find that when firms are unable to avoid the overhang problem efficiently, they both forgo and postpone positive NPV mining projects. Specifically, self-bonded reclamation liabilities negatively impact the likelihood a firm acquires new positive NPV mining rights, while externally-bonded reclamation liabilities and traditional debt liabilities do not. Further, I find that only firms with large exposures to self-bonded reclamation liabilities are significantly more likely to delay construction on positive NPV mining projects. Consistent with Myers (1977), each of these effects is more pronounced when firms' liabilities are plausibly more risky.

The fact that traditional debt imposes few overhang-related distortions does not mean debt overhang is unimportant in capital structure decisions. Rather, taken together, the results highlight exactly how important the agency costs of debt overhang are and why effective solutions have endogenously arisen to avoid them. Further, if we assume that a dollar of reclamation liability imposes similar overhang as a dollar of traditional debt, the ex post costs of overhang imposed by mine reclamation liabilities provide an upper bound estimate of the ex ante contracting costs of avoiding overhang from traditional debt.

My results indicate that these costs are nontrivial. In particular, firms passing on positive NPV mining rights translates to an expected loss in value of roughly \$0.63 million each year. Additionally, for firms with self-bonded reclamation liabilities that exceed the potential value of a new project, the average delay in construction is nearly two years. In time value of money terms, this equates to an expected loss of nearly a \$1.1 million each year. Given the average reclamation liability in my sample is held for 20 years, the present value of these annual costs roughly aggregated together is \$17.66 million, or 6.27% of market value for the median firm with at least one producing mine.

Due to the unique nature of the resource extraction setting, I explore several tests related to the external validity of the main results. First, I verify that the results from more traditional tests of debt overhang are similar in my sample.[2] Specifically, I correlate firms' liability leverage ratios with capital expenditures. Consistent with results from a typical panel of U.S. industrial firms, I find a strong negative relation between capital expenditures and a firm's traditional debt leverage ratio. This is despite my baseline finding that traditional debt is unrelated to investment in positive NPV projects. Together, these results suggest that previous studies that find a negative relation between leverage ratios and CapEx do not identify a debt overhang effect. Instead, the negative relation may identify a decrease in the firm's investment opportunity set or even a decrease in *negative* NPV projects.

Second, I show my general results hold in two separate samples of U.S. mining firms. The first sample uses hand-collected data on U.S. mining firms that voluntarily disclose information on the amount of their reclamation liabilities and bonding methods in their 10-Ks and other reports. The second uses mine-level data from the Mine Safety and Health Administration (MSHA). I hand-match each of these samples to Compustat and find that three separate measures of a firm's self-bonded reclamation liabilities are negatively related to its capital expenditures.

My paper makes three contributions to the corporate finance literature. First, I believe mine is the closest to testing Myers's (1977) theory, and thus the importance of debt overhang, directly. In doing so I find that self-bonded reclamation liabilities, and not traditional debt, negatively impact investment in mining projects that are shown to be valuable ex ante.

[2]E.g., see Lang et al. (1996), Ahn et al. (2006), Cai and Zhang (2011), Cantor (1990), Whited (1992), Opler and Titman (1994), Sharpe (1994), and Aivazian et al. (2005).

This suggests that firms incur much of the costs of overhang ex ante through contracting and debt composition solutions rather than ex post through investment distortions. Other recent studies have also highlighted the importance of debt overhang. For example, Melzer's (2017) and Bernstein's (2018) findings imply that the overhang problem is not confined to the corporate setting. Additionally, the ex ante NPV estimates allow me to roughly estimate the costs of overhang in a reduced-form way. Much of the evidence from structural models suggests that debt overhang may not be a first-order concern in capital structure decisions as the estimated agency costs are typically 1-2% of market value.[3] My results, however, suggest that the costs (6.27% of market value) could be even larger than the structural estimates in Titman and Tsyplakov (2007) and Moyen (2007).

Second, this paper contributes to the literature on potential solutions to asset substitution and debt overhang problems. In recent empirical work, both Gilje (2016) and Denes (2017) argue that covenants, debt composition and other economic mechanisms are efficient in reducing the incentives of firms to engage in risk-shifting. These papers offer considerable insight into the struggle to identify empirical evidence of Jensen and Meckling's (1976) asset substitution problem. The evidence on avoiding debt overhang is currently more segmented. For example, a branch of literature concentrates only on renegotiation.[4] In fact, in one of the most well-executed empirical approaches, Giroud et al. (2012) show that the renegotiation of debt contracts substantially improves a firm's operating performance. My results indicate that the collective set of solutions described above reduce the inefficiencies stemming from debt overhang.

Lastly, this paper complements a growing strand of literature that focuses on the impor-

[3] E.g., see Mello and Parsons (1992), Parrino and Weisbach (1999), Mauer and Ott (2000), Hennessy (2004), and Childs et al. (2005).

[4] E.g., see Aivazian and Callen (1980), Gertner and Scharfstein (1991), Mella-Barral and Perraudin (1997), Pawlina (2010), and Chu (2016). There is also a branch that focuses only on debt maturity. For example, Barclay and Smith (1995), Guedes and Opler (1996) and Barclay et al. (2003) find a negative correlation between debt maturity and growth opportunities, while Johnson (2003) shows that short debt maturity attenuates the negative relationship between leverage and growth opportunities. Additionally, Billett et al. (2007) focus on the endogenous evolution a firm's growth options with respect to leverage, debt maturity and covenants and Diamond and He (2014) analyze the specific conditions in which short maturity debt will lead to lower overhang.

tance of a firm's non-debt obligations. Similar to prior studies on pension (Rauh (2006)) and legal (Arena and Julio (2015), Bennett et al. (2018)) liabilities, I show non-debt liabilities have a first-order effect on firm investment policy.[5] Further, both Rauh (2009) and Akey and Appel (2018) document evidence of the effect non-debt liabilities can have on managerial risk-shifting behavior. Finally, Chen et al. (2018) and Chang et al. (2018) show that operating leverage and corporate environmental liabilities, respectively, are substitutes for traditional debt liabilities.

2 Institutional Setting

2.1 Resource Extraction in Canada

Canada consistently ranks among the world leaders in the global production of minerals and metals (Marshall (2017)). Due to this rich geology, a substantial number of resource extraction firms locate in Canada and ultimately list on either the Toronto Stock Exchange (TSX) or TSX Venture Exchange (TSXV). According to the National Resource Governance Institute, the TSX and TSXV account for 31% of the world's public mining firms, and 15% of the global mining market value. Richer La Flèche et al. (2016) report that more mining companies are listed on the TSX and TSXV than on any other exchange in the world.

Beyond the size of the public resource extraction industry in Canada, this setting offers an even more significant advantage. Following the Bre-X mining scandal in the 1990s, the Ontario Securities Commission introduced the National Instrument 43-101 Standards of Disclosure of Mineral Projects, a listing requirement for both the TSX and TSXV (Fox (2017)).[6] Upon their implementation in 2003, the main tenets of National Instrument 43-101 (NI 43-101) standardized the reporting of all scientific and technical information, and required this information to be prepared by or under the supervision of a "qualified person".

[5] While Rauh (2006), Arena and Julio (2015), and Bennett et al. (2018) each have a different mechanism in mind, their results are consistent with an overhang channel.

[6] Documentation on National Instrument 43-101 Standards of Disclosure on Mineral Projects can be found at http://web.cim.org/standards/documents/block484_doc111.pdf.

The qualified person must have a mining-specific academic and career background, among other credentials, and is *liable* for the content of the NI 43-101 technical reports (Fox (2017)).

These reports are "a summary of material scientific and technical information concerning mineral exploration, development, and production activities on a mineral property that is material to an issue" (Ontario Securities Commission (2011)). They include highly detailed information on the property itself, as well as resource and reserve quantities (both proven and probable), and the potential economic viability of the project. The economic viability is analyzed in a series of reports (preliminary economic assessment (PEA), the pre-feasibility report, and the feasibility report). Each is required to include capital and operations cost estimates, estimated mine life, forecasted production and revenues, and the overall estimated NPV of extracting the mineral.

These data are extremely rich and allow me to directly test Myers's (1977) prediction that firms will pass on positive NPV projects. It is possible, however, that the NPV estimates are uninformative, inaccurate, or biased. In recent studies on voluntary disclosure of gold feasibility studies in Australia, Ferguson and Pündrich (2015) and Ferguson et al. (2013) find that the technical reports contain information used by investors, suggesting these types of reports are at least partially informative. Further, Internet Appendix Table IA1 reports linear regression results in which the dependent variables are various cumulative abnormal returns around the disclosure of the first NPV estimate in a technical report. While small, the coefficient on NPV/Market capitalization$_{t-1}$ is positive and significant, suggesting that investors believe a higher estimate means a higher project value. Finally, holding the qualified person liable for inaccuracies or forgeries helps alleviate concerns of large misrepresentations.

2.2 Mine Reclamation and Financial Assurance

Mining regulations and regulatory bodies display significant heterogeneity around the world (e.g., see Richer La Flèche (2016)). Nearly unanimously, however, regulators direct mining operators to disturb as little land as possible, and to reclaim the disturbed areas when extraction is complete. While reclamation standards vary among different jurisdictions, most include an extensive amount of work to be completed after the productive life of a mine.[7] Because this creates a substantial, long-term obligation for mining operators, most jurisdictions require mining operators to post financial assurance, or a bond, that ensures the costs of reclamation will be borne by the mining company and not by the local government and taxpayers. There are essentially four main types of financial assurance: (1) surety bonds; (2) collateral bonds; (3) letters of credit; and (4) self-bonds.[8] The types of financial assurance that are deemed acceptable vary substantially across jurisdictions and through time.

Surety bonds are the most common type of financial assurance method, particularly for mines in the United States (Gorton (2009), Nazzaro (2005)). A surety company or other type of financial institution (the surety) will agree to act as an obligor should the mining operator fail to complete the reclamation. Collateral bonds typically take the form of cash trusts or certificates of deposit and are easily accessible to the regulator. Letters of credit are heavily used outside of mining, often to reduce counter-party trade risk. Typically the buyer receives a letter from a bank guaranteeing that payment for the goods or services will arrive on time and in full. Similarly, mining operators use letters of credit to guarantee that reclamation will be completed as agreed upon during permitting.

These first three methods of bonding are costly ex ante for the mining operators. Survey evidence indicates that surety premiums in the U.S. can range from 1-3.5% (Kuipers (2000)) to 5-6% (Chelimsky (1988)) with up to a 100% collateral requirement. Recent anecdotal evidence also suggests that surety bonds can be extremely costly. Bonogofsky et al. (2015) report that Cloud Peak Energy saved upwards of $2 million per year switching from surety to self-bonds. While letters of credit often have much lower annual premiums, banks require

[7] This often includes, but is not limited to, demolition of existing mining structures; sealing and stabilization of pits and shafts; recontouring of access roads, tailing ponds, trenches, pits and shafts; revegetation; and monitoring and evaluation of the site. For a checklist that details reclamation standards in the U.S., see Bureau of Land Management (2005).

[8] This list does not include some less common forms of financial assurance accepted by certain jurisdictions. For example, some states/provinces and countries have what is called a bond pool, where mining operators each contribute to a government sponsored pool which is used to perform reclamation. For more information on reclamation bonding options, see Cheng and Skousen (2017), Gorton (2009), and Nazzaro (2005).

collateral deposits upwards of 120-200% of total estimated liability to provide the guarantee (Kirschner and Grady (2003)). Thus, there are significant liquidity costs associated with letters of credit or cash bonds.

Unlike the first three bonding methods, a self-bond does not necessitate explicit collateral. Rather a self-bond requires that an operator, its parent, or a third-party provide the guarantee for the cost of reclamation. Because of financial and credit-worthiness criteria, as well as contracting costs with a third party, most self-bonds are parent guarantees, often called company or corporate guarantees (Nazzaro (2005)).[9]

2.3 Contracting Costs

There are several potential, yet costly, solutions to the overhang problem. First, a firm can simply choose to finance its projects through equity rather than debt and avoid the overhang problem altogether. Additionally, Myers (1977) notes that debt overhang can be resolved through a policy of (non-automatically) rolling over short maturity debt claims, or ex post through renegotiation. Finally, Stulz and Johnson (1985) suggest that the use of secured debt can mitigate the underinvestment problem.

Reclamation liabilities offer a unique instrument to study these potential solutions as the costs of implementing the first three listed above are significantly higher than they are for traditional debt.[10] For example, while not costless to finance solely through equity, the option nevertheless exists and is often exercised by firms (e.g., see Strebulaev and Yang (2013)). This option, however, is not applicable to reclamation liabilities. For any resource extraction firm wishing to extract minerals from the ground, a reclamation liability is simply part of the production function that cannot be avoided.

[9] Additionally, many jurisdictions that allow self-bonds only explicitly allow parent guarantees. For example, Missouri's metal mining statute, MO Rev. Stat. §444.368.1 states, "...the operator shall file a demonstration of financial assurance in the form of a bond, certificate of deposit, letter of credit, insurance, *company guarantee* [emphasis added], escrow agreement or other form of financial assurance as approved by the director."

[10] Additionally, my identification strategy (discussed below in Section 3.1) allows me to separate unsecured mine reclamation liabilities

Diamond (1991, 1993) and Sharpe (1991) point out that short maturity debt contracts impose costs due to liquidity risk. Even so, firms continue to issue large amounts of short maturity bonds. For example, Xu (2017) finds that the average maturity for U.S. corporate bonds over a sample period of 1997-2012 is 10.5 years. Additionally, Custódio et al. (2013) report that less than 20% of the median firm's total debt outstanding has a remaining maturity greater than five years. Both of these studies highlight that firms regularly issue and hold effectively short maturity debt, suggesting the potential overhang costs of long maturity debt exceed the costs of liquidity risk from short maturity debt.

On the other hand, the average mine in my sample has a productive life greater than 20 years, creating a liability with twice the maturity as the average bond. Strategically choosing short-life mines ex ante is likely to be extremely costly, as the life of a mine is highly correlated with the amount of mineral contained in the deposit and thus the NPV of the project. Additionally, because "retiring" a mine reclamation liability involves closing and reclaiming a mine, shortening the maturity of current a mine reclamation liability even by a few years could mean sacrificing millions of dollars in valuable resource extraction and production.

The best evidence of the costs of renegotiating debt contracts is the sheer prevalence with which renegotiation takes place. For example, Roberts (2015) finds that the average bank loan is renegotiated 3.5 times, while Roberts and Sufi (2009) report that 90% of private credit agreements are renegotiated prior to their original maturity. Further, Nikolaev (2017) finds that the unconditional probability a firm renegotiates at least one contract in a given year is 37%. Taken together, the results in these studies suggest that the overall costs of renegotiation for traditional debt are likely not prohibitive. This does not mean, however, that renegotiating debt contracts is never costly. For example, Chu's (2016) findings imply that syndicate loans with many lenders are more costly to renegotiate. This suggests that there is heterogeneity in the costs of avoiding overhang even in the cross-section of traditional debt contracts. Indeed, Internet Appendix Table IA2 supports this hypothesis using the sample from Table 4 below.

Mine reclamation liabilities, on the other hand, appear significantly more difficult to renegotiate ex post in a way that mitigates overhang. First, while traditional debt contracts can be renegotiated on price, level, maturity, or specific covenants, mine reclamation liabilities can only be renegotiated on level. Second, adjusting the level of a mining firm's liability down requires local regulators to transfer the obligation from large mining firms to local taxpayers. Third, it is unclear if regulators have the authority to renegotiate in all jurisdictions (e.g., see Socalar (1988)). Lastly, while reclamation liabilities and their associated bonds are often negotiated up, (e.g., see Walsh (2017)), using a comprehensive news search, I cannot find evidence of a single anecdote in which a reclamation liability is renegotiated down to a point that it no longer accurately represents the expected costs. All of this evidence supports the argument that reclamation liabilities have renegotiation costs that exceed those of traditional debt.

3 Empirical Strategy and Data

3.1 Identifying Reclamation Liabilities

The difference in contracting costs between traditional debt and reclamation liabilities allows me to assess the importance of debt overhang through a direct comparison of the different liabilities' impact on investment policy. While traditional debt is defined simply as the sum of a firm's book value of total long-term debt and the book value of debt in current liabilities, defining a firm's exposure to reclamation liabilities is more complex. Estimated reclamation liabilities are reported in project technical reports; however, even with the high disclosure standards, firms are not required to disclose the manner in which their reclamation liabilities are bonded. Rather, I exploit the cross-section and time-series variation in self-bonding regulations to identify quasi-exogenous variation in firms' ability to self-bond their mine reclamation liabilities. This variation allows me to separate reclamation liabilities that must be backed by explicit collateral.

There are two main endogeneity concerns with this identification strategy. The first is that the passage of self-bonding regulations is influenced by the prevailing political economy or specific firms that lobbied for or otherwise motivated the laws (e.g., see Karpoff and Wittry (2018)). Because the collection of lobbying expenditures and reports is a relatively new phenomena, the majority of law changes precede the beginning of databases such as Open Secrets. In any case, I can only identify 4 cases in which a single firm reports it is "monitoring the status of self-bonding legislation." Further, I cannot find a single news article that credits a specific firm for a change in the financial assurance provision in any jurisdiction's mining regulation.

The second concern arises from the fact that firms choose where and when to purchase new mining rights and which mining projects to move forward with permitting. This could potentially introduce a selection bias in which firms decide to purchase or permit mining projects in accord with the prevailing self-bonding regulations. While this firm-mine matching problem cannot be completely ruled out, Internet Appendix Table IA3 analyzes the number of new mining rights acquired and the number of new mining projects permitted in a jurisdiction surrounding the passage of a self-bonding provision. The results indicate that, on aggregate, self-bonding provisions do not influence firms' decisions to buy new rights or start the permitting process in one director or the other.[11]

Table 1 displays all self-bonding regulations for which I can locate documentation. These regulations cover over 90% of the permitted mines in my sample. In Australia and Canada, individual states and provinces, respectively, set bonding laws, while the U.S. federal government regulates financial assurance provisions that are specific to the mineral to be extracted.

[11] Internet Appendix Table IA3 mitigates not only this particular selection concern, but also a confounding variables problem based on geographic clustering. Specifically, if there are location economies, firms may choose to locate several mines in the same jurisdiction. Then, if the passage of self-bonding provisions is correlated with negative local economic or industry-specific shocks, it could be the case that firms with several mines located in such a jurisdiction will also have high levels of self-bonded reclamation liabilities and will invest less in the future due to the shock, not the amount of their reclamation liabilities. The results in Internet Appendix Table IA3 strongly suggest this is not the case.

Namely, the Surface Mining Coal Reclamation Act (SMCRA) of 1977 regulates coal mining, while 43 C.F.R §3809, passed in 2001, amended legislation for hardrock and metal mining. However, both federal provisions give states the option to be *more* stringent, that is, to prohibit self-bonding if the federal law allows it.[12] Like 43 C.F.R §3809, most of the mining regulations in Table 1 are part of a country's or state's mining reform, where previous legislation had required mines to be reclaimed without requiring financial assurance. I assume the time periods preceding reform, and jurisdictions with a reclamation requirement yet without a financial assurance provision, explicitly allow self-bonding, as the incentive to fulfill the obligation in all cases is very similar.

Under these assumptions, I define a mine as self-bonded if it was permitted in a state, province, or country, and in a year in which a self-bond or corporate guarantee was considered an acceptable form of financial assurance. All other mines are classified as externally-bonded. This definition is analogous to assuming that mining companies choose to self-bond whenever they are able. This seems reasonable for a few reasons. First, the other forms of financial assurance are costly ex ante. Annual premiums and collateral requirements can add millions of dollars to estimated reclamation costs. Second, the evidence supports that the option to self-bond is heavily exercised (e.g., see Interstate Mining Compact Commission (2014) and Nazzaro (2005)). Finally, the fact that some mines defined as self-bonded in my sample are not actually guaranteed through a self-bond should bias me against finding significant results.

Each self-bonded and externally-bonded mine in some stage of production contributes to a firm's overall self-bonded and externally-bonded reclamation liabilities, respectively. Using the definition of a self-bonded mine above, a firm's self-bonded liabilities in a given year are comprised of the estimated reclamation liabilities for all of its producing self-

[12] 43 C.F.R §3809 amended the law to prohibit new self-bonds but explicitly grandfathered existing self-bonded mines to that form of financial assurance. Thus, producing hardrock and metal mines were not required to provide additional financial assurance after the change. In other legislative changes around the world, it is much more difficult to discern the existence of and details regarding a grandfather provision. In this paper, I assume all current self-bonded mines are grandfathered in at the point of the law change, only needing to be re-bonded with a new form of financial assurance if the mine owner changes.

bonded mines. I use the short-hand notation SB to represent the total U.S. dollar amount of a firm's self-bonded reclamation liabilities in a given year. Formally, this is $\text{SB}_t = \sum_{i \in \text{P,S}} E[\text{Reclamation liability}_{it}]$, where P represents mines in production, S represents mines defined as self-bonded, and the estimated reclamation liability is reported in the technical reports prior to the mine being permitted. Similarly, EB represents the total U.S. dollar amount of a firm's externally-bonded reclamation liabilities in a given year. Formally $\text{EB}_t = \sum_{i \in \text{P,E}} E[\text{Reclamation liability}_{it}]$, where again P represents mines in production, but E represents mines defined as externally-bonded. These definitions of a firm's self-bonded and externally-bonded liabilities assume that once a mine is closed, the firm is no longer exposed to the liability.[13] This assumption alleviates concerns due to a particular financial constraint in which firms with self-bonded reclamation liabilities cannot fund new investment due to their ongoing clean-up costs.

Figure 1 presents an example of three coal mines located near the border of British Columbia and Alberta, Canada. Transalta permitted Highvale Coal Mine in Alberta in 2007. Because Section 21 of the Alberta Conservation and Reclamation Regulation, passed in 1993, permits self-bonds, the \$42.1M in estimated reclamation costs for Highvale is is added to SB for Transalta when Highvale entered production in 2008.

Teck Resources permitted the other two mines, Greenhills Coal Mine and Elkview Coal Mine, in British Columbia in 1992 and 2008, respectively. Section 30 of British Columbia's Bonding Act, passed in 1996, prohibits self-bonding. Because of the timing of the regulation in British Columbia, the \$153.2M in estimated reclamation costs for Greenhills is added to SB for Teck Resources when Greenhills began production in 1993, and the \$53.4M in estimated reclamation costs for Elkview is added to EB when Elkview continued production in 2009 after being acquired by Teck Resources.[14] All three of these mines are still in production

[13] This is equivalent to making the assumption that mining firms begin reclamation as soon as a mine is closed. The results are robust to alternative assumptions, such as the liability persists for 1, 2, or 3 years after ceasing production.

[14] Because of the time-series variation in regulations, even the same mine can be defined as self-bonded and externally-bonded over different parts of the sample if ownership changes following the passage of new regulation. For example, Highvale Coal Mine could be defined as externally-bonded if Teck Resources

today and so are still considered liabilities for the respective companies at the end of my sample in 2016.

3.2 Empirical Measures

In an ideal empirical setting to test debt overhang, a researcher would use exogenous variation in the value of a firm's option to default on their obligation (Merton (1974)) to study the impact on investment. Unfortunately, the value of a firm's put option is not directly observable. Thus, to facilitate the comparison between traditional debt and mine reclamation liabilities and analyze their impact on investment, I use two separate empirical measures that should be positively correlated with the value of the firm's option to default.

The first measure can essentially be thought of as a "leverage" ratio. For traditional debt, it is the market debt leverage ratio as defined in Frank and Goyal (2009). That is, debt leverage is a firm's total debt divided by the market value of its assets.[15] For reclamation liabilities, the denominator of market value of assets remains the same, but total debt is replaced by SB or EB, the total dollar amounts of a firm's respective reclamation liabilities. Thus, the "leverage" ratios for self-bonded and externally-bonded reclamation liabilities are SB/MV and EB/MV, respectively. While leverage ratios are commonplace in studying debt overhang, there is no strong theoretical basis for why debt or other obligations would impact firms' option to default, and ultimately investment policy in a linear way. Rather, it is much more likely that firms are impacted by overhang only after crossing some threshold.

The second empirical measure I use attempts to account for this nonlinearity. In doing so, I take advantage of the richness of the project-level data extracted from the NI 43-101 technical reports. I define indicator variables that take a value of 1 if a firm has existing liabilities—traditional debt, SB, or EB—that exceed the potential value created by the new mineral project. For example, if SB exceeds the estimated NPV of a potential mining project, $\mathbb{1}_{\text{SB}\geq\text{NPV}} = 1$. Ultimately, this measure is also imperfect. While it exactly identifies

transferred the rights to a new owner after 1996.

[15] See Appendex Table A1 for the details of this ratio.

the "wedge" in the baseline model, Myers (1977) provides a generalization of the problem for firms that hold more than one asset. In this model, the investment decision is slightly more complicated as the firm compares the NPV of the project ($\Delta V(s) - I(s)$ in Myers's notation) against the capital gain to bondholders if the option is exercised ($\Delta V_D(s)$). In this framework, using the entire liability (P) assumes the bondholders would get nothing if the firm does not exercise the option. This is unlikely the case for firms with assets in place. However, using the entire liability allows me to avoid making assumptions about asset allocations in bankruptcy. It also represents a necessary (but insufficient) condition for firms to be exposed to overhang (i.e. it must be the case that $\Delta V_D(s) \leq P$).

3.3 Data and Summary Statistics

The vast majority of mine-level data used in this paper are contained in public company filings. Extensive mine-level information is disclosed in regularly filed reports such as the Management Discussion and Analysis (MD&A) report. The NI 43-101 technical reports also contain detailed data on a company's mineral projects. A mining research firm called Mining Intelligence aggregated the information in these filings and provided me a database of nearly 800 publicly traded Canadian mining firms owning over 3,600 mining projects worldwide during my sample period of 1990-2016.[16] The data includes historical mine status and ownership, mine type and location, cost of acquisitions, and information extracted from the NI 43-101 technical reports on feasibility. I supplement the data provided by Mining Intelligence with hand-collected estimates of a mine's reclamation liabilities for each permitted mine in my sample.[17]

Table 2 reports summary statistics for the mining projects in my sample. Panel A shows that only 14% of the 22,379 mine-year observations are in some phase of production. The

[16]Mining Intelligence is a division of InfoMine, Inc., a private data intelligence firm that provides data solutions and services to mining companies, suppliers, educators and financiers. The company claims to cover over 14,000 resource extraction companies and 36,000 mining properties worldwide, while collecting data from over 1.8 million publicly filed documents.

[17]Estimates for the reclamation liabilities are most often found in the NI 43-101 technical reports filed prior to production and thus are not updated through time.

vast majority of mining projects in this sample are in the earlier stages of development, with two-thirds of the observations in prospecting, exploration or feasibility. Figure 2 displays these different stages for a typical mining project and highlights important milestones such as when different feasibility studies are often disclosed.

Panel B of Table 2 displays summary statistics for only mines that enter production at some point during the sample. These mines would have been required to submit a bond for their reclamation liabilities to the appropriate local authorities. The average liability of \$27.7 million and maximum of \$558 million highlight that reclamation liabilities are non-trivial. Using the self-bonded definition described in Section 3.1, I classify nearly 40% of the producing mines in my sample as self-bonded. The remainder of Panel B shows the distribution of mines by mine type, primary mineral extracted, and mine location. The mines in my sample are most likely to extract gold and be located in North America, byproducts of using Canadian mining firms.

Panel C of Table 2 reports the descriptive statistics for the data from the NI 43-101 technical reports and the NPV calculated for the acquisition of mining rights. The mean project (acquisition) NPV is over \$400 (\$200) million while the median is \$172 (\$68.5) million. Due to the high costs of exploratory drilling and the commissioning of the technical reports, it is reasonable to assume that firms only pursue feasibility studies on mining projects that are almost certainly positive NPV. This biases the sample towards including only valuable projects. However, this sample composition allows me to precisely test Myers's (1977) theory that firms will forgo *positive* NPV projects. Panel C also highlights the frequency at which firms are exposed to Myers's (1977) wedge. Of the 269 projects, nearly 18% are owned by a firm with debt liabilities that are greater than the estimated value of the project at the time of disclosure. Far fewer firms have enough exposure to self-bonded reclamation liabilities to surpass the estimated project NPV. However, as I show in Section 4, these liabilities have a large impact on investment decisions.

Table 3 summarizes firm-level variables on reclamation liabilities and the number of mining projects in various stages of development for the Canadian mining firms over the sample period of 1990 through 2016. This data is aggregated to the firm-year level from the mine-level data described above. The average firm in my sample with at least one mine in production has just under \$70 million in reclamation liabilities while the median firm has \$11.9 million. This suggests that reclamation liabilities, while typically smaller, are of the same order of magnitude as traditional debt liabilities. Table 3 also displays a firm's liability leverage ratios. The maximum values of SB/MV (10.54) and EB/MV (72.41) indicate that certain firms have very high exposure to reclamation liabilities. These statistics match evidence from the U.S. coal mining industry that indicates huge amounts of reclamation liabilities, self-bonded and otherwise, can be concentrated in a small number of firms (Hein et al. (2016)).

Finally, Table 3 displays the summary statistics for the firms' accounting variables taken from Compustat—North America. Compared to the typical U.S. industrial sample, the average Canadian mining firm tends to be smaller in terms of book (\$856.8M) and market (\$925M) value of assets, have lower leverage (0.109) and much more variable operating performance. They hold a comparable amount of cash as a percentage of book value of assets (0.243) and have similar growth opportunities (Tobin's $Q = 3.3$).

4 Results

4.1 Acquisition of Mining Rights

My first set of empirical tests analyzes firms' propensity to make an initial investment in a new project. For mining firms, the initial investment in a mining project is acquiring the rights to extract the mineral in a specific deposit. I view the option to purchase new mining rights as an auction. Thus, the option immediately expires upon completion of the auction for all but the highest bidder.[18] This framework allows me to test Myers's (1977) main

[18] In reality, it is possible the option to purchase new mining rights does not expire immediately. For example, two mining firms may have protracted negotiations regarding the transfer of rights that extends

empirical prediction that firms will completely forgo investing in positive NPV projects in some states of the world.

Table 4 reports the results of linear probability models (LPMs) in which the dependent variable takes a value of 1 if the firm acquires new mining rights in a given year, and 0 otherwise.[19] The sample consists of firms located in Canada and listed on the TSX or TSXV exchanges from 1990 through 2016. This first set of tests uses firm-year observations and because I am not analyzing a firm's decision with respect to a specific mineral project, I am constrained to using the first empirical measure, the liability leverage ratio. Each model includes both firm and year fixed effects and I cluster the standard errors at the firm level.

Firm fixed effects present a trade-off between controlling for firm-specific characteristics and perhaps limiting the exogenous variation in the reclamation liabilities. Specifically, when using firm fixed effects, the entire effect is identified by time-series variation within a given firm. Because the estimated reclamation costs are not updated through time, the majority of the variation in my measures of reclamation liabilities comes from starting new mines and closing old ones—both arguably endogenous to the decision to invest in new mines. Notwithstanding, I use firm fixed effects in all LPM specifications in my empirical tests for a few reasons. First, this does not limit the plausibly exogenous variation in defining liabilities as self-bonded or externally-bonded. If the results were driven completely by opening and closing mines, there should be no difference between self-bonded and externally-bonded reclamation liabilities. Second, the survival analysis in Section 4.2.2 should alleviate this concern completely as all of the identification in those tests comes from the cross-section of reclamation liabilities at the time of the first NPV estimate.

Models (1) and (2) of Table 4 analyze the acquisition of any mining project in my sample—not just those that are NPV positive. Model (1) reports the results without controls, while Model (2) adds covariates for a firm's size (log of book assets), internal capital con-

the life of the option.

[19]I use the LPM as it allows me to include high-dimensional year and firm fixed effects. See Angrist and Pischke (2008) for a discussion on the advantages of the LPM. My results remain largely unchanged when using a logistic or probit model without firm fixed effects.

straints (cash), profitability (ROA), growth opportunities (Tobin's Q and firm projects in each stage of development), and lifecycle effects (log of firm age). The results suggest that both self-bonded reclamation liabilities and traditional debt have a significant and negative impact on a firm's propensity to acquire new mining rights. In fact, the magnitude of the coefficient for traditional debt (-0.057) is nearly two times as large as that of self-bonded reclamation liabilities (-0.035). However, because it is unclear whether these mining projects are value increasing, there are several alternative explanations for the negative correlations reported for traditional debt. For example, one such explanation is that debt acts as a governance mechanism, limiting costly overinvestment due to Jensen's (1986) free cash flow problem.

In order to rule out these alternative explanations, Models (3) and (4) restrict the new mining rights to those that look to be NPV positive. I define the NPV of the mining rights as the initial NPV estimate in the NI 43-101 technical reports (or the remaining NPV estimated by this report if the mine is already producing) less the cost the paid for the individual mine at acquisition.[20] Thus, the dependent variable in Models (3) and (4) takes a value of 1 in a year the firm acquires new mining rights in which this NPV is positive, and 0 otherwise.

The inference from these models differs substantially from the first two. For example, in Model (4), the coefficient on SB/MV remains similar in magnitude at -0.025 and significant at the 1% level. The coefficient on the market debt leverage ratio, in contrast, is cut in half and not significant at conventional levels. The results in Models (3) and (4) support Myers's (1977) empirical prediction that firms with existing obligations will forgo positive NPV growth options in some states of the world, however only for firms with self-bonded reclamation liabilities.

The costs of this investment distortion are non-trivial. For a one standard deviation

[20] This mining rights NPV measure is noisy as individual mines are often sold in package deals as assets in the sale of a mining operator and it is difficult to assess the value of the mine without the operator's capital assets. However, in using the entire acquisition cost, once again I'm biasing the sample toward the *more* valuable projects. I exclude sales of mining operators that own more than one mining project. I control for the incidence of these multi-project acquisitions, as well as the acquisitions of mining rights that have no NI 43-101 technical report NPV estimate, in Model (4).

increase in SB/MV, the average firm is 12.3% less likely to acquire positive NPV mining rights relative to the baseline likelihood for firms with at least one producing mine. In expectation, this amounts to a $-0.123 \times 0.057 \times 89.1 = \$0.63$ million loss in value each year the firm maintains high exposure to self-bonded reclamation liabilities. Given the average reclamation liability in my sample is held for 20 years, the present value of these annual costs is $6.38 million, or 2.27% of market value for the median producing firm.[21] Overall, the results in Table 4 highlight that (1) overhang imposes meaningful costs, and (2) contracting mechanisms allow firms with traditional debt liabilities to avoid these costs, at least ex post. Ultimately, the costs imposed by mine reclamation liabilities offer an upper-bound estimate of the contracting costs firms pay ex ante to avoid traditional debt overhang.

4.2 Mining Projects as Real Options

Section 4.1 analyzed the impact of a mining firm's liabilities on its propensity to acquire *new* NPV positive mining rights in an auction framework in which the option to invest expires. Once a firm acquires the rights to mine a particular deposit, however, it has the exclusive right to extract the mineral for a considerable time period. The firm can choose to immediately make capital and infrastructure investments, or the firm can choose to delay construction until a future date. Thus, these existing mining projects represent real options for the firm.

Mello and Parsons (1992) and Mauer and Ott (2000) model debt overhang in a real options framework in which the firm's growth option does not simply expire. Both studies, using resource extraction for the setting of their contingent claim models, make similar empirical predictions—mainly the agency cost of debt overhang arises from suboptimal operating decisions. In particular, Mello and Parsons (1992) show that firms will delay opening (or reopening) a mine past the optimal trigger point when mineral prices are low. Similarly, Mauer and Ott (2000) find that firms will delay exercising the option to expand mining

[21] These calculations use the median NPV of $89.1 million for the mining rights and a market value $281.8 million (median for firms with at least one producing mine).

operations past the point which maximizes firm value.

In this section, I further utilize the data in the NI 43-101 technical reports to study these predictions. Specifically, I analyze the impact that liabilities have on firms' decision to start construction on positive NPV mining projects. The NI 43-101 technical reports allow me to use detailed data on project specifics in an attempt to control for the optimal trigger point. These controls include the NPV and capital costs associated with the mining project, the expected life of the mine, the primary mineral price, futures prices and implied volatility, among other things. Additionally, these data allow me to use Myers's wedge, the second empirical measure, which arguably does a better job classifying firms exposed to overhang.

The sample in this section is slightly different from the sample used in the previous section. While the analysis uses the same Canadian mining firms listed on the TSX or TSXV, it uses project-year observations rather than firm-year observations. Thus the sample uses annual data from the time a mineral project is estimated to have a positive NPV value (time t=0) until either the firm begins construction on the project or the sample period ends in 2016. The first year in which NPV estimates were provided in technical reports was 2003, creating a sample of over 800 project-year observations, covering almost 180 firms and over 200 mining projects from 2003 through 2016.

4.2.1 Linear Probability Model

Table 5 displays the results using a linear probability model in which the dependent variable takes a value of 1 if the project starts construction and 0 otherwise.[22] Once again, the models include year and firm fixed effects in the baseline specification, with robust standard errors clustered at the firm level.[23] Panel A continues to use the liability leverage ratios. Model (1) presents the results without additional control variables. Models (2) through (5)

[22] There is an additional advantage to using the LPM with the Myers's wedge indicators as Angrist and Pischke (2008) point out that models with categorical regressors do not satisfy the assumptions of the logistic or probit model as they are not continuous.

[23] Table IA3 in the Internet Appendix examines the robustness of Table 5 to the inclusion of alternative fixed effects, such has location, mine type, and primary mineral by year fixed effects.

in Table 5, Panel A, add a host of control variables. These covariates are meant to control for factors in a firm's decision to optimally exercise its real option to construct the mine. For example, Model (2) adds the standard accounting control variables that were used in Table 4, as well as project level controls for the NPV, capital costs, expected mine life, and number of projects in each developmental stage. Models (2) through (5) also attempt to control for a firm's investment opportunity set. Specifically, Model (2) controls for the total NPV of a firm's alternative projects and whether or not the firm begins construction on one of those alternative projects. Model (3) adds a control for the annual percentage change in the mineral price of the primary mineral to be extracted from the mine, Model (4) controls for the 12-month futures price, and Model (5) adds the implied volatility from historical 1-month at the money put-call straddles.

Consistent with the results from Section 4.1, each model in Panel (A) suggests that only self-bonded reclamation liabilities have a significant impact on firms' investment decisions. The economic magnitude of each coefficient is large, as Model (3) implies a one standard deviation increase in SB/MV leads to nearly a 25% decrease in the likelihood (relative to the baseline likelihood) a firm decides to begin construction on a positive NPV mining project in that year. While all insignificant, the coefficients on a firm's debt leverage ratio are negative in Models (2) through (5), while the coefficients on the externally-bonded reclamation liability ratio are positive.

Panel B of Table 5 switches to the second empirical measure—an indicator that equals 1 if the liability exceeds the estimated value of the mineral project and 0 otherwise. Once again, Model (1) examines the impact of the firm's liability on its propensity to begin construction without additional control variables, and the final four models repeat the analyses in Panel A with control variables that are firm-, project-, or mineral-specific and attempt to control for a firm's optimal trigger point. In each model, the coefficient on self-bonded reclamation liabilities is negative and significant while those on traditional debt and externally-bonded reclamation liabilities are near zero and insignificant. For example, the coefficient on Myers' wedge for self-bonded reclamation liabilities in Model (3) is -0.269 and is significant at the 5% level. This suggests firms with self-bonded reclamation liabilities that exceed the NPV of a mining project are nearly 27% less likely to begin construction on the mine in that year than otherwise similar firms.

A potential alternative explanation for the results in Table 5 is that a firm's existing self-bonded reclamation liabilities make it politically difficult for the firm to obtain permits for new mining projects. This explanation would be consistent with the negative relationship between the firm's liabilities and investment in new projects but has a very different interpretation than debt overhang. Table IA4 in the Internet Appendix addresses this concern by using the fact that many projects have a time gap between the permitting stage and the construction stage. These results show there is no discernible impact of a firm's self-bonding reclamation liabilities on the likelihood the firm obtains new permits.

Overall, the results in Table 5 are consistent with the predictions in Mello and Parsons (1992) and Mauer and Ott (2000) and suggest firm liabilities do impact a firm's decision to trigger the real option. However, as with the results in Table 4, this impact is concentrated in liabilities in which the contracting costs are high. In particular, a firm's ability to shorten the maturity of its debt or renegotiate it ex post limits the costs of the overhang problem. Further, the stark difference between self-bonded and externally-bonded reclamation liabilities in these tests offers support for the efficacy of secured obligations in mitigating overhang.

4.2.2 Survival Analysis

A second way to examine the effect of a firm's liabilities on its propensity to begin or delay construction is with survival analysis, which in this case, offers two advantages. First, survival analysis allows me to more directly test the empirical predictions in Mello and Parsons (1992) and Mauer and Ott (2000) that firms will delay the exercise of the real option. Additionally, I am able to roughly calculate an average delay imposed by a firm's liabilities. Second, survival analysis allows me to avoid the issue created by using firm fixed effects in the linear probability models. In particular, the variation in the Cox regressions comes from cross-sectional differences in firms' reclamation liabilities at the time of the initial NPV estimate, which is presumably a very similar point in development for each mine. To make this concrete, the liability leverage ratios and indicator variables are fixed at the time the NPV is estimated. Thus, the overhang measures are time-invariant across all observations for a specific project, from the point the NPV is estimated to the end of the sample in 2016, or when the event (construction) occurs, whichever comes first.

Figure 3 displays Kaplan-Meier nonparametric survivor functions for each of the three indicator variables identifying Myers's wedge. Consistent with earlier results, the delay for projects in which self-bonded reclamation liabilities exceed the estimated NPV is stark, while the survivor functions for externally-bonded reclamation liabilities and traditional debt show no discernible differences between projects in which the liabilities exceed the NPV and those that it does not.[24] It is possible, however, the nonparametric tests do not adequately account for factors relating to the optimal exercise date. Table 6 presents the results from Cox exponential proportional hazard regressions.[25] Each model includes year and primary mineral fixed effects.[26]

Panel A of Table 6 uses liability leverage ratios, while Panel B uses the indicator for Myers's wedge. Model (1) in Panel A displays the hazard ratios for the liability leverage ratios without additional control variables. The hazard ratio for SB/MV is 0.344 and is significant at the 10% level. The hazard ratio for EB/MV is also well under 1, but not significant at conventional levels. Finally, the hazard ratio for the market debt leverage ratio

[24]Further, the observed survivor functions for externally-bonded reclamation liabilities cross, a violation of the proportional hazard assumption.

[25]In unreported tests, I verify that the results are not sensitive to the assumed exponential proportional hazards distribution and remain qualitatively (and quantitatively) similar for the Weibull and Gompertz proportional hazards distributions.

[26]Firm fixed-effects are excluded from the Cox regressions in exchange for primary mineral fixed effects. Just as Kalbfleisch and Sprott (1970) find that logistic and probit models suffer from "incidental parameter bias" when using a large number of parameters, Allison (2002) uses simulations to show this same bias is nearly as severe in Cox regressions.

is very close to 1 and is insignificant, suggesting that a firm's traditional debt obligations are unrelated to its decision to exercise the real option.

Models (2) through (5) add the same additional control variables in the same progression as Table 5. The estimated project NPV, capital costs, and mine life are included as static covariates, while the rest are time-varying. Models (2) through (5) report results similar to those in Model (1) in that both self-bonded and externally-bonded reclamation liabilities seem to impact a firm's exercise decision. The main difference in these models is the hazard ratio for self-bonded reclamation liabilities is not significant at conventional levels. The biggest loss of significance happens when adding mineral-specific time-varying coefficients to the Cox models. This could reflect that when better controlling for factors that influence a firm's optimal trigger point, reclamation liabilities lose explanatory power. Alternatively, it could be the case that the time-varying coefficients make identifying true exposure to overhang, a tail event, more difficult for the liability leverage ratios.

Panel B of Table 6 displays the results using indicators for Myers's wedge with the same progression of static and time-varying control variables. The hazard ratios for self-bonded reclamation liabilities range between 0.329 and 0.476, and in contrast to Panel A, remain significant at the 5% level when including the additional control variables. This provides support for the argument that the linear measure for overhang struggles to properly identify when firms should be exposed to overhang. Again consistent with earlier results, externally-bonded reclamation liabilities and traditional debt do no impact the timing of the construction decision, suggesting that overhang is an important ex post concern only for liabilities that have high contracting costs.[27] These results provide direct support for the empirical predictions in Mello and Parsons (1992) and Mauer and Ott (2000).

The costs of the delay imposed by reclamation liabilities are significant. I use a back of the envelope calculation to approximate the delay induced by self-bonded reclamation liabilities.

[27]Even for liabilities with high contracting costs, other market mechanisms may arise to mitigate the costs of overhang. For example, Internet Appendix Table IA 6 shows that firms with high self-bonded reclamation liabilities attempt to sell positive NPV projects more often than otherwise similar firms.

Specifically, I calculate the difference in the survival function between $\mathbb{1}_{\text{SB}\geq\text{NPV}} = 1$ and $\mathbb{1}_{\text{SB}\geq\text{NPV}} = 0$ at each percentile of mining projects beginning construction. Firm-project pairs in which self-bonded reclamation liabilities exceed the NPV begin construction, on average, 1.9 years later than otherwise similar pairs. For the median project constructing nearly two years later, the new NPV is $172/1.075^{1.9} = \$149.9$ million, which represents a time-value-of-money loss of $22.1 million. Because the baseline likelihood of starting construction on a positive NPV mining project is 5.01%, firms with large self-bond reclamation liabilities face an expected loss each year of $1.1 million. The present value of holding this liability for 20 years (typical life of mines in my sample) is $14.6 million or 4.00% of market value for the median producing firm.

4.3 Risky Liabilities

One of the main assumptions in Myers's (1977) model is that the firm's existing liability is *risky* debt. Thus it seems that a debt overhang effect should be concentrated in firms with riskier liabilities. Once again, studying the mining industry offers an inherent advantage—the energy sector is extremely volatile. For example, S&P Global Ratings (2019) reports that Energy & Natural Resource's weighted average default rate from 1981-2018 is over 3%, placing it just behind Leisure Time/Media for highest defaulting industry. The fact that oil & gas and mining, in general, are prone to default suggests that a large percentage of the firms in my sample do indeed have risky liabilities.

Table 7 considers this prediction from Myers's model more directly. I take two approaches to mitigate the impact of safe cash flows and liabilities on my results. First, for each of my main tests (Model (4) from Table 4, and Model (2) from Panels A and B in Table 5), I exclude any firm that receives an investment grade security rating at any point in my sample period. These results appear in Models (1), (3), and (5). Second, I interact each overhang measure from my main tests with an indicator variable that equals one in the year of, and the year prior to, a security rating downgrade from either S&P or Moody's.[28] The interaction results appear in Models (2), (4), and (6).

Models (1), (3), and (5) of Table 7 very closely resemble the main tests using the full sample of firms as the coefficient on self-bonded reclamation liabilities is negative and significant at the 1% level in each. The point estimates are each slightly larger than those with the full sample, suggesting that the safest firms are mitigating the average effect of overhang from self-bonded reclamation liabilities. Models (2), (4), and (6) indicate that the likelihood for firms with self-bonded reclamation liabilities to acquire new rights or begin construction on NPV positive mining projects is significantly lower in downgrade periods. This effect is quite large. For example, the interaction term in Model (4) is -0.368, nearly 7 times as large as the average effect of -0.054.

The results in Table 7 confirm two points. First, the mining industry overall is considerably risky. For example, only a handful of firms earn an investment grade security rating over any sample period used. Further, while the interaction of self-bonded reclamation liabilities with the downgrade period is negative and significant, self-bonded reclamation liabilities on their own are significantly related to underinvestment. Second, the results are consistent with the hypothesis that overhang is more pronounced for riskier liabilities.[29]

5 External Validity

Studies on firm decisions from niche industries like resource extraction reasonably raise questions on external validity. My use of mine reclamation liabilities as an instrument to study debt overhang may amplify these concerns. However, while mine reclamation liabilities are certainly unique to resource extraction, other general liabilities exist that could induce similar investment behavior. In fact, Rauh (2006), and Arena and Julio (2015) and Bennett

[28]The rationale for using the year prior to, as well as the year of, a rating downgrade is that most times the downgrade significantly lags the precipitating event.

[29]Internet Appendix Table IA7 shows similar results for both U.S. samples discussed in the following section.

et al. (2018) find results consistent with an overhang effect when studying pension and legal liabilities, respectively.

Furthermore, resource extraction and in particular, resource extraction in Canada is a nontrivial economic sector. Worldwide, resource extraction is a multitrillion dollar industry and Section 2.1 argues that the mining in industry in Canada is among the world's largest. This mitigates a portion of external validity concerns. However, to go a step further, this section reports the results of several additional empirical tests exploring the effect reclamation liabilities have on firm investment.

5.1 Replicating the Negative Correlation Between Liabilities and Capital Expenditures

The first set of tests aims to use my sample on Canadian mining firms to replicate past results in the literature that have examined the effect of a firm's traditional debt liabilities on its capital expenditures. A firm's capital expenditures has become a standard proxy for investment in this literature (E.g., see Lang et al. (1996), Aivazian et al. (2005), Ahn et al. (2006), and Cai and Zhang (2011)). Each model includes year and firm fixed effects and robust standard errors are clustered at the firm level.

Model (2) of Table 8 examines the effect of debt, self-bonded and externally-bonded reclamation liabilities on the level of capital expenditures, including the standard control variables. Similar to prior results, the market debt leverage ratio is negatively related to capital expenditures and is significant at the 5% level. The coefficients on both self-bonded and externally-bonded reclamation liabilities are also negative and significant at the 1% level, although the size of the externally-bonded coefficient is an order of magnitude smaller than those of self-bonded reclamation or traditional debt liabilities. Next, Model (3) includes an interaction term between market leverage and Tobin's Q. The coefficient on this term is negative and significant at the 5%, suggesting that the decrease in capital expenditures among highly levered firms is concentrated in high growth firms. This result is consistent with prior studies and is often interpreted as evidence of debt overhang.

Overall, the results in Table 8 are consistent with panel regression results found in a typical U.S. industrial sample. Market leverage is negatively correlated with capital expenditures and this effect seems to be concentrated in high growth firms, even though in the same sample, traditional debt is unrelated to a firm's investment in positive NPV mining projects. These results imply that previous results showing a decrease in proxies for investment (CapEx, etc.) may be identifying a decrease in firms' overall investment opportunity sets, or a decrease in *negative* NPV projects. For example, it could be the case that financial covenants restrict Jensen's (1986) free cash flow problem. Additionally, these results provide some measure of reassurance that the results in Section 4 are not simply due to a unique sample.

5.2 Overhang and Investment in the U.S.

This set of empirical tests takes advantage of two different samples from the U.S. First, some mining firms voluntarily disclose the amount of their reclamation liabilities and the method in which these liabilities are bonded. I hand-collect this information from firms' 10-Ks and other reports over a sample period of 1992-2016. This yields just over 40 firms and around 350 firm-year observations. Second, I used data from the U.S. Mine Safety and Health Administration (MSHA) that reports mine-level information such as mine location and status through time. This sample yields nearly 5,000 mining firms, including 120 I can match to Compustat, over a period of 1983-2016.

Table 9 displays the results of these regressions where the dependent variable is capital expenditures in Models (1) through (6), and the incidence of a new mine in Model (7). Models (1) and (2) report the results using the firm's leverage liability ratios constructed from the hand-collected data. Consistent with earlier results, the coefficient on SB/MV is negative and significant in both cases. Models (3) and (4) confirms the negative relation using an indicator variable for firms that disclose any use of self-bonds. Finally, Models (5) through (7) use the MSHA data to create a measure that counts the number of self-bonded mines owned by each firm and show that this measure is negatively related to CapEx (Models (5) and (6)) and the likelihood a firm starts a new mine (Model (7)).

The results in Table 9 are consistent with the idea that self-bonded reclamation liabilities negatively impact firm investment and that this result is not unique to Canadian resource extraction firms. Furthermore, the results help ease concerns about the assumptions made for the identification of self-bonding reclamation liabilities in Section 4.

6 Conclusion

Debt overhang is a clearly modeled inefficiency that plays a central role in capital structure theory. However, contracting and debt composition mechanisms exist that could make debt overhang difficult to identify empirically. In addition, fully identifying the effects of overhang requires observing the firm's opportunity set. To mitigate these identification challenges, I exploit novel data on resource extraction that provides ex ante NPV estimates and firms that carry two major types of liabilities, traditional debt and reclamation liabilities, each with different costs associated with avoiding the overhang problem.

Consistent with debt overhang, I find that firms' investment decisions are significantly affected by the overhang imposed by unsecured mine reclamation liabilities. In particular, firms with such liabilities are more likely to forego the acquisition of new positive NPV mining rights, and to postpone construction in existing positive NPV mining projects than firms without such liabilities.

Firms' traditional debt, in contrast, is unrelated to investment in such positive NPV projects, consistent with the proposition that contracting and debt composition mechanisms exist that enable firms to avoid the debt overhang problem. This in true even in a sample in which firms' leverage ratios are negatively correlated with capital expenditures. Together, these results suggest that previous studies which use capital expenditures as a proxy for investment may identify a decrease in the overall opportunity set or even a decrease in negative NPV projects.

My findings imply that traditional debt, by itself, imposes few overhang-related investment distortions. This does not mean, however, that debt overhang is unimportant. Rather, my unique settings highlights exactly how important debt overhang is in capital structure decisions and why such effective solutions have endogenously arisen to mitigate it. Specifically, the overhang imposed by mine reclamation liabilities suggests that the costs of these ex ante solutions for traditional debt could be as large as 6.27% of firm value.

References

Ahn, Seoungpil, David J. Denis, and Diane K. Denis, 2006, Leverage and investment in diversified firms, *Journal of Financial Economics* 79, 317–337.

Aivazian, Varouj A., and Jeffrey L. Callen, 1980, Corporate leverage and growth, *Journal of Financial Economics* 8, 379–399.

Aivazian, Varouj A., Ying Ge, and Jiaping Qiu, 2005, The impact of leverage on firm investment: Canadian evidence, *Journal of Corporate Finance* 11, 277–291.

Akey, Pat, and Ian Appel, 2018, The limits of limited liability: Evidence from industrial pollution, Working paper, University of Toronto.

Allison, Paul D., 2002, Bias in fixed effects Cox regressions with dummy variables, Working paper, University of Pennsylvania.

Angrist, Joshua D., and Jörn-Steffen Pischke, 2008, Mostly harmless econometrics: An empiricist's companion, Princeton University Press, Princeton, NJ.

Arena, Matteo, and Brandon Julio, 2015, The effects of securities class action litigation on corporate liquidity and investment policy, *Journal of Financial and Quantitative Analysis* 50, 251–275.

Barclay, Michael J., Leslie M. Marx, and Clifford W. Smith, Jr., 2003, The joint determination of leverage and maturity, *Journal of Corporate Finance* 9, 149–167.

Barclay, Michael J., and Clifford W. Smith, Jr., 1995, The maturity structure of corporate debt, *Journal of Finance* 50, 609–631.

Bennett, Benjamin, Todd Milbourn, and Zexi Wang, 2018, Corporate investment under the cloud of litigation, Working paper, The Ohio State University.

Bernstein, Asaf, 2018, Negative equity, household debt overhang, and labor supply, Working paper, University of Colorado at Boulder.

Billett, Matthew T., Tao-Hsien Dolly King, and David C. Mauer, 2007, Growth opportunities and the choice of leverage, debt maturity, and covenants, *Journal of Finance* 62, 697–730.

Bonogofsky, Alexis, Amanda Jahshan, Hilary Yu, Dan Cohn, and Margie MacDonald, 2015, *Undermined promises II*, Joint publication by the National Wildlife Federation, the Natural Resources Defense Council, and the Western Organization of Resource Councils.

Bureau of Land Management, 2005, *BLM Nevada 3809 reclamation bonding guidelines*, Prepared by U.S. Department of the Interior, Washington D.C.

Cai, Jie, and Zhe Zhang, 2011, Leverage change, debt overhang, and stock prices, *Journal of Corporate Finance* 17, 391–402.

Cantor, Richard, 1990, Effects of leverage on corporate investment and hiring decisions, *Federal Reserve Bank of New York Quarterly Review (Summer)*, 31-41.

Chang, Xin, Kangkang Fu, Tao Li, Lewis H.K. Tam, and George Wong, 2018, Corporate environmental liabilities and capital structure, Working paper, Nangyang Technological University.

Chelimsky, Eleanor, 1988, *Surface mining: Cost and availability of reclamation bonds* (Report no. PEMD-88-17), Report to the Chairman, Subcommittee on Environment, Energy, and Natural Resources, Committee on Government Operations, House of Representatives, United States General Accounting Office, Washington, D.C.

Chen, Zhiyao, Jarrad Harford, and Avraham Kamara, 2018, Operating leverage, profitability, and capital structure, *Journal of Financial and Quantitative Analysis*, forthcoming.

Cheng, Linlin, and Jeffrey G. Skousen, 2017, Comparisons of international mine reclamation bonding systems with recommendations for China, *International Journal of Coal Science & Technology* 4, 67–79.

Childs, Paul D., David C. Mauer, and Steven H. Ott, 2005, Interactions of corporate financing and investment decisions: The effects of agency conflicts, *Journal of Financial Economics* 76, 667–690.

Chu, Yongqiang, 2016, Debt renegotiation and debt overhang: Evidence from lender mergers, Working paper, University of South Carolina.

Custódio, Cláudia, Miguel A. Gerreira, and Luís Laureano, 2013, Why are US firms using more short-term debt?, *Journal of Financial Economics* 108, 182–212.

Denes, Matthew, 2017, When do firms risk shift? Evidence from venture capital, Working paper, Carnegie Mellon University.

Diamond, Douglas W., 1991, Debt maturity structure and liquidity risk, *Quarterly Journal of Economics* 106, 709–737.

Diamond, Douglas W., 1993, Seniority and maturity of debt contracts, *Journal of Financial Economics* 33, 341–368.

Diamond, Douglas W., and Zhiguo He, 2014, A theory of debt maturity: The long and short of debt overhang, *Journal of Finance* 69, 719–762.

Ferguson, Andrew, Alexey Feigin, and Stephen Kean, 2013, Gold mine feasibility study disclosure in Australia: Determinants and implications, *Resources Policy* 38, 8–17.

Ferguson, Andrew, and Gabriel Pündrich, 2015, Does industry specialist assurance of non-financial information matter to investors?, *Auditing: A Journal of Practice & Theory* 34, 121–146.

Fox, Kenneth A., 2017, The usefulness of NI 43-101 technical reports for financial analysts, *Resources Policy* 51, 225–233.

Frank, Murray Z., and Vidhan K. Goyal, 2009, Capital structure decisions: Which factors are reliably important?, *Financial Management* 38, 1–37.

Gertner, Robert, and David Scharfstein, 1991, A theory of workouts and the effects of reorganization law, *Journal of Finance* 46, 1189–1222.

Gilje, Erik P., 2016, Do firms engage in risk-shifting? Empirical evidence, *Review of Financial Studies* 29, 2925–2954.

Giroud, Xavier, Holger M. Mueller, Alex Stomper, and Arne Westerkamp, 2012, Snow and leverage, *Review of Financial Studies* 25, 680–710.

Gorton, William T., 2009, Understanding the reclamation surety relationship before and after operator default, Report to the Rocky Mountain Mineral Law Foundation Special Institute on Mine Closure, Financial Assurance, and Final Reclamation, Westminster, CO.

Guedes, Jose, and Tim Opler, 1996, The determinants of the maturity of corporate debt issues, *Journal of Finance* 51, 1809–1833.

Hein, Jayni F., Ben Snow, Sean Stefanik, and Lauren Webb, 2016, Self-bonding in an era of coal bankruptcy: Recommendations for reform, *Institute for Policy Integrity*, NYU School of Law.

Hennessy, Christopher A., 2004, Tobin's Q, debt overhang, and investment, *Journal of Finance* 59, 1717–1742.

Interstate Mining Compact Commission, 2014, *Self-bonding survey.*

Jensen, Michael C., 1986, Agency costs of free cash flow, corporate finance, and takeovers, *American Economic Review Papers & Proceedings* 76, 323–329.

Jensen, Michael C., and William H. Meckling, 1976, Theory of the firm: Managerial behavior, agency costs and ownership structure, *Journal of Financial Economics* 3, 305–360.

Johnson, Shane A., 2003, Debt maturity and the effects of growth opportunities and liquidity risk on leverage, *Review of Financial Studies* 16, 209–236.

Kalbfleisch, John D., and D.A. Sprott, 1970, Application of likelihood methods to models involving large numbers of parameters, *Journal of the Royal Statistical Society* 32, 175–208.

Karpoff, Jonathan M., and Michael D. Wittry, 2018, Institutional and legal context in natural experiments: The case of state antitakeover laws, *Journal of Finance* 73, 657–714.

Kirschner, Lisa A., and Edward B. Grady, 2003, Mining and the vanishing surety bond market, *Natural Resources & Environment* 17, 152–154, 187–189.

Kuipers, James R., 2000, Hardrock reclamation bonding practices in the Western United States, Report to the National Wildlife Federation, Boulder, CO.

Lang, Larry, Eli Ofek, and René M. Stulz, 1996, Leverage, investment, and firm growth, *Journal of Financial Economics* 40, 3–29.

Mauer, David C., and Steven H. Ott, 2000, Agency costs, underinvestment, and optimal capital structure: The effect of growth options to expand, in Michael J. Brennan and Lenos Trigeorgis, eds.: *Project Flexibility, Agency, and Competition* (Oxford University Press, New York, NY), 151-179.

Mella-Barral, Pierre, and William Perraudin, 1997, Strategic debt service, *Journal of Finance* 52, 531–556.

Mello, Antonio S., and John E. Parsons, 1992, Measuring the agency cost of debt, *Journal of Finance* 47, 1887–1904.

Melzer, Brian T., 2017, Mortgage debt overhang: Reduced investment by homeowners at risk of default, *Journal of Finance* 72, 575–612.

Merton, Robert C., 1974, On the pricing of corporate debt: The risk structure of interest rates, *Journal of Finance* 29, 449–470.

Moyen, Nathalie, 2007, How big is the debt overhang problem?, *Journal of Economic Dynamics & Control* 31, 433–472.

Myers, Stewart C., 1977, Determinants of corporate borrowing, *Journal of Financial Economics* 5, 147–175.

Nazzaro, Robin N., 2005, *Hardrock mining: BLM needs to better manage financial assurances to guarantee coverage of reclamation costs* (Report no. GAO-05-377), Report to the Ranking Minority Member, Committee on Homeland Security and Governmental Affairs, U.S. Senate, United States Government Accountability Office, Washington D.C.

New Generation Research, Inc., 2016, *Energy sector bankruptcies: A revew of bankruptcy activity in the energy sector: 2015-2016.*

Nikolaev, Valeri V., 2017, Scope for renegotiation in private debt contracts, *Journal of Accounting and Economics*, forthcoming.

Ontario Securities Commission, 2011, Form 43-101F1 Technical Report.

Opler, Tim C., and Sheridan Titman, 1994, Financial distress and corporate performance, *Journal of Finance* 49, 1015–1040.

Parrino, Robert, and Michael S. Weisbach, 1999, Measuring investment distortions arising from stockholder-bondholder conflicts, *Journal of Financial Economics* 53, 3–42.

Pawlina, Grzegorz, 2010, Underinvestment, capital structure, and strategic debt restructuring, *Journal of Corporate Finance* 16, 679–702.

Rauh, Joshua D., 2006, Investment and financing constraints: Evidence from the funding of corporate pension plans, *Journal of Finance* 61, 33–71.

Rauh, Joshua D., 2009, Risk shifting versus risk management: Investment policy in corporate pension plans, *Review of Financial Studies* 22, 2687–2733.

Richer La Flèche, Eirk, ed., 2016, *The Mining Law Review*, Law Business Research, London, UK.

Richer La Flèche, Erik, David Masse, and Jennifer Honeyman, 2016, Chapter 24: Canada capital markets, in Erik Richer La Flèche ed.: *The Mining Law Review* 5 (Law Business Research Ltd, London, UK), 264-273.

Roberts, Michael R., 2015, The role of dynamic renegotiation and asymmetric information in financial contracting, *Journal of Financial Economics* 116, 61–81.

Roberts, Michael R., and Amir Sufi, 2009, Renegotiation of financial contracts: Evidence from private credit agreements, *Journal of Financial Economics* 93, 159–184.

Sharpe, Steven A., 1991, Credit rationing, concessionary lending, and debt maturity, *Journal of Banking and Finance* 15, 581–604.

Sharpe, Steven A., 1994, Financial market imperfections, firm leverage, and the cyclicality of employment, *American Economic Review* 84, 1060–1074.

Socalar, Milton J., 1988, *Comments on OSMRE authority to waive or negotiate reclamation fees* (Report no. B-229068.4), Report prepared for the Department of the Interior's Office of Surface Mining Reclamation and Enforcement, Comptroller General of the United States, Washington, D.C.

S&P Global Ratings, 2019, *2018 annual global corporate default and ratio transition study.*

Strebulaev, Ilya A., and Baozhong Yang, 2013, The mystery of zero-leverage firms, *Journal of Financial Economics* 109, 1–23.

Stulz, René M., and Herb Johnson, 1985, An analysis of secured debt, *Journal of Financial Economics* 14, 501–521.

Titman, Sheridan, and Sergey Tsyplakov, 2007, A dynamic model of optimal capital structure, *Review of Finance* 11, 401–451.

Walsh, Dan, 2017, *Final Bond Determination*, Report prepared for Golden Sunlight Mines, Inc., Montana Department of Environmental Quality, Helena, MT.

Whited, Toni M., 1992, Debt, liquidity constraints and corporate investment: Evidence from panel data, *Journal of Finance* 47, 1425–1460.

Xu, Qiping, 2017, Kicking maturity down the road: Early refinancing and maturity management in the corporate bond market, *Review of Financial Studies* 31, 3061–3097.

Figure 1: SB vs EB example. This figure displays three coal mines on the border of British Columbia and Alberta, Canada as an example of the specifics of my empirical strategy. The data on estimated reclamation liabilities were hand-collected from firms' public disclosures. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1.

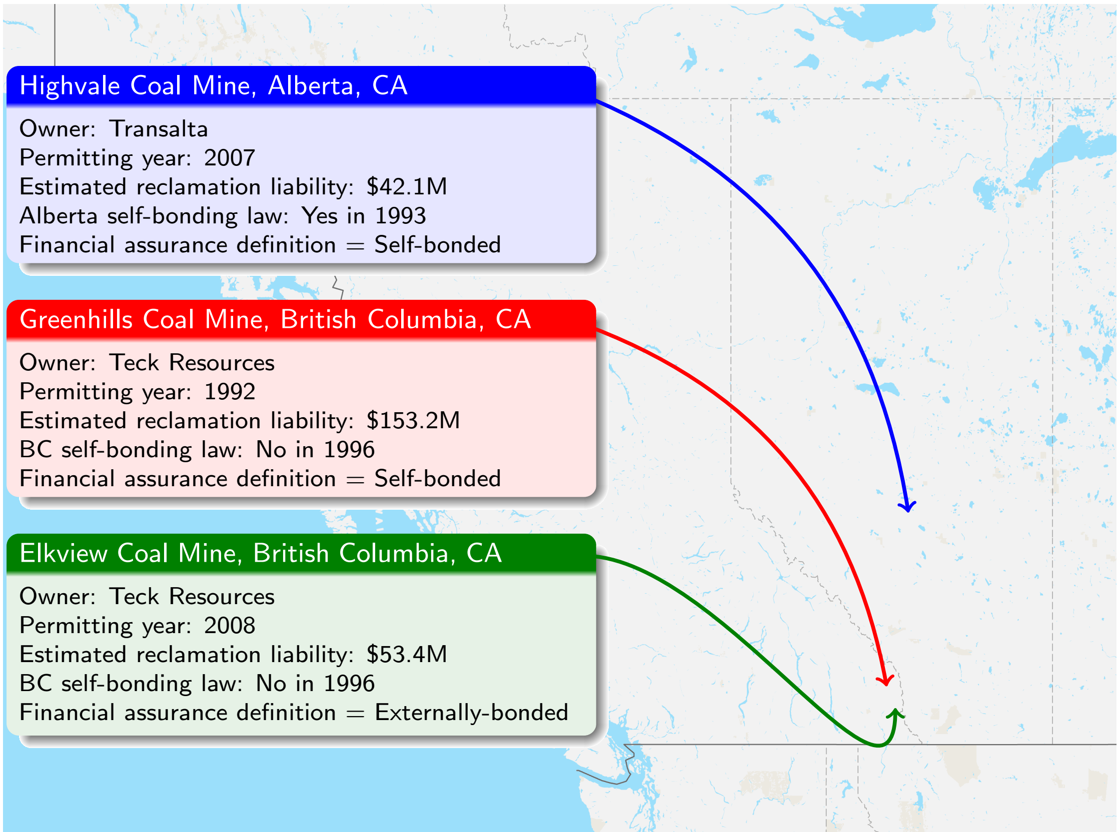


Figure 2: The life cycle of a typical mining project. This figure displays the life cycle of a typical mining project. The source for this figure is KPMG International 2012. This figure also highlights the main milestones at each stage for a mining project. The mines in my sample match very closely with the ranges given for the various stages. The median and mean numbers for mining projects spent in exploration (prospecting and exploration) in my sample are 6 and 7.5 years, respectively. The median and mean numbers for mining projects spent in evaluation (feasibility) in my sample are 5 and 5.2 years, respectively. The median and mean numbers for mining projects spent in development (construction/permitting) in my sample are 3 and 4.75 years, respectively. The median and mean numbers for mining projects spent in production in my sample are 12 and 19.9 years, respectively. Finally, the median and mean numbers for mining projects spent in closure in my sample are 10 and 11.1 years, respectively. It is, however, very difficult to discern when a mining project finishes the closure stage.

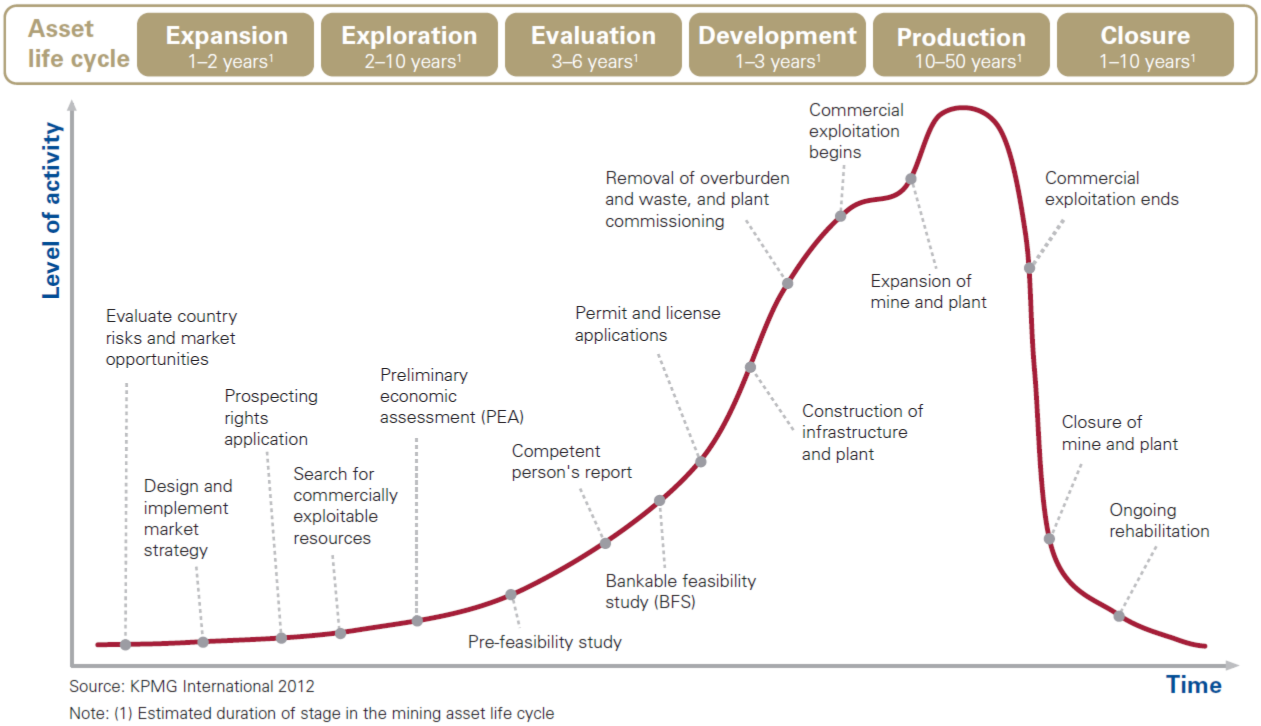


Figure 3: Survival analysis. This figure displays Kaplan-Meier survival functions. The sample is comprised of project-year observations for mining projects from the year the firm publicly discloses the NI 43-101 technical report that includes the initial NPV estimate to the year the firm begins construction on the mine, or the sample period ends, whichever comes first. The sample consists of firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada over the sample period of 2003 to 2016. A project experiences an "event" when it begins construction. Solid lines represent the observed non-parametric survival functions while dashed lines depict the predicted functions. Each sub-figure is split on whether the liability—self-bonded reclamation liabilities, externally-bonded reclamation liabilities, or total debt—exceeds the initial NPV estimate. The data on estimated reclamation liabilities was hand-collected from firms' public disclosures. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1.

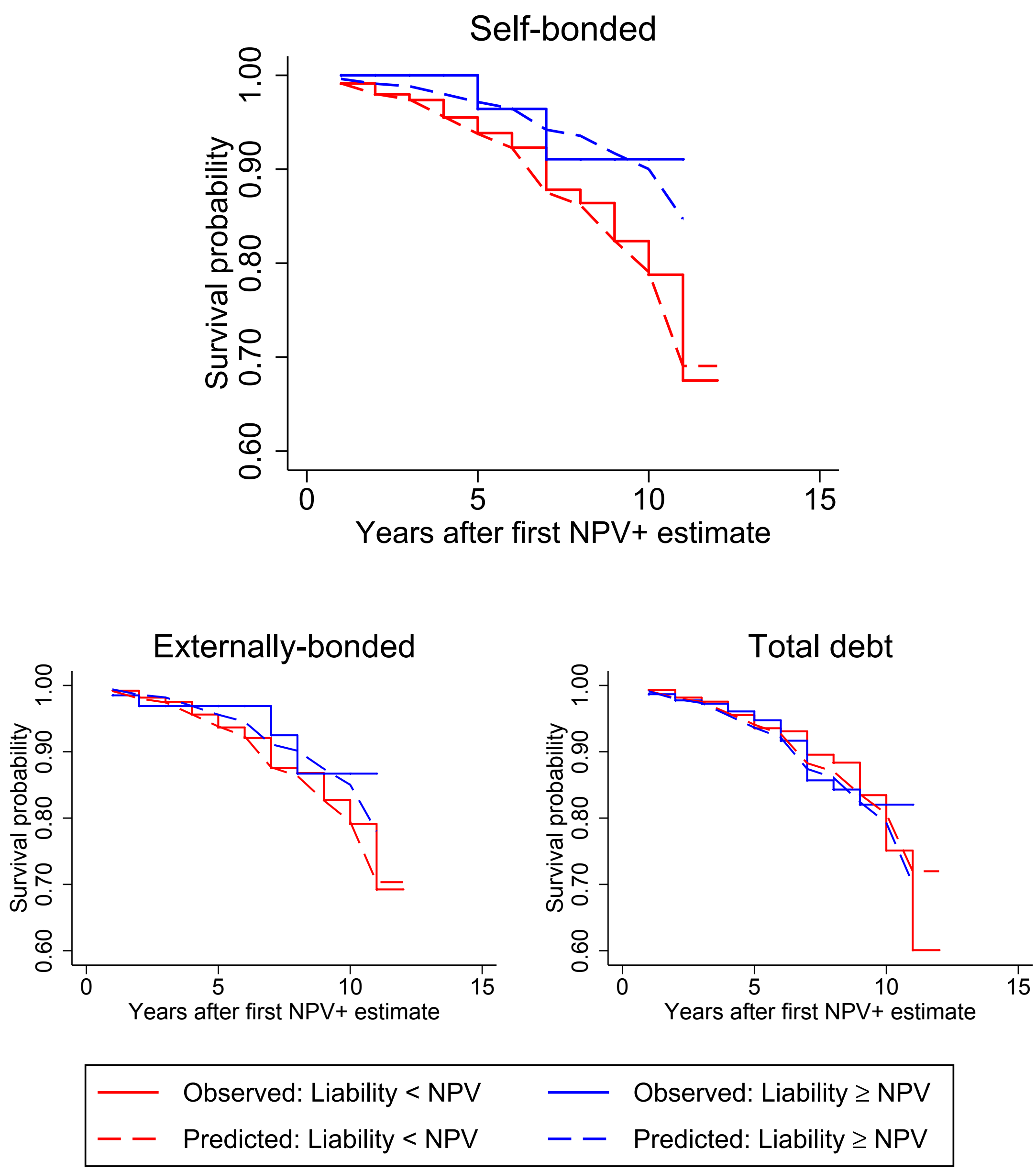


Table 1: Self-bonding mining regulations. This table displays the self-bonding regulations for the United States (panel A) and the rest of the world (panel B). The sources for the mining regulations are the various annotated state statutes and country legislative acts. While this is not an exhaustive list, the regulations listed in this table cover over 90% of the permitted mines in my sample. For the purposes of this paper, states and countries without a bonding regulation, and those time periods before a law is passed to allow (or disallow) self-bonding, are treated as if they explicitly allow self-bonding.

Panel A: United States

	Coal Mining			**Hardrock and Metal Mining**		
State	**Federal or State Statute**	**Year**	**Self-bonding allowed**	**Federal or State Statute**	**Year**	**Self-bonding allowed**
All	Surface Mining Control and Reclamation Act (SMCRA)	1977	Yes	43 C.F.R. §3809.555, .571(c)	2001	No
Alabama	ALA. CODE §§9-16-89(c)	1982	Yes	ALA. CODE §§9-16-8	1969	No
Alaska	ALASKA STAT. §§27.21.160	1982	Yes	ALASKA STAT. §§27.19.040(e)	1990	Yes
Arizona				ARIZ. REV. STAT. §§27-991	1994	Yes
Arkansas	ARK. CODE ANN. §§15-58-509	1979	Yes	ARK. CODE ANN. §§15-57-317	1991	Yes
California				CAL. PUB. RES. CODE §2773.1	1975	Yes
Colorado	COLO. REV. STAT. §34-33-113(3)	1979	Yes	COLO. REV. STAT. §34-32-117(3)(f)	1977	Yes
Delaware				DEL. CODE ANN. tit. 7 §6115	1953	Yes
Florida				FLA. STAT. ANN. §§253.571	1969	Yes
Georgia				GA. CODE ANN. §§12-4-75(3)	1968	No
Hawaii	HAW. REV. STAT. §§182-3	1963	Yes	HAW. REV. STAT. §§182-3	1963	Yes
Idaho				IDAHO CODE §47-1512 (surface)	1955	No
				IDAHO CODE §47-1317 (dredge)	1971	Yes
Illinois	225 ILL. COMP. STAT. 720/6.01(b)	1981	Yes	225 ILL. COMP. STAT. 715/8	1983	No
Indiana	IND. CODE 14-34-7	1982	Yes	IND. CODE 14-36-1-24	1995	Yes
Iowa	IOWA CODE §207.10	1985	No	IOWA CODE §208.23	1985	No
Kansas	KAN. STAT. ANN. §§49-615(a)	1994	No	KAN. STAT. ANN. §§49-615(a)	1994	No
Kentucky	KY. REV. STAT. ANN. §350.064(2)	1980	No			
Louisiana	LA. REV. STAT. ANN. §§30:909(c)	1978	No			
Maine				ME. REV. STAT. ANN. tit. 38 §490-RR(3)	1979	Yes
Maryland	MD. CODE ANN., ENVIR. §§15-612	1976	No	MD. CODE ANN., ENVIR. §§15-507, 15-823	1975	No
Michigan				MICH. COMP. LAWS §§324.63211	1968	Yes
Minnesota				MINN. STAT. §§93.49	1969	Yes
Mississippi	MISS. CODE ANN. §§53-9-31	1979	Yes	MISS. CODE ANN. §§53-7-37	1977	Yes
Missouri	MO. ANN. STAT. §§444.950	1982	Yes	MO. ANN. STAT. §§444.368 (metals)	1989	Yes
				MO. ANN. STAT. §§444.778 (other)	1971	No
Montana	MONT. CODE ANN. §§82-4-223	1971	No	MONT. CODE ANN. §§82-4-338	1971	Yes
Nebraska						
Nevada				NEV. REV. STAT. §§519A.160(4)	1989	Yes
New Hampshire				N.H. REV. STAT. §12-E:6	1979	Yes
New Mexico	N. M. STAT. ANN. §§69-25A-13	1979	Yes	N. M. STAT. ANN. §§69-36-7(q)	1994	No
New York				N.Y. ENVTL. CONSERV. LAW §§23-2701	1976	Yes
North Carolina				N.C. GEN. STAT. §§74-54	1971	Yes
North Dakota	N.D. CENT. CODE §§38-14.1-16.8	1979	Yes	N.D. CENT. CODE §§38-14.1-16.8	1979	Yes
Ohio	OHIO REV. CODE ANN. §§1513.01(W)	2000	Yes	OHIO REV. CODE ANN. §§1514.04	2002	No
Oklahoma	OKLA. STAT. tit. 45 §§745.6(e)	1979	Yes	OKLA. STAT. tit. 45 §§728(e)	1982	No
Oregon				OR. REV. STAT. §§517.810	1971	Yes
Pennsylvania	25 PA. CONS. STAT. §§86.159	1982	Yes	25 PA. CONS. STAT. §§77.222	1990	No
South Carolina	S.C. CODE ANN. §§48-20-110	1990	Yes	S.C. CODE ANN. §§48-20-110	1990	Yes
South Dakota				S.D. CODIFIED LAWS §§45-6B-23	1982	No
Tennessee	TENN. CODE ANN. §§59-8-408	1987	No	TENN. CODE ANN. §§59-8-207	1972	No
Texas	TEX. NAT. RES. CODE ANN. §§134.123	1995	Yes	TEX. NAT. RES. CODE ANN. §§134.123	1995	Yes
Utah	UTAH CODE ANN. §§40-10-15(3)	1979	Yes	UTAH CODE ANN. §§40-8-14(3)	1975	Yes
Virginia	VA. CODE ANN. §§45.1-241	2014	No	VA. CODE ANN. §§45.1-183	1968	No
Washington				WASH. REV. CODE §§78.56.110(1) and 78.44.087(3)	1994	No
West Virginia	W. VA. CODE §§22-3-11(c)	1987	Yes			
Wisconsin				WIS. STAT. §§295.59 and 293.51	1973	No
Wyoming	WYO. STAT. ANN. §§35-11-417	1973	Yes	WYO. STAT. ANN. §§35-11-417	1973	Yes

Table 1—*Continued*

Panel B: Non-US Countries

Country	State or Province	Regulation	Year	Bonding Regulation	Self-bonding allowed
Argentina		1997 Mining Code	1997	No	
Australia		Mining regulations at the state level			
	New South Wales	Protection of the Environment Operations Act 1997 No 156 Part 9.4	1997	Yes	Yes
	Queensland	Environmental Protection Act 1994 (EM1010)	1994	Yes	No
	South Australia	Opal Mining Act	1995	Yes	Yes
	Tasmania	Mineral Resources Development Act 1995 Sec 14, 53, and 75	1995	Yes	No
	Victoria	Mineral Resources (Sustainable Development) Act 1990 (MR(SD)A) Sec 78	1990	Yes	No
	Western Australia	Mining Act 1978 Sec 126	1978	Yes	Yes
Bolivia		1997 Mining Code	1997	No	
Botswana		Mines and Minerals act of 1999	1999	Yes	Yes
Burkina Faso		The Mining Code, Article 12 of Decree No. 2017-068	2017	Yes	No
Brazil		NRM 20 and DN 127	2001, 2009	No	
Canada		Mining regulations at the province level			
	Alberta	Conservation and Reclamation Regulation Sec 21	1993	Yes	Yes
	British Columbia	Bonding Act (RSBC 1996) Chapter 30	1996	Yes	No
	Manitoba	Mine Closure Regulation, 1999 (Mines and Mineral Act)	1999	Yes	Yes
	New Brunswick	Mining Act	1989	Yes	No
	Newfoundland	Mining Act Chapter M-15.10	1999	Yes	No
	Northwest Territories	NWT Waters Act	1992	Yes	No
	Nova Scotia	Minerals Resources Act S.N.S. 1990, c.18 Sec 77	1990	Yes	Yes
	Nunavut	Nunavut Water Regulations Sec 10.3	2013	Yes	No
	Ontario	Ontario Mining Act	2000	Yes	Yes
	Quebec	Quebec Mining Act	1998	Yes	Yes
	Saskatchewan	The Mineral Industry and Environmental Protection Regulations	1996	Yes	No
	Yukan Territory	Yukon Water Act	1992	Yes	Yes
Chile		1983 Mining Code	1983	No	
Colombia		Law No. 685 (The Mining Code)	2001	No	
Congo		Law No. 007/2002 (The Mining Code)	2002	Yes	No
Dominican Republic		Environmental Law (Law No. 64-00)	2000	Yes	No
Ethiopia		Environmental Protection Authority Establishment Proclamation No. 9/1995	1995	Yes	No
Finland		Environmental Protection Act Section 43a (647/2011)	2011	Yes	No
Ghana		EPA Act 494 and LI 1652	1999	Yes	No
Indonesia		R.I. Government Regulation No. 78	2010	Yes	No
Mali		The Mining Law	1999	No	
Mexico		NOM-141-SEMARNAT-2003	2003	Yes	No
Mongolia		Mongolia Minerals Law	1997	Yes	No
Panama		Code of Mineral Resources, Law 13 of 2012	2012	Yes	No
Papua New Guinea		Mining Act of 1992	1992	No	
Peru		Law No. 28090, Law that Rules the Closing of Mines (Ley que Regula el Cierre de Minas)	2003	Yes	No
Philippines		DENR Administrative Order No. 2010-21 (Mining Act IRR)	2010	Yes	No
South Africa		National Environment Management Act (NEMA)	1998	Yes	No
Tanzania		The Mining Act	2010	Yes	Yes
Turkey		Regulation on Reclamation of Lands Disturbed by Mining	2007	No	
Vietnam		Mineral Law	1996	Yes	No
Zambia		Mines and Minerals Development Act No. 7 of 2008	2008	Yes	No
Zimbabwe		Environmental Management Act [Chapter 20:27]	2002	Yes	No

Table 2: Mine-level summary statistics. This table displays summary statistics for mines owned by firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada during the period 1990 to 2016. Observations in Panel A are at the mine-year level, and observations in Panel B and Panel C are at the mine-level. Data on firms' mines, including mine status, mine type, primary mineral extracted, mine location, and project-level data on estimated NPV, capital costs, discount rates, and mine life were provided by Mining Intelligence. Other mining data, including estimated reclamation liabilities and the surface area to be reclaimed were hand-collected from firms' public disclosures. Self-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1. All monetary variables are reported in U.S. dollars, using historical exchange rates from OFX when values are reported in other currencies.

Panel A: All mining projects (mine-year observations)					
Variable	**Obs.**	**Mean/ Percentage**	**Median**	**Min**	**Max**
Mine status (in %)					
Prospect/exploration	22,379	0.623			
Feasibility	22,379	0.041			
Construction/Permitting	22,379	0.015			
Production	22,379	0.142			
Closed	22,379	0.027			
Panel B: Permitted mines (mine observations)					
Surface area to be reclaimed (in km^2)	580	106.5	42.5	0.04	2663
Estimated reclamation liabilities (in $Ms)	580	27.7	6.6	0	558
Self-bonded mines (in %)	580	0.398			
Estimated self-bonded reclamation liabilities (in $Ms)	188	24.2	6.5	0.0	400
Mine type (in %)					
Open-pit or surface	580	0.553			
Underground	580	0.436			
Primary mineral extracted (in %)					
Gold	580	0.478			
Copper	580	0.083			
Coal	580	0.071			
Silver	580	0.045			
Zinc	580	0.045			
Uranium	580	0.043			
Other or combination	580	0.236			
Mine location (in %)					
Canada	580	0.288			
United States	580	0.200			
Mexico	580	0.112			
Australia	580	0.047			
Chile	580	0.036			
Brazil	580	0.029			
Other	580	0.290			

Table 2—*Continued*

Panel C: Projects with Estimated NPV (mine observations)					
Variable	**Obs.**	**Mean/ Percentage**	**Median**	**Min**	**Max**
Initial NPV estimate ($Ms)	269	402.4	172	-48.9	7114.6
Initial capital costs estimate ($Ms)	269	535.7	223	1.2	7899.0
Discount rate (%)	269	6.8	7.5	5	15
Estimated mine life (years)	269	14.1	11	1	50
Projects undertaken by 2016 (%)	269	0.283			
$\mathbb{1}_{\text{SB}\geq\text{NPV}}$	269	0.043	0	0	1
$\mathbb{1}_{\text{EB}\geq\text{NPV}}$	269	0.072	0	0	1
$\mathbb{1}_{\text{TD}\geq\text{NPV}}$	269	0.177	0	0	1
Acquisition of mining rights					
Total acquisition cost ($Ms)	191	38.5	9.1	0.9	532.0
NPV of mining rights ($Ms)	191	200.6	68.5	-115.8	1129.0
Primary mineral extracted (in %)					
Gold	269	0.442			
Copper	269	0.171			
Silver	269	0.041			
Uranium	269	0.041			
Other or combination	269	0.305			
Mine location (in %)					
Canada	269	0.428			
United States	269	0.112			
Mexico	269	0.112			
Peru	269	0.045			
Other	269	0.303			

Table 3: Firm-level summary statistics. This table reports firm-level summary statistics for firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada over the sample period of 1990 to 2016. Data on firms' mines, including the number of mining projects, were provided by Mining Intelligence. The data on estimated reclamation liabilities, as well as the surface area to be reclaimed, were hand-collected from firms' public disclosures. *SB/MV* is the sum of estimated reclamation liabilities of all a firm's mines defined as self-bonded divided by the market value of the firm's assets. *EB/MV* is the sum of estimated reclamation liabilities of all a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) divided by the market value of the firm's assets. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1. Accounting data is from Compustat—North America. Variable definitions are located in Appendix A1. All monetary variables are reported in U.S. dollars, using historical exchange rates from OFX when values are reported in other currencies. All ratio variables are winsorized at the 1% and 99% levels.

Variable	**Obs.**	**Mean**	**Median**	**Min**	**Max**
Mining variables					
Estimated reclamation liabilities (in \$Ms)	7,986	10.2	0.0	0.3	1,609
SB/MV	7,079	0.010	0.000	0.000	10.540
EB/MV	7,079	0.058	0.000	0.000	72.410
Surface area to be reclaimed (km^2)	7,986	38.4	0.0	0.0	3,109
Self-bonded surface area (%)	7,986	0.07	0.00	0.00	1.00
Likelihood of acquiring new mining rights (all)	7,986	0.194	0	0	1
Likelihood of acquiring new mining rights (NPV+)	7,986	0.037	0	0	1
Prospect/exploration	7,986	1.6	0	0	73
Feasibility	7,986	0.1	0	0	4
Construction/permitting	7,986	0.0	0	0	6
Production	7,986	0.3	0	0	15
Closed	7,986	0.1	0	0	7
Mining variables (firms with producing mines)					
Estimated reclamation liabilities (in \$Ms)	1,182	68.7	11.9	0.3	1,609
SB/MV	1,094	0.062	0.004	0.000	10.540
EB/MV	1,094	0.373	0.009	0.000	72.410
Surface area to be reclaimed (km^2)	1,182	259.3	87.1	0.0	3,109
Self-bonded surface area (%)	1,182	0.48	0.34	0.00	1.00
Likelihood of acquiring new mining rights (all)	1,094	0.265	0	0	1
Likelihood of acquiring new mining rights (NPV+)	1,094	0.057	0	0	1
Prospect/exploration	1,094	2.9	1	0	73
Feasibility	1,094	0.2	0	0	4
Construction/permitting	1,094	0.2	0	0	6
Production	1,094	2.4	1	1	15
Closed	1,094	0.3	0	0	7
Accounting variables					
Capital expenditures (% of book assets)	7,498	0.132	0.081	0.000	0.824
Book value of assets (BV) (in \$Ms)	7,609	856.8	21.0	0.0	76,467
Market value of assets (MV) (in \$Ms)	7,079	925.0	23.3	0.0	61,511
Short-term debt (STD) (in \$Ms)	7,601	18.1	0.0	0.0	7,338
Long-term debt (LTD) (in \$Ms)	7,604	155.5	0.0	0.0	13,173
Market leverage (%)	7,079	0.109	0.000	0.000	0.911
Return on assets (ROA) (%)	7,516	-0.795	-0.098	-21.182	0.314
Cash (% of book assets)	7,603	0.243	0.134	0.000	0.988
Tobin's Q	7,077	3.3	1.1	0.1	64.5
Annual stock return (%)	6,685	0.435	-0.100	-0.907	11.750

Table 4: Firm liabilities and the likelihood of acquiring the rights to new positive NPV mining projects. This table reports the results of linear probability models in which the dependent variable is the likelihood a firm acquires the rights to a new positive net present value (NPV+) mining project. The NPV of the mining rights is defined as the value of the NPV estimate in the NI 43-101 technical reports discounted back to the acquisition year, less the cost the acquiring firm paid for the individual mine at acquisition. The sample consists of firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada over the sample period of 1990 to 2016. *SB/MV* is the sum of estimated reclamation liabilities of all a firm's mines defined as self-bonded divided by the market value of the firm's assets. *EB/MV* is the sum of estimated reclamation liabilities of all a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) divided by the market value of the firm's assets. The data on estimated reclamation liabilities were hand-collected from firms' public disclosures. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1. *Acquire mining rights (other)* is an indicator variable that equals 1 if a firm acquires rights to a mine that I cannot calculate an NPV number for. Accounting variables (defined in Appendix Table A1) are constructed using data from Compustat—North America and the data on firms' mining projects were provide by Intelligence Mining. Robust standard errors, clustered at the firm level, are reported in parentheses. *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable =	Likelihood of acquiring rights to any project		Likelihood of acquiring rights to NPV+ projects	
	(1)	(2)	(3)	(4)
SB/MV	-0.030***	-0.035**	-0.022***	-0.025***
	(0.009)	(0.015)	(0.008)	(0.006)
EB/MV	0.002	0.003	0.001	0.001
	(0.003)	(0.003)	(0.001)	(0.001)
Market leverage	-0.046	-0.057*	-0.014	-0.027
	(0.031)	(0.032)	(0.015)	(0.019)
Log of book assets		0.018***		0.004*
		(0.005)		(0.002)
Cash		0.013		0.010
		(0.023)		(0.011)
ROA		0.000		0.000
		(0.002)		(0.001)
Tobin's Q		0.001		0.000
		(0.001)		(0.000)
Log of firm age		0.001		-0.025**
		(0.021)		(0.011)
# of mines in exploration		0.026***		0.002
		(0.007)		(0.002)
# of mines in feasibility		0.015		0.013
		(0.023)		(0.018)
# of mines in construction/permitting		0.024		0.047**
		(0.030)		(0.019)
# of producing mines		0.008		0.014**
		(0.016)		(0.007)
# of closed mines		0.024		0.045**
		(0.030)		(0.020)
Aquire mining rights (other)				-0.071***
				(0.007)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Number of firms	790	775	790	775
Observations	7,083	6,747	7,083	6,747
R^2	0.275	0.301	0.128	0.166

Table 5: Firm liabilities and the likelihood of beginning construction on positive NPV projects. This table reports the results of linear probability models in which the dependent variable is the likelihood of beginning construction on a positive net present value (NPV+) project. The sample is comprised of project-year observations for mining projects from the year the firm publicly discloses the NI 43-101 technical report that includes the initial NPV estimate to the year the firm begins construction on the mine, or the sample period ends, whichever comes first. The sample consists of firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada over the sample period of 2003 to 2016. SB/MV is the sum of estimated reclamation liabilities of all of a firm's mines defined as self-bonded divided by the market value of the firm's assets. EB/MV is the sum of estimated reclamation liabilities of all of a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) divided by the market value of the firm's assets. $\mathbb{1}_{\mathrm{SB}\geq\mathrm{NPV}}$ is an indicator variable that equals 1 if the sum of estimated reclamation liabilities of all of a firm's mines defined as self-bonded exceeds the initial NPV estimate of the mining project. $\mathbb{1}_{\mathrm{EB}\geq\mathrm{NPV}}$ is an indicator variable that equals 1 if the sum of estimated reclamation liabilities of all of a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) exceeds the initial NPV estimate of the mining project. $\mathbb{1}_{\mathrm{TD}\geq\mathrm{NPV}}$ is an indicator variable that equals 1 if the firm's total debt exceeds the initial NPV estimate of the mining project. The data on estimated reclamation liabilities were hand-collected from firms' public disclosures. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1. Accounting variables (defined in Appendix Table A1) are constructed using data from Compustat—North America, the project data were provide by Intelligence Mining and futures data is from Bloomberg. Robust standard errors, clustered at the firm level, are reported in parentheses. *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Panel A: Liability ratios

Dependent variable =	Likelihood of beginning construction on an NPV+ project				
	(1)	(2)	(3)	(4)	(5)
SB/MV	-0.046***	-0.054***	-0.055***	-0.048**	-0.124***
	(0.014)	(0.014)	(0.016)	(0.020)	(0.024)
EB/MV	0.020	0.003	0.011	0.012	0.122
	(0.061)	(0.062)	(0.125)	(0.132)	(0.137)
Market leverage	0.010	-0.007	-0.030	-0.038	-0.070
	(0.058)	(0.056)	(0.057)	(0.060)	(0.069)
Log of book assets		0.018	0.023	0.023	0.017
		(0.014)	(0.016)	(0.017)	(0.019)
Cash		0.043	0.034	0.042	0.042
		(0.058)	(0.072)	(0.074)	(0.079)
ROA		-0.015	-0.021	-0.022	-0.018
		(0.012)	(0.016)	(0.016)	(0.017)
Tobin's Q		0.000	-0.001	-0.002	-0.002
		(0.005)	(0.007)	(0.007)	(0.008)
Log of firm age		-0.055	-0.049	-0.059	-0.037
		(0.064)	(0.074)	(0.078)	(0.074)
Project NPV ($100Ms)		0.014**	0.016*	0.013	0.012
		(0.007)	(0.009)	(0.009)	(0.008)
Project Capital Costs ($100Ms)		-0.014**	-0.012*	-0.009	-0.008
		(0.007)	(0.007)	(0.008)	(0.007)
Expected mine life		-0.002	-0.002	-0.004	-0.009
		(0.004)	(0.005)	(0.010)	(0.012)
Total NPV of alternative projects ($100Ms)		-0.014	-0.013	-0.013	-0.012
		(0.011)	(0.011)	(0.011)	(0.011)
Start alternative NPV+ project		-0.042	-0.032	-0.017	-0.013
		(0.111)	(0.117)	(0.114)	(0.113)
Primary mineral price (% ch.)			-0.001		
			(0.064)		
Futures price				0.007	-0.009
				(0.015)	(0.012)
Options-implied volatility					-0.058
					(0.065)
Controls for projects in each stage	No	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Primary mineral FE	No	No	No	Yes	Yes
Number of firms	177	174	143	140	126
Observations	838	822	679	662	589
R^2	0.289	0.299	0.294	0.296	0.306

Table 5—*Continued*

Panel B: Liability indicators					
Dependent variable =	Likelihood of beginning construction on an NPV+ project				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}_{\mathrm{SB}\geq\mathrm{NPV}}$	-0.232**	-0.264**	-0.269**	-0.243**	-0.378*
	(0.116)	(0.109)	(0.106)	(0.106)	(0.222)
$\mathbb{1}_{\mathrm{EB}\geq\mathrm{NPV}}$	0.013	-0.019	0.001	0.024	0.046
	(0.072)	(0.082)	(0.086)	(0.097)	(0.104)
$\mathbb{1}_{\mathrm{TD}\geq\mathrm{NPV}}$	-0.018	-0.034	-0.028	-0.043	-0.040
	(0.066)	(0.070)	(0.071)	(0.074)	(0.075)
Log of book assets		0.022	0.025	0.025	0.014
		(0.014)	(0.017)	(0.017)	(0.018)
Cash		0.049	0.042	0.047	0.054
		(0.058)	(0.070)	(0.072)	(0.076)
ROA		-0.015	-0.020	-0.021	-0.016
		(0.012)	(0.016)	(0.016)	(0.017)
Tobin's Q		0.000	-0.001	-0.001	-0.003
		(0.005)	(0.007)	(0.007)	(0.008)
Log of firm age		-0.052	-0.047	-0.059	-0.042
		(0.064)	(0.074)	(0.078)	(0.075)
Project NPV ($100Ms)		0.012*	0.012	0.012	0.012
		(0.006)	(0.008)	(0.009)	(0.009)
Project Capital Costs ($100Ms)		-0.013*	-0.011	-0.009	-0.007
		(0.007)	(0.007)	(0.007)	(0.008)
Expected mine life		-0.004	-0.005	-0.006	-0.009
		(0.004)	(0.005)	(0.009)	(0.012)
Total NPV of alternative projects ($100Ms)		-0.015	-0.015	-0.014	-0.011
		(0.011)	(0.011)	(0.011)	(0.011)
Start alternative NPV+ project		-0.023	-0.014	-0.003	0.012
		(0.109)	(0.114)	(0.111)	(0.110)
Primary mineral price (% ch.)			0.016		
			(0.061)		
Futures price				0.010	0.005
				(0.015)	(0.013)
Options-implied volatility					-0.081
					(0.064)
Controls for projects in each stage	No	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes
Primary mineral FE	No	No	No	Yes	Yes
Number of firms	177	174	143	140	126
Observations	838	822	679	662	589
R^2	0.290	0.300	0.295	0.297	0.300

Table 6: Firm liabilities and time until beginning construction on positive NPV projects. This table reports the results of Cox proportional hazards model regressions. The reported coefficients are the hazard ratios. The sample is comprised of project-year observations for mining projects from the year the firm publicly discloses the NI 43-101 technical report that includes the initial NPV estimate to the year the firm begins construction on the mine, or the sample period ends, whichever comes first. The sample consists of firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada over the sample period of 2003 to 2016. The "event" is when the firm begins construction on the project. *SB/MV* is the sum of estimated reclamation liabilities of all of a firm's mines defined as self-bonded divided by the market value of the firm's assets. *EB/MV* is the sum of estimated reclamation liabilities of all of a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) divided by the market value of the firm's assets. $\mathbb{1}_{\text{SB}\geq\text{NPV}}$ is an indicator variable that equals 1 if the sum of estimated reclamation liabilities of all of a firm's mines defined as self-bonded exceeds the initial NPV estimate of the mining project. $\mathbb{1}_{\text{EB}\geq\text{NPV}}$ is an indicator variable that equals 1 if the sum of estimated reclamation liabilities of all of a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) exceeds the initial NPV estimate of the mining project. $\mathbb{1}_{\text{TD}\geq\text{NPV}}$ is an indicator variable that equals 1 if the firm's total debt exceeds the initial NPV estimate of the mining project. The data on estimated reclamation liabilities were hand-collected from firms' public disclosures. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1. In both Panels A and B, the unreported time-varying coefficients used in Model (2) are log of book assets, cash, ROA, and Tobin's Q, log of firm age, the number of projects in each mining stage, the total NPV of the firm's alternative mining projects, and a dummy variable if the firm starts construction on an alternative project. The unreported time-varying coefficients used in Model (3) include those in Model (2), as well as the annual percentage change in the price of the primary mineral extracted. Models (4) and (5) add the 12-month futures price and the implied volatility from historical put-call straddles, respectively. These variables (defined in Appendix Table A1) are constructed using data from Compustat—North America, Mining Intelligence, and Bloomberg. Robust standard errors, clustered at the firm level, are reported in parentheses. *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Panel A: Liability ratios

	Survival analysis				
	(1)	(2)	(3)	(4)	(5)
SB/MV	0.344*	0.657	0.699	0.707	0.646
	(0.191)	(0.328)	(0.296)	(0.300)	(0.300)
EB/MV	0.552	0.725	0.746	0.759	0.143
	(0.449)	(0.402)	(0.416)	(0.367)	(0.306)
Market leverage	0.926	1.347	1.043	1.096	0.925
	(0.405)	(0.717)	(0.539)	(0.572)	(0.461)
Project capital costs ($100Ms)		0.978	0.973	0.985	1.006
		(0.033)	(0.035)	(0.052)	(0.049)
Project NPV($100Ms)		1.026	1.018	1.016	1.017
		(0.023)	(0.025)	(0.025)	(0.023)
Expected mine life		1.027	1.043	1.041	1.011
		(0.046)	(0.046)	(0.048)	(0.032)
Year FE	Yes	Yes	Yes	Yes	Yes
Primary mineral FE	Yes	Yes	Yes	Yes	Yes
Time-varying coefficients	No	Yes	Yes	Yes	Yes
Number of firms	191	189	158	155	144
Observations	955	944	823	811	754
Pseudo-R^2	0.108	0.127	0.114	0.115	0.143

Table 6—*Continued*

Panel B: Liability indicators					
	Survival analysis				
	(1)	(2)	(3)	(4)	(5)
$\mathbb{1}_{\text{SB}\geq\text{NPV}}$	0.329***	0.433**	0.467**	0.476**	0.470**
	(0.121)	(0.162)	(0.167)	(0.173)	(0.163)
$\mathbb{1}_{\text{EB}\geq\text{NPV}}$	0.709	0.800	0.864	0.889	0.643
	(0.380)	(0.447)	(0.479)	(0.496)	(0.373)
$\mathbb{1}_{\text{TD}\geq\text{NPV}}$	0.917	0.834	0.867	0.866	0.866
	(0.289)	(0.307)	(0.314)	(0.313)	(0.315)
Project capital costs ($100Ms)		0.951	0.949	0.956	0.945
		(0.043)	(0.042)	(0.049)	(0.051)
Project NPV($100Ms)		1.035	1.027	1.025	1.032
		(0.024)	(0.024)	(0.024)	(0.022)
Expected mine life		1.023	1.041	1.041	1.029
		(0.040)	(0.038)	(0.039)	(0.037)
Year FE	Yes	Yes	Yes	Yes	Yes
Primary mineral FE	Yes	Yes	Yes	Yes	Yes
Time-varying coefficients	No	Yes	Yes	Yes	Yes
Number of firms	191	189	158	155	144
Observations	955	944	823	811	754
Pseudo-R^2	0.108	0.126	0.113	0.114	0.125

Table 7: Risky firm liabilities and positive NPV projects. This table reports the results of linear probability models to analyze how more plausibly risky firm liabilities affect investment in positive NPV projects. The dependent variable and additional control variables for Models (1) and (2) match those of Table 4, Model (5). The dependent variable and additional control variables for Models (3) and (4) match those of Table 5, Model (2). Finally, the dependent variable and additional control variables for Models (5) and (6) match those of Table 6, Model (2). The additional control variables (defined in Appendix Table A1) are constructed using data from Compustat—North America, Mining Intelligence, and Bloomberg. Models (1), (3), and (5) exclude firms that have an investment grade bond rating at some point over the sample period. For Models (2), (4), and (6), the independent variables of interest are the measures of a firm's liabilities (self-bonded and externally-bond reclamation liabilities, and traditional debt), and the interaction of these measures with an indicator variable that equals one in the year a firm receives a credit downgrade from S&P or Moody's, and the year directly preceding the downgrade. Robust standard errors, clustered at the firm level, are reported in parentheses. *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable =	Acquire rights		Begin construction			
	(1)	(2)	(3)	(4)	(5)	(6)
SB/MV	-0.023***	-0.021***	-0.058***	-0.054***		
	(0.007)	(0.008)	(0.012)	(0.014)		
EB/MV	0.001**	0.002***	-0.004	0.004		
	(0.001)	(0.001)	(0.063)	(0.061)		
Market leverage	-0.026*	-0.018	-0.004	-0.006		
	(0.015)	(0.017)	(0.055)	(0.062)		
$\mathbb{1}_{\mathrm{SB}\geq\mathrm{NPV}}$					-0.414***	-0.284*
					(0.157)	(0.148)
$\mathbb{1}_{\mathrm{EB}\geq\mathrm{NPV}}$					-0.019	0.013
					(0.097)	(0.086)
$\mathbb{1}_{\mathrm{TD}\geq\mathrm{NPV}}$					0.003	-0.052
					(0.132)	(0.121)
SB/MV × downgrade period		-0.102**		-0.368**		
		(0.045)		(0.141)		
EB/MV × downgrade period		0.056***		0.052		
		(0.017)		(0.174)		
Market leverage × downgrade period		-0.006		-0.002		
		(0.005)		(0.037)		
$\mathbb{1}_{\mathrm{SB}\geq\mathrm{NPV}}$ × downgrade period						-0.170*
						(0.097)
$\mathbb{1}_{\mathrm{EB}\geq\mathrm{NPV}}$ × downgrade period						0.092
						(0.197)
$\mathbb{1}_{\mathrm{TD}\geq\mathrm{NPV}}$ × downgrade period						0.137
						(0.138)
Downgrade period		0.003		0.005		0.028
		(0.003)		(0.020)		(0.085)
Additional controls	Yes	Yes	Yes	Yes	Yes	Yes
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Number of firms	756	775	170	174	170	174
Observations	6,361	6,747	791	822	791	822
R^2	0.133	0.135	0.312	0.300	0.312	0.302

Table 8: Firm liabilities and capital expenditures. This table reports the results of linear regression models in which the dependent variable is capital expenditures (as a percentage of a firm's total book assets). In Models (1) through (3) the sample consists of firms listed on the Toronto Stock Exchange (TSX) or Toronto Stock Exchange Venture (TSXV) and located in Canada over the sample period of 1990 to 2016. The sample period in Model (4) is 2003 to 2016 to match the sample used in the project-level analysis. SB/MV is the sum of estimated reclamation liabilities of all a firm's mines defined as self-bonded divided by the market value of the firm's assets. EB/MV is the sum of estimated reclamation liabilities of all a firm's mines defined as externally-bonded (guaranteed with a surety bond, letter of credit, etc.) divided by the market value of the firm's assets. The data on estimated reclamation liabilities were hand-collected from firms' public disclosures. Self-bonded and externally-bonded mines are classified using the self-bonding regulations in Table 1 according to the description in Section 3.1. Accounting variables (defined in Appendix Table A1) are constructed using data from Compustat-North America and the data on firms' mining projects were provide by Intelligence Mining. Robust standard errors, clustered at the firm level, are reported in parentheses. *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable =	Capital expenditures			
	(1)	(2)	(3)	(4)
SB/MV	-0.023***	-0.022***	-0.022***	-0.021***
	(0.003)	(0.003)	(0.003)	(0.003)
EB/MV	-0.004***	-0.003***	-0.003***	-0.003***
	(0.001)	(0.001)	(0.001)	(0.001)
Market leverage	-0.022	-0.038**	-0.028	-0.033*
	(0.017)	(0.017)	(0.017)	(0.018)
Market leverage × Tobin's Q			-0.003**	
			(0.001)	
Log of book assets		0.004	0.004	0.010***
		(0.004)	(0.004)	(0.004)
Cash		-0.085***	-0.086***	-0.101***
		(0.013)	(0.013)	(0.014)
ROA		-0.012***	-0.012***	-0.013***
		(0.002)	(0.002)	(0.002)
Tobin's Q		-0.002**	-0.001*	-0.002***
		(0.001)	(0.001)	(0.001)
Log of firm age		-0.034***	-0.034***	-0.039***
		(0.010)	(0.010)	(0.010)
# of mines in exploration		0.001	0.001	0.000
		(0.001)	(0.001)	(0.001)
# of mines in feasibility		0.002	0.002	0.002
		(0.006)	(0.006)	(0.006)
# of mines in construction/permitting		0.025**	0.025**	0.024**
		(0.011)	(0.011)	(0.011)
# of producing mines		-0.002	-0.002	-0.002
		(0.003)	(0.003)	(0.004)
# of closed mines		0.000	0.000	0.003
		(0.006)	(0.006)	(0.009)
Firm FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Number of firms	790	775	775	764
Observations	7,029	6,697	6,697	5,904
R^2	0.354	0.387	0.388	0.415

Table 9: Reclamation liabilities and investment in the U.S. This table reports the results of linear regression models in which the dependent variable of interest is either capital expenditures (as a percentage of a firm's total book assets) or the incidence of a new mine (Model (7)). In Models (1) through (4), the sample consists of mining firms incorporated in the United States over the sample period of 1992 to 2016 that self-disclosed reclamation liabilities in their annual reports. *SB/MV* is total amount of disclosed self-bonded reclamation liabilities divided by the market value of the firm's assets. *EB/MV* is total amount of disclosed externally-bonded (guaranteed with a surety bond, letter of credit, etc.) reclamation liabilities divided by the market value of the firm's assets. $\mathbb{1}_{\text{SB}\geq 0}$ is an indicator variable that equals 1 if the firm has a positive amount of self-bonds. In Models (5) through (7), I use data from the Mine Safety and Health Administration (MSHA) on the location and permitting of a firm's mines to calculate the number of mines defined as self-bonded and externally-bonded using the self-bonding regulations in Table 1 according to the description in Section 3.1. Accounting variables (defined in Appendix Table A1) are constructed using data from Compustat-North America. Robust standard errors, clustered at the firm level, are reported in parentheses. *,**, and *** denote significance at the 10%, 5%, and 1% level, respectively.

Dependent variable =	Capital expenditures						New mine
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
SB/MV	-0.080*	-0.112**					
	(0.040)	(0.043)					
EB/MV	-0.016	-0.011					
	(0.014)	(0.008)					
$\mathbb{1}_{\text{SB}\geq 0}$			-0.016***	-0.022***			
			(0.004)	(0.008)			
# of self-bonded mines					-0.003*	-0.003*	-0.015**
					(0.002)	(0.001)	(0.006)
# of externally-bonded mines					0.003*	0.003**	0.016***
					(0.001)	(0.001)	(0.006)
Market leverage	-0.037	-0.002	-0.040	-0.015	-0.026**	-0.009	
	(0.028)	(0.032)	(0.026)	(0.028)	(0.012)	(0.012)	
Log of book assets		0.011		0.010		-0.004	
		(0.016)		(0.015)		(0.003)	
Cash		-0.080**		-0.035		-0.092***	
		(0.035)		(0.037)		(0.017)	
ROA		-0.043*		-0.040*		0.043	
		(0.024)		(0.023)		(0.027)	
Tobin's Q		0.028***		0.024***		0.015***	
		(0.006)		(0.005)		(0.005)	
Log of firm age		-0.001		-0.000		0.001	
		(0.001)		(0.001)		(0.005)	
Firm fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Data used	10-k	10-k	10-k	10-k	MSHA	MSHA	MSHA
Number of firms	39	39	42	42	120	120	4,983
Observations	338	338	359	359	1,453	1,453	33,876
R^2	0.621	0.680	0.629	0.682	0.559	0.585	0.238

Table A1: Data Appendix. This table defines the variables used in the empirical tests in the main portion of the paper and lists the data source(s) for each variable.

Variable name	Source	Definition
SB	Technical reports, Mining Intelligence	Sum of estimated reclamation liabilities of producing mines defined as self-bonded
EB	Technical reports, Mining Intelligence	Sum of estimated reclamation liabilities of producing mines defined as externally-bonded
SB/MV	Technical reports, Mining Intelligence, Compustat	SB/((prcc_f*cshpri) + dlc + dltt + pstkl − txditc)
EB/MV	Technical reports, Mining Intelligence, Compustat	EB/((prcc_f*cshpri) + dlc + dltt + pstkl − txditc)
Market leverage	Compustat	(dltt+dlc)/((prcc_f*cshpri) + dlc + dltt + pstkl − txditc)
$\mathbb{1}_{\text{SB}\geq\text{NPV}}$	Technical reports, Mining Intelligence	Indicator for SB that exceed the estimated project net present value
$\mathbb{1}_{\text{EB}\geq\text{NPV}}$	Technical reports, Mining Intelligence	Indicator for EB that exceed the estimated project net present value
$\mathbb{1}_{\text{TD}\geq\text{NPV}}$	Technical reports, Mining Intelligence	Indicator for total debt liabilities that exceed the estimated project net present value
Log of book assets	Compustat	log(at)
Capital expenditures	Compustat	capx/at
Return on Assets (ROA)	Compustat	ebitda/at
Cash	Compustat	che/at
Tobin's Q	Compustat	((prcc_f*cshpri) + dlc + dltt + pstkl − txditc)/at
Log of firm age	Datastream	Log of years since firm listed on TSX or TSXV
Project NPV	Mining Intelligence	Initial estimate of project net present value
NPV of mining rights	Mining Intelligence	NPV estimate from NI 43-101 techinical report (or the remaining NPV estimated by this reported if the mine is already producing), less the cost the acquiring firm paid for the individual mine at acquisition
Project capital costs	Mining Intelligence	Initial estimate of project capital costs
Total NPV of alternative projects	Mining Intelligence	The total NPV of the firm's alternative mining projects in a given year
Start alternative NPV+ project	Mining Intelligence	Indicator for firm starting an alternative NPV + project in a given year
Primary mineral price (% ch.)	FRED global price series or producer price indices	$(\text{P}_t - \text{P}_{t-1})/\text{P}_{t-1}$, where P is the January price in year t
Futures price	Bloomberg	12-month futures prices for the primary mineral mined
Options-implied volatility	Bloomberg	Volatility implied by historical 1-month, at the money put-call straddles
Mine status	Mining Intelligence	Status (exploration, feasibility, production, etc) of mining project
Mine type	Mining Intelligence	Type (Surface, open pit, or underground) of mining project
Primary commodity	Mining Intelligence	Primary mineral to be extracted
Downgrade period	Compustat, Moody's reports	Indicator for the year of, and year prior to, a security rating downgrade from S&P or Moody's
From Table 8		
SB/MV	10-k's, Compustat	Self-disclosed self-bonded reclamation liabilities/((prcc_f*cshpri) + dlc + dltt + pstkl − txditc)
EB/MV	10-k's, Compustat	Self-disclosed externally-bonded (surety bond, letter of credit, etc) reclamation liabilities/((prcc_f*cshpri) + dlc + dltt + pstkl − txditc)
$\mathbb{1}_{\text{SB}>0}$	10-k's	Indicator for positive levels of self-bonded reclamation liabilities
# of self-bonded mines	Mine Health & Safety Administration	Number of mines permitted in a jurisdiction and at a time when self-bonding was allowed
# of externally-bonded mines	Mine Health & Safety Administration	Number of mines permitted in a jurisdiction and at a time when self-bonding was not allowed

Do Nonfinancial Firms Use Financial Assets to Risk-Shift?

Evidence from the 2014 Oil Price Crisis

Zhiyao Chen

Chinese University of Hong Kong, Hong Kong

Ran Duchin

University of Washington, Seattle WA 98195, USA

May 2019

Zhiyao Chen is at the CUHK Business School, Chinese University of Hong Kong, Shatin, Hong Kong. Email: nicholaschen@baf.cuhk.edu.hk. Ran Duchin is at the Foster School of Business, University of Washington, PO Box 353226, Seattle, WA 98195-3226. Email: duchin@uw.edu. We thank Philip Bond, Sudipto Dasgupta, Vyacheslav (Slava) Fos, Jarrad Harford, Ron Kaniel, Philip Strahan, and seminar participants at Case Western, CUHK Business school, Oklahoma Price business school, and the University of Toronto for valuable comments.

Do Nonfinancial Firms Use Financial Assets to Risk-Shift?

Evidence from the 2014 Oil Price Crisis

Abstract

Using hand-collected data on financial portfolios of firms in the oil industry, we investigate risk-shifting around the 2014 oil crisis. Following the crisis, firms with high leverage, particularly short-term, uncollateralized, and unhedged, substantially increase their investments in risky financial assets, including corporate debt, equity, and mortgage-backed securities. In contrast, they do not invest in riskier real assets, which are more visible, restricted by debt covenants, and carry higher transaction costs and delayed payoffs. Overall, we provide first evidence that distressed firms risk-shift using financial assets camouflaged as cash reserves, highlighting the role of debt maturity, collateral, and hedging in risk-shifting.

1. Introduction

A vast body of theoretical work predicts that firms will invest in riskier projects as they become distressed (Modigliani and Miller (1958), Fama and Miller (1972), Jensen and Meckling (1976), Stiglitz and Weiss (1981), Acharya and Viswanathan (2011), and Della Seta, Morellec, and Zucchi (2019)). Despite the prominence of these theories, the empirical evidence on risk-shifting is mixed. On the one hand, Andrade and Kaplan (1998), Rauh (2009), and Gilje (2016), among others, find that firms do not undertake riskier investments as they become distressed. On the other hand, studies such as Eisdorfer (2008), Becker and Stromberg (2012), and Pryshchepa, Aretz, and Banerjee (2013) identify settings in which distressed firms may increase their risk-taking. Recently, Denes (2018) shows that government leverage in VC funding results in risk-shifting by distressed portfolio firms.

We attempt to shed new light on this topic by extending the literature on risk-shifting in several important ways. First, contrary to prior empirical studies that focus exclusively on real investments, this paper investigates firms' financial investments as a novel conduit for risk-shifting. We argue that financial assets can be better conduits for risk-shifting than real assets. Compared with traditional real assets, financial assets are more liquid, easier to access, and carry substantially lower transaction costs. Furthermore, trading in risky financial assets is less visible, does not require an upfront investment in physical or human capital, and can generate immediate/accelerated payoffs.[1] Finally, financing contracts such as debt covenants that are used to counter managerial agency typically restrict real investment and payout (Chava et al. (2010)); however, they do not restrict nor consider financial investments. In fact, financial assets are typically reported on the balance sheet as corporate cash holdings, and consequently, allow firms to camouflage their risk-shifting as investments in seemingly safe asset classes.

[1] The merit of liquidity and accelerated payoffs is well demonstrated by the actions of Fred Smith, the founder of FedEx, who saved his company in 1971 by gambling its last $5,000 in Las Vegas. The immediate payoffs from his speculative positions, which came at the expense of the company's creditors, allowed FedEx to operate for another week and consequently to this very day.

Second, we highlight the importance of debt maturity in risk-shifting. While prior studies argue that short-term debt can mitigate agency problems by disciplining management (e.g., Calomiris and Kahn (1991) and Leland and Toft (1996)), the roll-over risk of short-term debt also introduces incentives to take risk to improve interim debt repricing and avoid inefficient liquidation (Della Seta, Morellec, and Zucchi (2018)). Short-term debt likely plays a particularly important role in the context of a transitory oil price crisis since it is more likely to become due before the crisis ends.

Third, we investigate the role of collateral in risk-shifting. We argue that collateralized financial leverage is less likely to trigger risk-shifting behavior because collateral prevents wealth transfers from debtholders to equity holders. Specifically, collateral guarantees the value of the debt claims, mitigates information asymmetry between equity holders and debtholders (Besanko and Thakor (1987) and Boot, Thakor and Udell (1991)), and reduces costly monitoring by lenders (Rajan and Winton, 1995). Thus, risk-shifting should occur primarily in highly levered firms with low levels of collateral.

Finally, we provide novel evidence on the role of hedging in weakening risk-shifting incentives. Theory suggests that corporate hedging can effectively mitigate risk and enhance firm value (e.g., Froot et al. (1993), Leland (1998), Chidambaran et al. (2001)). Prior empirical studies, however, do not offer conclusive evidence that hedging matters for firm policies and value (e.g., Guay and Kothari (2003), Jin and Jorion, (2006), Bartram et al. (2011)). In contrast, this paper shows that risk-shifting is an important channel through which hedging can affect corporate policies and outcomes.

The identification strategy exploits the 2014 oil price crisis as a natural quasi experiment. The 2014 oil price drop was an exogenous, market-wide shock that affected the entire oil and gas industry.[2] As Figure 1 shows, the downward oil price pressure persisted through 2016. In fact, Figure 1 shows that oil prices decreased dramatically by the end of 2016. The rationale for using the crisis is that the substantial drop in oil price reduced operating income across all oil and gas firms. The cash flow squeeze, coupled with high short-term leverage, triggered financial distress at affected firms and created an incentive to risk-shift.

To study the role of financial assets in risk-shifting, we collect information on financial investments following the method of Duchin, Gilbert, Harford, and Hrdlicka (DGHH, 2017). DGHH find that U.S. industrial firms invest heavily in non-cash, risky financial assets (e.g., corporate debt, equity, mortgage-backed securities). We hand-collect data on financial investments for firms in the oil and gas industry from 2012 to 2016. The data are collected by exploiting the 2009 Statement of Financial Accounting Standards (SFAS) No. 157 that requires all firms to report the fair value of major asset classes on their balance sheets. The empirical analyses examine whether levered firms increased their risky financial investments in response to the 2014 oil price crisis.

The empirical setting has two important advantages. First, the relative risk of financial assets can be assessed more precisely ex-ante than the risk of real assets. For instance, it is clear that corporate debt, equity, and mortgage-backed securities are riskier than treasury bonds and notes. In contrast, it is difficult to determine which well exploration is riskier ex-ante. Consequently, most previous studies relied on measures of realized ex-post cash flow volatility (e.g., Eisdorfer (2008),

[2] By the end of 2008, the price of oil had bottomed out at $40. The economic recovery that began the following year sent the price of oil back over $100; it hovered between $100 and $125 until 2014, when it experienced another steep drop to $37. Numerous factors contributed to the 2014 drop in oil prices. On the one hand, overproduction caused an excess supply of oil. Countries such as the U.S. and Canada increased their efforts to produce oil. In the U.S., private companies began extracting oil from shale formations in North Dakota using a process known as fracking. Meanwhile, Canada went to work extracting from Alberta's oil sands, the world's third-largest crude oil reserve. As a result of this local production, the two North American countries were able to cut their oil imports sharply, which put further downward pressure on worldwide oil prices. Moreover, faced with a decision between letting prices continue to drop and ceding market share by cutting production in an effort to send prices upward again, Saudi Arabia kept its production stable, deciding that low oil prices offered more of a long-term benefit than giving up market share. By supporting low oil prices, Saudi Arabia hoped that countries such as the U.S. and Canada would be forced to abandon their costlier production methods, such as fracking, due to a lack of profitability. On the other hand, the demand has declined since 2010. Economies such as China, whose rapid growth and expansion greatly increased demand for oil in the 2000’s, began to slow after 2010. Other large emerging economies such as Russia, India and Brazil experienced a similar decline after 2010.

Becker and Stromberg (2012), Pryshchepa, Aretz, and Banerjee (2013), Chen, Strebulaev, Zhang and Xing (2018), and Chen and Strebulaev (2018)). These measures, however, suffer from a look-ahead bias, are not necessarily attributable to risk-taking, and can be driven by omitted variables.

Second, the empirical design mitigates concerns about reverse causality. While losses from risky financial investments have no direct effect on operating performance, losses from risky real investments do affect operating performance. Consequently, studies of real investment struggle to identify the direction of causality, that is, whether deteriorating operating performance results in risky investments or whether risky investments result in deteriorating operating performance. Since financial investments are operating-performance-neutral, this setting is less subject to reverse causality critique.

In the main analyses, we find that risky financial assets increased by $244.65 million, or 51%, on average, among firms with high levels of short-term debt. Consequently, the ratio of risky financial assets to total financial assets increased from 21% in 2013 to 33% in 2016 among firms with high short-term debt, and the ratio of risky financial assets to total book assets increased from 2.2% in 2013 to 4.8% in 2016. The significant economic increase in risky financial assets is also highly statistically significant at conventional levels and continues to hold in difference-in-difference panel regressions that include year and firm fixed effects as well as time-varying firm-level attributes such as expenditures and sales. The results also hold robustly across alternative measures of short-term debt and in subsamples of oil producers.

Overall, these estimates imply that highly levered oil firms gravitated considerably toward riskier financial investments following the oil price crisis. By the end of 2016, their risky financial investments accounted for a nontrivial fraction of both their financial and total operations.

To further address the endogeneity problem associated with risk-taking, we estimate a first-stage regression that identifies exogenous changes in profitability from 2013 to 2016 based on the 2014 oil price shock and use the fitted changes in profitability from the first stage regressions to explain changes in risk-taking from 2013 to 2016 in the second stage regressions. The estimates suggest that the decline in profitability following the oil price crisis was accompanied by an increase in financial risk-taking at highly levered firms. These results hold across different measures of risky financial assets and are highly statistically significant. They suggest that declines in firm profitability around the oil price crisis played an important role in risk-taking at firms with high short-term debt.

We further confirm the role of distress in financial risk-taking by hand-collecting data on bankruptcy filings in the oil and gas industry following the 2014 oil price crisis. The estimates suggest a significant increase of 25.1% in the ratio of risky financial assets to total financial assets in the year before filing for bankruptcy.

In the second set of analyses, we investigate the role of collateral in corporate risk-shifting. First, we double-sort firms on short-term liabilities and collateral. The estimates suggest that the increase in financial risk-taking at highly levered firms is concentrated in firms with low collateral. At highly levered firms with low levels of collateral, the ratio of risky financial assets to total financial assets increases from 27% in 2013 to 46% in 2016 and the difference is statistically significant at the 1% level. In contrast, at highly levered firms with high levels of collateral, the ratio of risky financial assets to total financial assets only increases from 16% in 2013 to 21% in 2016, and the difference is statistically insignificant at conventional levels.

We obtain similar results in triple difference-in-differences regressions where the key explanatory variable is the triple interaction *Low Collateral x High short-term liabilities x Crisis.* In this setting, which includes year and firm fixed effects, the above term is economically large and has a positive sign, suggesting that risky financial investments increase considerably more at levered firms with low collateral.

We also investigate the role of oil price hedging in risk-shifting by hand-collecting data on firms' derivative hedging positions at the onset of the oil price crisis. In univariate and regression analyses, we find that the increase in financial risk-taking at highly levered firms is concentrated in unhedged firms. At highly levered unhedged firms, the ratio of risky financial assets to total financial assets increases from 26% in 2013 to 45% in 2016. We obtain similar results in triple difference-in-differences regressions. These findings provide new evidence on the effect of hedging on corporate policies through its role in risk-shifting incentives.

In the final set of analyses, we compare the changes in financial investments and real investments among oil and gas firms after the onset of the oil price crisis. Following Gilje (2016), we hand-collect data on real investment risk from the 10-K disclosures of all publicly traded U.S. domiciled oil and gas firms from 2012 to 2016. We define risky real investment as investment in exploratory wells. Consistent with prior evidence (Gilje (2016)), we do not find that highly levered firms increased their risky real investments following the crisis. The differences between risky real investments in 2012 and in 2016 (measured in dollars or as a percentage of total assets) are mostly negative and statistically insignificant. We find similar results in difference-in-differences regressions that include firm and year fixed effects. These findings indicate that following the oil price crisis, highly levered firms did not risk-shift using their real assets. In fact, the findings are mostly consistent with a reduction in real investment risk, consistent with the evidence in Gilje (2016).

We also investigate the relation between financial risk taking and real investment risk. The results suggest that at firms with high short-term liabilities at the onset of the crisis, higher financial risk-taking is correlated with lower real investment risk. One interpretation of these findings is that distressed firms substitute risky financial investments for risky real investments when facing short-term debt maturities, possibly due to the accelerated payoffs of financial investments relative to real investments. To the extent that real investment is costlier to adjust, this evidence indicates that firms that face lower real investment risk increase their overall risk by investing in risky financial assets.

Overall, we contribute to the existing literature by providing new evidence on corporate risk-shifting along several dimensions. First, we investigate firms' financial investments as an alternative conduit for risk-shifting. As shown by Ang et al. (2006) and Frazzini and Petersen (2014), investments in financial assets with high systematic risk or high idiosyncratic risk lead to lower (and negative) stock returns or lower alphas, respectively. Second, we highlight the importance of collateral and hedging in the incentives to risk-shift. Third, our identification strategy exploits the 2014 oil price crisis as an unexpected exogenous shock to firms' operating performance, thus mitigating concerns about the endogeneity of risk-shifting.

2. Empirical Strategy, Data and Variables

A. Sample

To study risk-shifting around the oil price crisis, the sample period spans a six years window around the onset of the oil price crisis in 2014. The sample period starts in 2011, three years before the onset of the crisis, and ends in 2016, three years after the onset of the crisis. The start of the sample period in 2011 occurs two years after SFAS No. 157 went into effect, requiring all firms to report the fair value of their financial asset classes in their annual 10-K reports.

The empirical analyses focus on the oil and gas industry because it is the only industry that experienced a substantial negative shock since 2009, when SFAS No. 157 went into effect. An additional benefit of studying oil and gas firms is that it also allows us to measure the risk of real investments based on oil wells' risk, as proposed and implemented by Gilje (2016).

To construct the sample, we merge the hand-collected data on financial asset portfolios with both quarterly and annual Compustat data. The quarterly data are used to construct the time-invariant measures of short-term liabilities and collateral as of the onset of the crisis, that is, as of the end of the second quarter of 2014. The annual data are used to construct the time-varying control variables, whose annual frequency matches that of firms' financial asset reporting.

Specifically, we begin constructing the sample with all firms in the oil and gas industry (SIC codes between 1300 and 2999) with nonmissing observations on financial assets and positive total assets from 2011-2016.[3] This procedure yields a sample of 122 distinct oil and gas firms and 732 firm-year observations, which are used in univariate analyses (Tables 1 and 2) to describe the financial asset portfolios around the oil price crisis. For the subsequent multivariate regression analyses, we also require that the other control variables, such as *Market-to-book*, *Profitability*, and *Capital investment*, have nonmissing values. Consequently, we lose 5 firm-year observations, and end up with the final sample that includes 122 distinct oil firms and a total of 727 firm-year observations. We note that detailed variable definitions are given in Appendix A.

B. Empirical Strategy

To analyze the role of financial assets in risk-shifting following the onset of the oil price crisis, we employ a difference-in-differences approach in which we compare the risk of the financial asset portfolio of firms before and after the onset of the crisis as a function of their financial position (short-term debt and collateral), controlling for firm and year fixed effects, as well as observable time-varying firm attributes such as expenditures and sales.

We are mostly interested in studying the role of firms' financial positions (debt and collateral) in risk-shifting during the oil price crisis. Inferences may be confounded, however, if variation in these financial positions as the crisis unfolds is endogenous. To address this concern, we purge our specifications of this variation by using only the firm's financial positions measured at the end of the quarter before the onset of the crisis. Thus, our main framework is akin to an instrumental variables approach. The identifying assumption is that firms did not predict the oil price shock, and therefore, their financial positions before the onset of the crisis are independent of

[3] SIC codes 1300-2999 include the following sub-sectors: oil and gas extraction (1300-1300), crude petroleum and natural gas (1310-1319), natural gas liquids (1320-1329), petroleum and natural gas (1330-1339), petroleum and natural gas (1370-1379), oil and gas field services (1380-1380), drilling oil and gas wells (1381-1381), oil-gas field exploration (1382-1382), oil and gas field services (1389-1389), petroleum refining (2900-2912), and miscellaneous petroleum products (2990-2999).

the ensuing crisis. This empirical design is similar to the empirical design in recent papers such as Almeida, Campello, Laranjeira, and Weisbenner (2012).

More specifically, we calculate firms' short-term liabilities and collateral values as of the end of the second quarter of 2014, just before the onset of the crisis. Short-term liabilities are defined as the ratio of total current liabilities (Compustat item LCTQ) to total assets (Compustat item ATQ). Following Vassalou and Xing (2004) and Campbell, Hilscher, and Szilagyi (2008), we use current liabilities rather than current debt because a firm's financial burden comprises its total financial liabilities and not just debt. To measure a firm's collateral value, we follow the capital structure literature (e.g., Rampini and Viswanathan, 2013) and define collateral as the ratio of physical assets (Compustat item PPNETQ) to total book assets (Compustat item ATQ).

C. Measures of Risky Financial Assets

To study the risk of firms' financial investments, we follow DGHH (2017) and collect data on firms' financial portfolios that comprise: (1) the balance sheet accounts "cash and cash equivalents" and "short-term investments," which constitute Compustat's data item CHE, the standard measure of cash holdings in the literature, and (2) any additional financial assets reported as "long-term investments" or "other assets".

We hand-collect detailed information on the asset classes that constitute firms' financial portfolios from the footnotes of annual reports for all oil and gas firms.[4] To classify the riskiness of firms' financial assets, we follow DGHH (2017) and define safe financial assets as those that fall into the following asset classes: cash, cash equivalents, demand deposits, money market securities, treasury bills, treasury notes and treasury bonds. All other financial assets are considered risky. Restricted assets, deferred executive compensation, and derivative hedging are excluded from the analyses.

[4] Firms do not disclose the asset class information in their quarterly report.

Using these data, we construct three measures of firms' risky financial assets. The first measure is the logarithm of the dollar amount (in $ millions) of risky financial assets:

$$Log\ risky\ financial\ assets\ = log(1 +\ risky\ financial\ assets)$$

This measure is unscaled by a firm's investments or assets, and therefore is unaffected by changes in other firm-level attributes unrelated to its risky financial investments.

The second measure is the ratio of risky financial assets to total financial assets, defined as follows:

$$Risky\ financial\ assets/finanical\ assets = \frac{Risky\ financial\ assets}{Safe\ financial\ assets +\ Risky\ financial\ assets}.$$

This measure captures the percentage of financial investments allocated to risky financial assets. It measures the composition of a firm's total financial asset portfolio, reflecting the relative allocation to both safe and risky assets. As such, an increase in financial risk-taking according to this measure can reflect an increase in risky assets or a decrease in safe assets, holding constant the size of the financial asset portfolio.

The third measure is the ratio of risky financial assets to total book assets:

$$Risky\ financial\ assets/book\ assets = \frac{Risky\ financial\ assets}{Total\ book\ assets}.$$

This measure scales a firm's risky financial assets by its size. It therefore captures the importance of a firm's risky financial assets in its overall operations or value. We caution the reader, however, that one possible concern with this measure is that the value of a firm's total book assets likely declines following the onset of the oil price crisis. Consequently, an increase in the ratio of a firm's risky financial assets to total book assets following the crisis might be driven by a decline in its book assets rather than by an active decision to increase its risky financial asset holdings. Nevertheless, we include this measure in our analyses to provide evidence on the economic importance of risky financial asset holdings.

D. Measures of Risky Real Assets

We construct a measure of firms' risky real assets by hand-collecting data on investment risk from the 10-K reports of the Crude Oil & Natural Gas firms (SIC 1311) in the sample. These data are collected from firms' disclosures on "Costs Incurred in Natural Gas and Oil Exploration and Development, Acquisitions and Divestitures," which provide information related to expenditures on exploratory wells and development wells. Exploratory wells are those drilled to find a new field or to find a new reservoir in a field in which another reservoir has previously produced oil or gas. Development wells are those drilled within the proven area of an oil or gas reservoir to the depth of a stratigraphic horizon known to be productive.

Following Gilje (2016), we categorize all activities associated with exploratory drilling as risky investments. This includes both the capital to drill and the capital to acquire unproven acreage in which to drill. All activities associated with development drilling, which include the drilling of development wells and the acquisition of proven/producing acreage for development drilling are classified as safe investments.

Using these data, we construct three measures of risky real assets investments, which are similar to the measures of risky financial assets discussed above. The first measure is the logarithm of the amount (in $ millions) of risky real assets:

$$Log\ risky\ real\ assets\ = log(1 +\ exploratory\ oil\ wells)$$

The second measure is the ratio of investment in exploratory wells to total investment in wells:

$$Risky\ real\ assets/real\ assets = \frac{Exploratory\ oil\ wells}{Exploratory + development\ wells}$$

The third measure is investment in exploratory wells scaled by the total book assets:

$$Risky\ real\ assets/book\ assets = \frac{Exploratory\ wells}{Total\ book\ assets}$$

F. Summary Statistics

Table 1 presents summary statistics for the sample. Based on Panel A, the dollar amount of risky financial assets ranges from 0 to $30,985 million, with a mean value of $683 million and a median value of $16 million. The ratio of risky financial assets to total financial assets is highly right-skewed, with a mean value of 0.32, and a median value of 0.21. Similarly, the ratio of risky financial assets to total book assets is also right-skewed with a mean value of 0.04 and a median value of 0.01. These values are smaller compared to the risky financial assets of S&P 500 firms reported by DGHH (Mean = 0.40 for the ratio of risky financial assets to total financial assets; Mean = 0.06 for the ratio of risky financial assets to total book assets). This is consistent with the findings of DGHH that larger firms invest more in risky financial assets.

Panel A of Table 1 also provides summary statistics for firms' real investments. The dollar amount of risky real assets (exploratory wells) ranges from 0 to $112,591 million, with a mean value of $84 million and a median value of $731 million. The ratio of exploratory wells to total wells is right-skewed, with a mean value of 0.33 and a median value of 0.24. Similarly, the ratio of risky real assets to total book assets is also right- skewed, with a mean value of 0.09 and a median value of 0.04.

Panel B describes the other variables employed in this study. The ratio of short-term liabilities to total assets has mean and median values of 0.13 and 0.11, respectively. Collateral values of oil and gas firms are relatively high (Mean = 0.69), likely because these are manufacturing firms that rely heavily on fixed assets. The logarithm of total sales (firm size) has mean and median values of 6.53 and 6.59, respectively. Further, the oil and gas firms in our sample are on average profitable (Mean profitability = 0.09) and have Market-to-book ratios close to 1. About half of the oil and gas firms in our sample do not pay dividends, consistent with the findings in prior studies (Brav, Graham, Harvey, and Michaely, 2005). Lastly, total liabilities are, on average, 46% of book assets.

Panel C of Table 1 investigates the covariate balance between low- and high-short-term-liabilities firms at the onset of the crisis, that is, at the end of the second quarter of 2014. Since we sort firms on their short-term liabilities at the onset of the crisis, we first verify that there are significant differences between low- and high-short-term-liabilities firms. As Panel C shows, the difference in short-term liabilities economically large (High minus Low = 0.132) and statistically significant at the 1% level (t-statistic = 11.25). In contrast, the differences across other firm observable firm attributes are small and statistically indistinguishable from zero. These findings suggest that the two sets of firms are observationally identical at the onset of the crisis on all dimensions but the treatment variable – short-term liabilities.

4. Empirical Results

This section studies how the variation in financial positions at the onset of the oil price crisis affects firms' risk-taking behavior. We begin with an analysis of short-term liabilities and financial asset portfolios, proceed with an investigation of the effects of collateral, and conclude with an examination of real investments.

A. Univariate Evidence

We begin our analysis by presenting univariate results on the relation between firms' outstanding short-term liabilities at the onset of the oil price crisis and the riskiness of their financial asset portfolios. Table 2 reports annual average risky financial asset holdings in 2011-2016 across firms with low and high short-term liabilities. Specifically, we sort firms into terciles based on their short-term liabilities at the end of the second quarter of 2014, and label firms in the top tercile as high short-term liabilities firms and firms in the bottom two terciles as low short-term liabilities firms. Panel A corresponds to firms with low short-term liabilities and Panel B corresponds to firms with high short-term liabilities.

The results in Panel A of Table 2 show that firms with low short-term liabilities outstanding at the onset of the oil price crisis did not change their holdings of risky financial assets significantly. Across all three measures of risky financial assets, the differences-in-means from 2013 to 2016 are economically small and change signs. For example, the ratio of risky financial assets to total financial assets has decreased by 4% from 2013 to 2016, whereas the ratio of risky financial assets to total book assets has increased by 1.3%. These differences also are statistically insignificant at conventional levels.

In contrast, the results in Panel B show that firms with high short-term liabilities have increased their risky financial asset holdings substantially. The average investment in risky financial assets has increased from $484 million in 2013 to $728 million in 2016. Similarly, the ratio of risky financial assets to total financial assets has increased by 12.1% from 2013 to 2016, and the ratio of risky financial assets to total book assets has increased by 2.6%. These increases are all highly statistically significant at the 1% level.

Taken together, the univariate evidence suggests that firms that entered the oil price crisis with high levels of short-term debt increased their holdings of risky financial assets substantially. This increase persisted throughout 2016, when oil prices reached their lowest point. These findings provide first evidence that financially distressed firms increased their risk by investing in risky financial assets, consistent with theories of risk-shifting.

B. Multivariate Regression Evidence

Table 3 shows multivariate evidence on the effect of short-term liabilities at the onset of the oil price crisis on the riskiness of financial asset portfolios with a full system of controls and fixed effects. The table provides estimates from difference-in-differences regressions explaining the riskiness of firms' financial asset portfolios. In columns 1-3, the dependent variable is the logarithm of the dollar value of firms' risky financial assets. In columns 4-6, the dependent variable is the ratio of risky of financial asset to total financial assets. Finally, in columns 7-9, the dependent variable is the ratio of risky financial assets to total book assets. For each dependent variable, we report three regression models with different fixed effects. The first model does not include firm or year fixed effects. The second model includes firm fixed effects. The third model includes both firm and year fixed effects.

The main variable of interest it the interaction term *High short-term liabilities x Crisis*, which captures the effect high outstanding short-term liabilities at the onset of the crisis on the risk of firms' financial asset portfolios following the onset of the crisis. The variable *High short-term liabilities* is an indicator variable that equals 1 for firms in the top tercile on short-term liabilities at the onset of the crisis (Average ratio of short-term liabilities to book assets = 0.221) and 0 for firms in the bottom two terciles (Average ratio of short-term liabilities to book assets = 0.089).

Columns 1-3 show that firms with high outstanding short-term liabilities have increased substantially the dollar amount invested in risky financial assets. The coefficient on *High short-term liabilities x Crisis* is positive and highly statistically significant at the 1% level across all three specifications. The economic magnitudes are highly stable across the different regression models and imply that following the onset of the crisis highly levered firms increased their investments in risky financial assets by 49.6% to 53.0% more compared to unlevered firms. These findings hold even after controlling for unobservable time-invariant differences across firms (firm fixed effects) and macroeconomic time trends (year fixed effects).

In columns 4-6, we consider risky financial assets relative to total financial investments. The findings suggest that highly levered firms have increased the fraction of their financial portfolio invested in risky financial assets by 9.0%-9.5% compared to unlevered firms following the onset of the oil price crisis. These findings are highly statically significant at the 5% level and continue to hold after controlling for firm and year fixed effects. Finally, in columns 7-9, we consider the investment in risky financial assets relative to the total assets of the firm. We find that highly levered firms have increased their risky financial investments relative to their overall book assets by 1.0%-1.3% compared to unlevered firms. These results hold across the different regression models, and are statically significant at the 5% level except for column 9.

Taken together, the regression results are consistent with theories of risk-shifting at highly levered, distressed firms. They uncover a new channel through which firms increase their risk-taking, which is relatively cheap to execute and unmonitored by creditors and investors.

C. Robustness and Extensions

In Table 4 we provide several robustness tests and extensions. First, in columns 1-3, we replace the indicator variable *High short-term liabilities* with the continuous variable *Short-term liabilities*, defined as the ratio of short-term liabilities to total book assets as of the onset of the oil price crisis at the end of the second quarter of 2014. Second, in column 4-6, we scale short-term liabilities by total liabilities rather than total book assets. The purpose of these analyses is to ensure that the results do not depend on the definition of *High short-term liabilities*. As before, we consider three different measures of risky financial assets as the dependent variables, and, for brevity, only report the results from the most conservative regression model that includes year and firm fixed effects. The key independent variable is the interaction term *Short-term liabilities x Crisis*.

The results in columns 1-6 of Table 4 indicate that firms with higher levels of short-term liabilities invested more in risky financial assets. The interaction term *Short-term liabilities x Crisis* is positive, economically meaningful, and mostly statistically significant across the different specifications. Based on columns 1-3, for example, an increase of one standard deviation in short-term liabilities (Standard deviation = 0.094) corresponds to an increase of 3.95% in the ratio of risky financial assets to total financial assets, and an increase of 0.46% in the ratio of risky financial assets to total book assets. Overall, these findings are consistent with the evidence in Table 3 that firms with the highest level of short-term liabilities at the onset of the crisis increased their investments in risky financial assets during the crisis.

In columns 7-9, we re-estimate the regressions in a subsample of 102 oil producers that exclude large oil firms with diversified operations such as oil refinement. The purpose of these analyses is to investigate whether the results continue to hold after excluding large firms that may vary in their exposure to the oil price crisis and consequently have different degrees of risk-shifting incentives. For instance, operating in oil refinement entails a directionally opposite exposure to oil price declines. Despite the lower test power in the smaller sample of focused oil producers, the results are consistent with the full-sample estimates. Highly levered oil producers have increased the fraction of their financial portfolio invested in risky financial assets by 7.6% compared to unlevered firms following the onset of the oil price crisis.

To further address the endogeneity problem associated with risk-taking and identify the channel through which the oil price crisis induced firms to increase risk-taking, we estimate first-stage regressions that identify exogenous changes in *Profitability* and *Profitability x High short-term liabilities* from 2012 to 2016 based on the 2014 oil price shock and use the fitted changes in from the first stage regressions to explain changes in risk-taking from 2012 to 2016 in the second stage regressions.

These results are shown in Table 5. Columns 1 and 2 report estimates from the first stage regressions. In column 1, the dependent variable is *Profitability*, defined as the Return on Assets (ROA), or the ratio of net income to book assets. The coefficient on the indicator variable *Crisis* is negative and highly statistically and economically significant. Following the onset of the oil price crisis, the ROA of oil and gas firms has decreased, on average, by 19%. This result is highly statistically significant at the 1% level. It suggests that the oil price crisis had a material negative impact on firms' cash flows. In column 2, the dependent variable is *Profitability x High short-term liabilities*. Here, too, the coefficient on the indicator variable *Crisis* is negative and highly statistically and economically significant, suggesting that the oil price crisis had a significant negative effect on the profitability of highly levered firms.

In columns 3-8, we investigate the impact of the negative cash flow shock resulting from the oil price crisis on firms' risk-taking. We regress the three measures of firms' risky financial asset portfolios on the fitted profitability from the first stage and the fitted interaction of profitability with the indicator *High short-term liabilities.* The results indicate that highly levered firms with more negative profitability shocks have increased their holdings of risky financial assets following the onset of the crisis. The coefficients on the interaction term *Predicted profitability x High short-term liabilities* are negative and statistically significant at the 5% or 10% levels. A decline of one standard deviation in profitability at highly-levered firms implies an increase of 24.87% in the ratio of risky financial assets to total financial assets, and an increase of 2.41% in the ratio of risky financial asset to total book assets.

These findings collectively show that the oil price crisis had an economically substantial negative effect on the profitability of oil and gas companies. In response to the cash flow squeeze, levered firms with short-term debt positions exposed to rollover and bankruptcy risks, increased the risk of their financial asset portfolios. These findings are consistent with the predications of risk-shifting theories (e.g., Jensen and Meckling (1976)), providing novel evidence on the role of risky financial assets held by industrial firms in risk-shifting, and highlighting the importance of short-maturity liabilities amid transient cash flow shocks.

In Table 6, we provide direct evidence on the role of distress and subsequent bankruptcy in risk-taking by collecting information on bankruptcy filings of oil and gas companies following the onset of the oil price crisis. We obtain these data from the UCLA-LoPucki Bankruptcy Research Database for a total of 5 bankruptcy filings in 2015 and 16 bankruptcy filings in 2016.

To investigate the role of bankruptcy in risky financial asset holdings, we construct an indicator variable *Bankruptcy* that equals 1 for firms that file for bankruptcy in the subsequent one or two years, respectively, and 0 otherwise. As before, we estimate panel regressions explaining risky financial investments that include time-varying controls and firm and year fixed effects.

Table 6 shows that distressed firms, which end up filing for bankruptcy following the oil price crisis, invest more in risky financial assets. The coefficients on *Bankruptcy* are positive across all measures of risky financial assets and statistically significant in 5 out of the 9 cases. The economic magnitudes are also nontrivial. Future bankruptcy filings imply an increase of 22.9% in the ratio of risky financial assets to total financial assets, and an increase of 0.2% in the ratio of risky financial asset to total book assets with frim and time fixed effects. These findings provide direct evidence that distressed firms used their financial asset portfolios to risk-shift following the onset of the oil price crisis.

D. The Role of Collateral

In this subsection, we investigate the role of collateral in corporate risk-shifting. We argue that collateralized debt is less prone to the agency problem of asset substitution or risk-shifting by borrowers (Jensen and Meckling (1976)). Prior work has shown that collateral mitigates information asymmetry between equity holders and debtholders (Besanko and Thakor (1987) and Boot, Thakor and Udell (1991)), and reduces costly monitoring by lenders (Rajan and Winton, 1995). Consequently, risk-shifting should occur primarily in uncollateralized debt.

In Tables 7 and 8, we introduce collateral into the analyses. Table 7 presents univariate evidence on the role of collateral in risk-shifting by double-sorting firms on short-term debt and collateral. Panel A considers firms with low outstanding short-term debt positions at the onset of the oil price crisis, whereas Panel B considers firms with high outstanding short-term debt positions. Next, each subsample is sorted around the median collateral value, defined as the ratio of tangible assets to total book assets, into two subsamples of low and high collateral value. For each subsample, Table 7 reports the average risky financial asset holdings in each sample year from 2011 to 2016. As before, we consider three measures of risky financial asset holdings based on their dollar value, ratio to total financial asset holdings, and ratio to total book assets.

The estimates in Table 7 suggest that the increase in financial risk-taking at highly levered firms is concentrated in firms with low collateral. At highly levered firms with low levels of collateral, the ratio of risky financial assets to total financial assets increases from 26.7% in 2013 to 46.4% in 2016 and the difference is statistically significant at the 1% level (t-statistic = 3.63). In contrast, at highly levered firms with high levels of collateral, the ratio of risky financial assets to total financial assets only increases from 16% in 2013 to 21% in 2016, and the difference is statistically insignificant at conventional levels (t-statistic = 0.73). We obtain similar results for the other two measures of risky financial asset holdings. The average dollar value of risky financial assets increases significantly by 66.26% from 2013 to 2016 (t-statistic =2.165) at high short-term debt, low-collateral firms. The ratio of risky financial assets to total book assets increases from 3.4% in 2013 to 7.9% in 2016 (t-statistic = 3.141) at these firms. In contrast, risky financial asset holdings do not increase significantly at levered firms with high collateral value or at unlevered firms irrespective of their collateral value.

In Table 8, we present regression evidence on the role of collateral values in corporate risk-shifting. The table provides estimates from triple-differences regressions explaining the riskiness of firms' financial asset portfolios. In columns 1, 4, and 7, the dependent variable is the logarithm of the dollar value of firms' risky financial assets. In columns 2, 5, and 8 the dependent variable is the ratio of risky of financial assets to total financial assets. In columns 3, 6, and 9, the dependent variable is the ratio of risky financial assets to total book assets. The regression models include the set of time-varying controls from the previous tables, as well as different combinations of firm and year fixed effects.

The main variable of interest it the interaction term *High short-term liabilities x Low collateral x Crisis*, which captures the effect of high outstanding short-term liabilities and low collateral values at the onset of the crisis on the risk of firms' financial asset portfolios following the onset of the crisis.

The results in Table 8 suggest that the increase in risky financial assets following the onset of the oil price crisis is concentrated in firms with high outstanding short-term positions and low collateral values. In particular, the interaction term *High short-term liabilities x Low collateral x Crisis* is economically large and has a positive sign across all regression specifications. The point estimates imply an increase of 14.2% in the ratio of risky financial assets to total financial assets, and an increase of 4.7% in the ratio of risky financial asset to total book assets at highly levered firms with low collateral values. These estimates are statistically significant at conventional levels in 6 out of the 9 regression models.[5]

E. The Role of Hedging

In this subsection, we investigate the influence of derivative hedging on corporate risk-shifting. While theory on corporate risk management suggests that hedging can effectively mitigate risk and enhance firm value (e.g., Froot et al. (1993), Leland (1998), Chidambaran et al. (2001)), prior empirical studies do not offer conclusive evidence that hedging matters for firm policies and value (e.g., Guay and Kothari (2003), Jin and Jorion, (2006), Bartram et al. (2011)). We seek to provide novel evidence on an unexplored channel through which hedging can influence incentives, policies, and outcomes in firms, namely the risk-shifting channel. We conjecture that derivative hedging can weaken the incentives to risk-shift by reducing the exposure of highly levered firms to adverse shocks.

To investigate the role of derivative hedging in risk-shifting, we hand-collect detailed data on oil and gas firms' use of oil derivative contracts at the onset of the 2014 oil price crisis. Using these data, we classify the sample firms into hedging and non-hedging firms, and estimate the effect of hedging on risky financial investments in Tables 9 and 10.

[5] In unreported analyses, we obtain similar results in Two-Stage Least Squares regressions. The first-stage regressions identify exogenous changes in *Profitability, Profitability x High short-term liabilities*, and *Profitability x High short-term liabilities x Low collateral* from 2012 to 2016 based on the 2014 oil price shock. The second-stage regressions use the fitted changes in from the first stage regressions to explain changes in risk-taking from 2012 to 2016.

Table 9 presents univariate evidence on the role of hedging in risk-shifting by double-sorting firms on short-term debt and hedging. Panel A considers firms with low outstanding short-term debt positions at the onset of the oil price crisis, whereas Panel B considers firms with high outstanding short-term debt positions. Each subsample is then sorted based on whether or not the firm uses derivatives to hedge oil price risk, and the table reports the average risky financial asset holdings in each subsample year from 2011 to 2016.

The estimates in Table 9 suggest that the increase in financial risk-taking at highly levered firms is considerable larger in unhedged firms. At highly levered unhedged firms, the ratio of risky financial assets to total financial assets increases from 25.8% in 2013 to 44.8% in 2016. In contrast, at highly levered hedged firms, the ratio of risky financial assets to total financial assets only increases from 13.5% in 2013 to 22.1% in 2016.

In Table 10, we present regression evidence on the role of hedging in corporate risk-shifting. The table provides estimates from triple-differences regressions explaining the riskiness of firms' financial asset portfolios. The main variable of interest it the interaction term *High short-term liabilities x No hedging x Crisis*, which captures the effect of high outstanding short-term liabilities and the absence of oil price hedging at the onset of the crisis on the risk of firms' financial asset portfolios following the onset of the crisis.

The results in Table 10 suggest that the increase in risky financial assets following the onset of the oil price crisis is larger in unhedged firms with high outstanding short-term positions. In particular, the interaction term *High short-term liabilities x No hedging x Crisis* is economically nontrivial and has a positive sign across all regression specifications. For instance, the point estimates imply an increase of 8-9% in the ratio of risky financial assets to total financial assets. Note, however, that the estimates are only statistically significant at conventional levels in 4 out of the 9 regression models.[6]

5. Real Assets

In this section, we investigate the relation between risky financial investments and risky real investments among oil and gas firms after the onset of the oil price crisis. We start by providing univariate evidence on risky real investments during the sample period 2011-2016. These estimates are shown in Table 11.

We consider three measures of risky real assets that are conceptually similar to the three measures of risky financial assets consider in the previous analyses. The first measure, *Log risky real assets*, is the logarithm of the dollar value of firms' investment in exploratory oil wells. The second measure, *Risky real assets/total real assets* is the ratio of firms' investment in exploratory oil wells to their total investment in both exploratory and established wells. The third measure, *Risky real assets/total book assets* is the ratio of firms' investment in exploratory oil wells to their total book assets.

As before, Table 11 sorts the sample firms based on their outstanding short-term liabilities at the onset of the oil price crisis and reports the average investment in risky real assets in each year from 2011 to 2016. In contrast to risky financial investments, Table 11 shows no increase in risky real investments following the onset of the oil price crisis. The differences between average risky real investments in 2013 and 2016 are economically small, flip signs, and are never statistically significant at conventional levels. These results hold true for both sets of firms – those with low outstanding short-term liabilities at the onset of the crisis and those with high outstanding short-term liabilities.

[6] In unreported analyses, we obtain similar results in Two-Stage Least Squares regressions. The first-stage regressions identify exogenous changes in *Profitability, Profitability x High short-term liabilities*, and *Profitability x High short-term liabilities x No Hedging* from 2012 to 2016 based on the 2014 oil price shock. The second-stage regressions use the fitted changes in from the first stage regressions to explain changes in risk-taking from 2012 to 2016.

Table 12 provides multivariate regression evidence on the riskiness of real investments at oil and gas firms following the onset of the 2014 oil price crisis. The table provides estimates from difference-in-differences regressions explaining the riskiness of firms' investments in risky real assets, measured by their investment in exploratory oil wells. In columns 1-3, the dependent variable is the logarithm of the dollar value of firms' risky real investments. In columns 4-6, the dependent variable is the ratio of risky of real investments to total real investments. Finally, in columns 7-9, the dependent variable is the ratio of risky real investments to total book assets. For each dependent variable, we report three regression models with different fixed effects. The first model does not include firm or year fixed effects. The second model includes firm fixed effects. The third model includes both firm and year fixed effects. As before, the main variable of interest it the interaction term *High short-term liabilities x Crisis*. This variable captures the effect high outstanding short-term liabilities at the onset of the crisis on the risk of firms' real investments following the onset of the crisis.

The results in Table 12 show that firms with high outstanding short-term liabilities have not materially changed their investment in risky real assets following the onset of the oil price crisis. While the coefficients on *High short-term liabilities x Crisis* are always negative, they are never statistically significant at conventional levels.

Collectively, the findings suggest that following the oil price crisis, highly levered oil and gas firms increased their risk primarily through risky financial investments rather than risky real investments. There are several reasons why firms would prefer to risk-shift by investing in risky financial assets rather than risky real asset. First, compared with traditional real assets, financial assets are more liquid, easier to access, and carry substantially lower transaction costs. Second, trading in risky financial assets is less visible, does not require an upfront investment in physical or human capital, and can generate immediate/accelerated payoffs. Third, debt covenants typically restrict real investment and payout (Chava et al. (2010)), but do not restrict financial investments.

We also provide descriptive evidence on the relation between investments in risky financial assets and risky real assets. Table 13 provides estimates from triple-differences regressions explaining the riskiness of firms' investments in risky real assets. In columns 1-3, the dependent variable is the logarithm of the dollar value of firms' risky real investments. In columns 4-6, the dependent variable is the ratio of risky real investments to total real investments. Finally, in columns 7-9, the dependent variable is the ratio of risky real investments to total book assets. For each dependent variable, we include the corresponding measure of risky financial investments, and report three regression models that correspond to three different thresholds of short-term liabilities: tope tercile, top quartile, and top quintile. All regression models include year and firm fixed effects.

The main variable of interest it the interaction term *High short-term liabilities x Crisis x Risky financial assets*. This variable captures the relation between investment in risky real assets and risky financial assets during the oil price crisis at firms with high outstanding short-term liabilities at the onset of the crisis. As mentioned above, the indicator variable, *High short-term liabilities*, equals one if a firm's short term liabilities are classified in the top tercile, top quantile and top quintile at the end of the second quarter of 2014.

The results in Table 13 suggest an inverse relation between investments in risky assets and risky financial assets. Across all 9 columns, the triple interaction term *High short-term liabilities x Crisis x Risky financial assets* is negative. It is also statistically significant in 6 of the 9 regressions. Furthermore, as the threshold defining *High short-term liabilities* increases, the relation between risky real investments and risky financial investments becomes more negative, that is, there is more substitution between the two types of risky investments.

These findings indicate that at highly levered firms, higher financial risk-taking is correlated with lower real investment risk, and more so in more highly levered firms. One interpretation of these findings is that firms substitute risky financial investments for risky real investments, and increasingly so as they face greater short-term debt pressures that require accelerated payoffs. To the extent that real investment is costlier to adjust, this evidence indicates that firms that face lower real investment risk increase their overall risk by investing in risky financial assets. Furthermore, these findings reconcile the predictions of agency models of risk-shifting (Jensen and Meckling (1976)) with the results in Gilje (2016), who shows that distressed firms did not increase their investment in risky real assets (exploratory oil wells). The findings in this paper suggest that distressed firms do in fact increase their risk - they do so through their financial asset portfolio rather than their real investments studies by prior research.

In the final set of analyses, we investigate the effect of risky financial investments on firms' overall risk. If firms substitute risky financial investments for risky real investments, the overall effect on the firm's risk can go either way because the reduction in the riskiness of the real asset portfolio can offset the increase in the riskiness of the financial asset portfolio. Moreover, risky financial investments can eliminate idiosyncratic risk by diversifying the firm's investment portfolio. Under this scenario, the link between risky financial investments and risk-shifting becomes less clear.

We investigate this possibility by estimating the relation between risky financial investments and the firm's overall risk, as measured by the volatility of the firm's cash flows, profitability, or asset growth rates following the onset of the 2004 oil price crisis. These results are presented in Table 14, which report estimates from regressions in which the dependent variable is one of the above measures of the firm's overall risk.

Table 14 shows that across all measures of volatility, higher risky financial investments following the onset of the crisis lead to higher levels of volatility. This result is evident from the positive coefficient on the interaction term *Risky financial assets x Crisis*. The estimates are statistically significant at conventional levels in 7 of the 9 regressions. The economic magnitudes are nontrivial. For example, an increase of one standard deviation in *Risky financial assets/book* *assets* leads to an increase of 39.7% in the annualized of cash flows, 0.8% in the annualized volatility of profits, and 1.0% in the annualized volatility of asset growth. Overall, these results suggest that risky financial investments do not reduce overall risk by substituting for risky real investments or by diversifying a firm's holdings portfolio, consistent with their role in corporate risk-shifting.

6. Conclusion

We study how nonfinancial distressed firms use financial securities to increase their risk. Exploiting the 2014 oil price shock as an exogenous negative shock to firm profitability, and hand-collecting detailed data on firms' financial asset portfolios, we find that firms with large outstanding positions of short-term debt, primarily uncollateralized and unhedged, at the onset of the crisis substantially increased their investments in risky financial assets such as corporate bonds, stocks, and mortgage backed securities. In contrast, these firms did not increase their investment in risky real assets such as exploratory oil wells. Thus, while most empirical research on agency problems at distressed firms focused on real investments, often with no or mixed results, our evidence shows that distressed firms take on more risk through their financial asset portfolios.

We put forth several reasons why firms would prefer to risk-shift by investing in risky financial assets rather than risky real assets. First, compared with traditional real assets, financial assets are more liquid, easier to access, and carry substantially lower transaction costs. Second, trading in risky financial assets is less visible, does not require an upfront investment in physical or human capital, and can generate immediate/accelerated payoffs. Third, debt covenants typically restrict real investment and payout, but do not restrict financial investments. Fourth, financial assets are typically reported on the balance sheet as corporate cash holdings, and consequently, camouflage risk-shifting as investments in seemingly safe asset classes.

Our findings have important implications because the financial asset portfolios of nonfinancial firms are large in size, typically opaque, with poor disclosure requirements and little monitoring, and therefore can be used to risk-shift, with the expected detrimental consequences of the agency/moral hazard problem of asset substitution. Our findings suggest that increased disclosure standards and monitoring of corporate financial investments, including through debt covenants, may alleviate concerns about risk-shifting.

Overall, our paper highlights several key factors in corporate risk-shifting. First, it demonstrates that investment in risky financial securities rather than risky real assets can be a preferred conduit for risk-shifting. Second, it highlights the importance of debt maturity in risk-shifting by showing that outstanding, time-pressing short-term obligations rather than long-term obligations trigger risk-shifting behavior. Third, it shows that collateralized debt is significantly less likely to result in risk-shifting at distressed firms. Fourth, it provides one of the first evidence on the role of derivative hedging in curbing risk-shifting. We believe that future empirical work should continue to investigate the role of these three crucial ingredients – financial asset portfolios, debt maturity, collateral value, and derivative hedging – in corporate risk-shifting in other settings.

References

Andrade, Gregor, and Steven N. Kaplan. "How costly is financial (not economic) distress? Evidence from highly leveraged transactions that became distressed." The Journal of Finance 53, no. 5 (1998): 1443-1493.

Almeida, Heitor, Murillo Campello, Bruno Laranjeira, and Scott Weisbenner. "Corporate Debt Maturity and the Real Effects of the 2007 Credit Crisis." Critical Finance Review 1, no. 1 (2012): 3-58.

Andrade, Gregor, and Steven N. Kaplan. "How costly is financial (not economic) distress? Evidence from highly leveraged transactions that became distressed." The Journal of Finance 53, no. 5 (1998): 1443-1493.

Ang, Andrew, Robert J. Hodrick, Yuhang Xing, and Xiaoyan Zhang. "The cross- section of volatility and expected returns." The Journal of Finance 61, no. 1 (2006): 259-299.

Aretz, Kevin, Shantanu Banerjee, and Oksana Pryshchepa. "In the Path of the Storm: Does Financial Distress Cause Non-Financial Firms to Risk-Shift?". Review of Finance, forthcoming, (2018).

Acharya, Viral V., and S. Viswanathan. "Leverage, moral hazard, and liquidity." The Journal of Finance 66, no. 1 (2011): 99-138.

Bartram, S., Brown, G. and Conrad, J. (2011), The effects of derivatives on firm risk and value, Journal of Financial and Quantitative Analysis 46, 967-999.

Becker, Bo, and Per Strömberg. "Fiduciary duties and equity-debtholder conflicts." The Review of Financial Studies 25, no. 6 (2012): 1931-1969.

Besanko, David, and Anjan V. Thakor. "Competitive equilibrium in the credit market under asymmetric information." Journal of Economic Theory 42, no. 1 (1987): 167-182.

Boot, Arnoud WA, Anjan V. Thakor, and Gregory F. Udell. "Secured lending and default risk: equilibrium analysis, policy implications and empirical results." The Economic Journal (1991): 458-472.

Brav, Alon, John R. Graham, Campbell R. Harvey, and Roni Michaely. "Payout policy in the 21st century." Journal of financial economics 77, no. 3 (2005): 483-527.

Calomiris, Charles W., and Charles M. Kahn. "The role of demandable debt in structuring optimal banking arrangements." The American Economic Review (1991): 497-513.

Campbell, John Y., Jens Hilscher, and Jan Szilagyi. "In search of distress risk." The Journal of Finance 63, no. 6 (2008): 2899-2939.

Chava, Sudheer, and Amiyatosh Purnanandam. "Is default risk negatively related to stock returns?" The Review of Financial Studies 23, no. 6 (2010): 2523-2559.

Chen, Zhiyao, and Strebulaev, Ilya. A., "Macroeconomic risk and idiosyncratic risk-taking", Review of Financial Studies, forthcoming, 2018.

Chen, Zhiyao, Strebulaev, Ilya. A., Xing, Yuang, & Zhang, Xiaoyan, 2018. "Strategic Risk Shifting and the Idiosyncratic Volatility Puzzle: An Empirical Investigation", Rock Center for Corporate Governance at Stanford University, Working Paper No. 182.

Chidambaran, N.K., Fernando, C. and Spindt, P. (2001), Credit Enhancement through financial engineering: Freeport-McMoran's gold-denominated depositary shares, Journal of Financial Economics 60, 487-528.

Della Seta, Marco, Erwan Morellec, and Francesca Zucchi. "Short-term debt and incentives for risk-taking". Swiss Finance Institute Working Paper, 2018.

Denes, Matthew , "When Do Firms Risk Shift? Evidence from Venture Capital." Carnegie Mellon University Working Paper, 2018.

Duchin, Ran, Thomas Gilbert, Jarrad Harford, and Christopher Hrdlicka. "Precautionary savings with risky assets: When cash is not cash." The Journal of Finance 72, no. 2 (2017): 793-852.

Eisdorfer, Assaf. "Empirical evidence of risk shifting in financially distressed firms." The Journal of Finance 63, no. 2 (2008): 609-637.

Fama, Eugene F., and Merton H. Miller. The theory of finance. Holt Rinehart & Winston, 1972.

Frazzini, Andrea, and Lasse Heje Pedersen. "Betting against beta." Journal of Financial Economics 111, no. 1 (2014): 1-25.

Froot, K., Scharfstein, D. and Stein, J. (1993), Risk management: Coordinating corporate investments and financing policies, Journal of Finance 48, 1629-1658.

Gilje, Erik P. "Do firms engage in risk-shifting? Empirical evidence." The Review of Financial Studies 29, no. 11 (2016): 2925-2954.

Guay, W. and Kothari, S.P. (2003), How much do firms hedge with derivatives? Journal of Financial Economics 70, 423-461.

Jensen, Michael C., and William H. Meckling. "Theory of the firm: Managerial behavior, agency costs and ownership structure." Journal of financial economics 3, no. 4 (1976): 305-360.

Jin, Y. and Jorion, P. (2006), Firm value and hedging: Evidence from U.S. oil and gas producers, Journal of Finance 61, 893-919.

Kaplan, Steven N., and Joshua Rauh. "Wall Street and Main Street: What contributes to the rise in the highest incomes?" The Review of Financial Studies 23, no. 3 (2009): 1004-1050.

Leland, Hayne E. "Agency costs, risk management, and capital structure." The Journal of Finance 53, no. 4 (1998): 1213-1243.

Leland, Hayne E., and Klaus Bjerre Toft. "Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads." The Journal of Finance 51, no. 3 (1996): 987-1019.

Modigliani, Franco, and Merton H. Miller. "The cost of capital, corporation finance and the theory of investment." The American economic review 48, no. 3 (1958): 261-297.

Pryshchepa, Oksana, Kevin Aretz, and Shantanu Banerjee. "Can investors restrict managerial behavior in distressed firms?." Journal of Corporate Finance 23 (2013): 222-239.

Rajan, Raghuram, and Andrew Winton. "Covenants and collateral as incentives to monitor." The Journal of Finance 50, no. 4 (1995): 1113-1146.

Rampini, Adriano A., and S. Viswanathan. "Collateral and capital structure." Journal of Financial Economics 109, no. 2 (2013): 466-492.

Rauh, Joshua D. "Risk shifting versus risk management: investment policy in corporate pension plans." The Review of Financial Studies 22, no. 7 (2008): 2687-2733.

Rauh, Joshua D., and Amir Sufi. "Capital structure and debt structure." The Review of Financial Studies 23, no. 12 (2010): 4242-4280.

Stiglitz, Joseph E., and Andrew Weiss. "Credit rationing in markets with imperfect information." The American economic review 71, no. 3 (1981): 393-410.

Vassalou, Maria, and Yuhang Xing. "Default risk in equity returns." The journal of finance 59, no. 2 (2004): 831-868.

Appendix A: Variable Definitions

A.1. Measures of Risky Financial Assets

log(1 + Risky financial assets) = the logarithm of the dollar amount (in millions) of risky financial assets.

Risky financial assets/financial assets = risky financial assets divided by total financial assets (the sum of safe financial assets and risky financial assets).

Risky financial assets/ book assets = risky financial assets divided by total book assets.

A.2. Measures of Risky Real Assets

log(1 + Risky real assets) = the logarithm of the dollar amount (in millions) invested in risky exploratory wells.

Risky real assets/ real assets = investment in risky exploratory wells divided by the total investment in exploratory and development wells.

Risky real assets/book assets = investment in exploratory wells divided by total book assets.

A.3. Other Regression Variables

Short-term liabilities/book assets = current short-term liabilities (Compustat item LCTQ) divided by total book assets (Compustat item ATQ), calculated as of the onset of the oil price crisis in the second quarter of 2014.

High short-term liabilities = an indicator that equals 1 for firms in the top tercile of *Short-term liabilities/book assets* and 0 otherwise.

Crisis = an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016.

Collateral/book assets = physical assets (Compustat item PPNETQ) divided by total book assets (Compustat item ATQ), calculated as of the onset of the oil price crisis in the second quarter of 2014.

Log(sale) = proxy for firm size, calculated as the logarithm of sales (Compustat item Sale).

Profitability = operating income before depreciation (Compustat item OIBDP) divided by lagged total book assets (Compustat item AT).

Market-to-book = the sum of the equity (Compustat items PRCC_F*CSHO) and total debt (Compustat items DLTT + DLC) divided by total book assets (Compustat item AT).

Dividend ratio = cash dividends (Compustat item DVC) divided by total book assets (Compustat item AT).

Capital investment = capital expenditure (Compustat item CAPEX) divided by total book assets (Compustat item AT).

Total liabilities/book assets = total liabilities (Compustat items DLTT + LCT) divided by total book assets (Compustat item AT).

Bankruptcy = an indicator that equals 1 if a firm will go bankrupt in the next two years.

Hedging = an indicator that equals 1 for firms that have oil derivatives contracts in 2014 in SEC 10-K filing (Item 7A in 10-K reports).

Volatility of cash flow = the annualized standard deviation of quarterly operating income percentage growth rates, measured over the 12 quarters (Compustat item OIBDPQ).

Volatility of profitability = the annualized standard deviation of quarterly profits, measured over the 12 quarters.

Volatility of asset growth rates = the annualized standard deviation of quarterly asset percentage growth rates, measured over the 12 quarters (Compustat item ATQ).

Figure 1. Brent Crude Oil Price

This figure plots the Brent crude oil price (USD per barrel) from Jan 2012 to Jan 2017. (Source: https://www.focus-economics.com/commodities/energy/brent-crude-oil.)

Table 1. Summary Statistics

This table reports summary statistics. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. All variable definitions appear in Appendix A.

Panel A: Risky Assets

Variable	N	Mean	Standard deviation	Min	25th Percentile	Median	75th Percentile	Max
Risky financial assets ($millions)	732	683	3217	0	0.014	16	144	30985
Risky financial assets/financial assets	732	0.320	0.324	0.000	0.003	0.210	0.573	1.000
Risky financial assets/book assets	732	0.037	0.105	0.000	0.000	0.010	0.033	1.000
Risky real assets ($millions)	492	731	5188	0	5	84	412	112591
Risky real assets/real assets	492	0.327	0.307	0.000	0.059	0.237	0.550	1.000
Risky real assets/book assets	492	0.094	0.160	0.000	0.010	0.040	0.112	1.000

Panel B: Firm-level Variables

Variable	N	Mean	Standard deviation	Min	25th Percentile	Median	75th Percentile	Max
Short-term liabilities/book assets	727	0.134	0.087	0.000	0.081	0.113	0.163	0.604
Total liabilities/book assets	727	0.464	1.080	0.000	0.252	0.346	0.435	10.000
Collateral/book assets	727	0.687	0.225	0.000	0.583	0.754	0.868	0.952
Log(sale)	727	6.525	2.567	0.001	5.022	6.594	8.174	12.980
Profitability	727	0.091	0.232	-1.500	0.063	0.134	0.205	0.566
Market-to-book	727	1.328	1.249	0.090	0.768	1.001	1.388	10.000
Capital investment	727	0.164	0.116	0.000	0.065	0.140	0.234	0.432
Dividend ratio	727	0.018	0.036	0.000	0.000	0.000	0.020	0.299
Volatility of cash flow growth	627	5.829	6.176	0.221	1.131	3.908	8.417	31.97
Volatility of profitability	632	0.140	0.160	0.006	0.037	0.076	0.188	0.860
Volatility of asset growth rates	633	0.367	0.296	0.014	0.1465	0.279	0.494	1.588

Panel C: Covariate Balance (2nd Quarter of 2014)

	High short-term liabilities firms		Low short-term liabilities firms			
Variable	Mean	Standard deviation	Mean	Standard deviation	Difference	t-statistics
Short-term liabilities/book assets	0.221	0.094	0.089	0.033	0.132	11.250
Total Liabilities/book assets	0.504	3.016	0.383	2.457	0.121	3.365
Log(sale)	5.652	0.031	5.151	0.028	0.501	0.974
Profitability	0.033	1.512	0.034	0.694	-0.001	-0.178
Market-to-book	1.661	0.013	1.385	0.019	0.276	1.369
Capital investment	0.072	0.062	0.091	0.059	-0.019	-1.635
Dividend ratio	0.008	0.189	0.012	0.184	-0.004	-1.199

Table 2. Univariate Evidence on Short-term Liabilities and Risky Financial Investments

This table presents evidence on the relation between a firm's outstanding short-term liabilities at the onset of the oil price crisis (the second quarter of 2014) and its investments in risky financial assets. Firms in the bottom two terciles of short-term liabilities (measured in the second quarter of 2014) are classified as *Low short-term liabilities* (Panel A) and those in the top tercile are classified as *High short-term liabilities* (Panel B). The table reports annual average investments in risky financial assets for the sample period 2011-2016 as well as a difference-in-means test between 2013 and 2016. T-statistics are reported in brackets. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. All variable definitions appear in Appendix A.

Panel A. Low Short-term Liabilities

Year	2011	2012	2013	2014	2015	2016	2016 - 2013
Risky financial assets($millions)	838 [1.773]	781 [1.942]	755 [1.797]	812 [1.858]	733 [1.780]	565 [1.510]	-190 [-0.648]
Risky financial assets/financial assets	0.344 [9.107]	0.384 [10.239]	0.369 [9.641]	0.394 [9.893]	0.333 [8.524]	0.329 [9.157]	-0.04 [-0.892]
Risky financial assets/book assets	0.041 [3.038]	0.042 [3.077]	0.038 [2.874]	0.041 [3.085]	0.039 [2.949]	0.051 [3.028]	0.013 [1.043]
N of Firms	81	81	81	81	81	81	81

Panel B. High Short-term Liabilities

Year	2011	2012	2013	2014	2015	2016	2016 - 2013
Risky financial assets($millions)	421 [2.678]	441 [2.272]	484 [2.109]	519 [2.286]	734 [1.946]	729 [2.345]	245 [1.966]
Risky financial assets/financial Assets	0.221 [5.211]	0.184 [4.546]	0.212 [5.605]	0.241 [5.789]	0.269 [5.887]	0.334 [7.659]	0.121 [2.705]
Risky financial assets/book assets	0.025 [4.250]	0.022 [3.838]	0.022 [4.665]	0.024 [4.717]	0.032 [3.595]	0.048 [5.506]	0.026 [3.397]
N of Firms	41	41	41	41	41	41	41

Table 3. Regression Evidence on Short-term Liabilities and Risky Financial Investments

This table reports estimates from difference-in-differences regressions on the relation between short-term liabilities and investments in risky financial assets around the 2014 oil price crisis. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. *High short-term liabilities* is an indicator that equals 1 for firms in the top tercile of Short-term liabilities/book assets and 0 otherwise. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Log risky financial assets			Risky financial assets/ financial assets			Risky financial assets/ book assets		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-1.460***			0.160**			0.080*		
	[-2.910]			[2.180]			[1.730]		
High short-term liabilities	-0.730*			-0.160***			<0.001		
	[-1.930]			[-3.410]			[-0.470]		
Crisis	-0.380***	-0.037		-0.040	-0.006		-0.010	<0.001	
	[-3.520]	[-0.400]		[-1.270]	[-0.189]		[-1.130]	[-0.005]	
High short-term liabilities × Crisis	0.530***	0.526***	0.496***	0.090**	0.093**	0.095**	0.010	0.013*	0.010
	[2.850]	[2.895]	[2.852]	[2.260]	[2.223]	[2.565]	[1.260]	[1.814]	[1.577]
Log(sale)	0.760***	0.237**	0.243***	0.030***	-0.001	-0.001	<0.001	-0.001	-0.001
	[10.520]	[2.355]	[2.627]	[3.490]	[-0.054]	[-0.031]	[-0.870]	[-0.229]	[-0.278]
Profitability	-1.500***	-0.157	0.104	-0.070	-0.036	-0.017	0.030	-0.003	0.004
	[-3.730]	[-0.659]	[0.386]	[-1.020]	[-0.661]	[-0.296]	[1.390]	[-0.388]	[0.567]
Market-to-book	0.070	0.031	0.037	-0.010	-0.012	-0.012	-0.010	<0.001	0.001
	[0.850]	[0.595]	[0.578]	[-0.330]	[-1.092]	[-1.092]	[-1.190]	[0.201]	[0.686]
Dividend ratio	8.250*	0.306	0.144	2.050**	-0.002	-0.018	1.270	-0.019	-0.024
	[1.900]	[0.179]	[0.078]	[2.390]	[-0.007]	[-0.052]	[1.580]	[-0.458]	[-0.546]
Capital investment	-1.570	0.316	0.749	0.030	0.350**	0.379**	-0.130**	0.004	0.017
	[-1.520]	[0.581]	[1.402]	[0.190]	[1.993]	[2.098]	[-1.990]	[0.238]	[1.162]
Total liabilities	-0.120	-0.046	-0.059	<0.001	-0.019*	-0.019*	<0.001	-0.001	-0.001*
	[-1.270]	[-0.948]	[-1.225]	[-0.360]	[-1.771]	[-1.868]	[0.170]	[-1.219]	[-1.811]
N_obs	727	727	727	727	727	727	727	727	727
R^2	0.500	0.904	0.906	0.135	0.667	0.668	0.242	0.936	0.936
Firm FE	N	Y	Y	N	Y	Y	N	Y	Y
Year FE	N	N	Y	N	N	Y	N	N	Y

Table 4. Robustness and Extensions

This table reports robustness tests from panel regressions on the relation between short-term liabilities and investments in risky financial assets around the 2014 oil price crisis. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. In columns 1-3, *Short-term liabilities* are measured as a continuous variable. In columns 4-6, *Short-term liabilities* are scaled by total liabilities rather than total book assets. In columns 7-9, the sample is restricted to a subset of 102 oil and gas producers (sic code from 1300 to 1399). In columns 1, 4, and 7, the dependent variable is *Log risky financial assets*. In columns 2, 5, and 8, the dependent variable is *Risky financial assets/ financial assets*. In columns 3, 6, and 9, the dependent variable is *Risky financial assets/ book assets*. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All regressions include year and firm fixed effects. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

	Continuous measure of short-term liabilities			Alternative measure of short-term liabilities			Subsample of oil and gas producers		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Short-term liabilities × Crisis	2.456*** [2.692]	0.420*** [3.263]	0.049 [1.485]	0.412** [2.314]	0.111** [2.584]	0.009 [1.211]	0.503* [1.959]	0.076* [1.811]	0.015 [1.614]
Log(sale)	0.215* [1.946]	-0.003 [-0.225]	-0.001 [-0.302]	0.289*** [3.147]	0.009 [0.474]	0.000 [-0.017]	0.301*** [2.868]	0.010 [0.464]	0.001 [0.331]
Profitability	0.191 [0.672]	0.005 [0.131]	0.005 [0.763]	0.144 [0.510]	-0.008 [-0.130]	0.005 [0.645]	0.012 [0.043]	-0.045 [-0.763]	0.006 [0.763]
Market-to-book	0.052 [0.789]	-0.007 [-0.933]	0.001 [0.888]	0.063 [1.001]	-0.005 [-0.434]	0.001 [1.105]	0.037 [0.584]	-0.012 [-1.085]	0.001 [1.082]
Dividend ratio	-0.167 [-0.097]	-0.090 [-0.353]	-0.021 [-0.600]	-0.218 [-0.141]	-0.144 [-0.433]	-0.031 [-0.795]	-1.582 [-1.355]	0.042 [0.087]	-0.042 [-0.924]
Capital investment	0.645 [1.230]	0.238* [1.907]	0.015 [1.132]	0.610 [1.181]	0.338** [1.989]	0.014 [1.015]	0.702 [1.313]	0.333* [1.750]	0.021 [1.394]
Total liabilities	-0.055 [-1.160]	-0.014* [-1.951]	-0.001* [-1.754]	-0.059 [-1.313]	-0.019* [-1.851]	-0.001* [-1.724]	-0.044 [-0.925]	-0.018* [-1.767]	-0.001* [-1.686]
N_obs	727	727	727	727	727	727	605	605	605
R^2	0.906	0.690	0.909	0.9053	0.6703	0.9361	0.9041	0.6691	0.9518
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 5. Two Stage Least Square Regression Evidence

This table reports estimates from two-stage least-squares (2SLS) regressions on the relation between short-term liabilities and investments in risky financial assets around the 2014 oil price crisis. Columns 1-2 estimate a first-stage regression that identifies exogenous changes in profitability from 2012 to 2016 based on the 2014 oil price shock. Columns 3-8 use the fitted changes in profitability from the first stage regressions to explain investments in risky financial assets in the second stage regressions. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. *High short-term liabilities* is an indicator that equals 1 for firms in the top tercile of Short-term liabilities/book assets and 0 otherwise. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

	First Stage		Second Stage					
Dependent Variable	Profitability	Profitability × High short-term liabilities	Log risky financial assets		Risky financial assets/ financial assets		Risky financial assets/ book assets	
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Predicted profitability			-14.793*** [-12.416]	24.116*** [132.767]	-1.033*** [-5.543]	2.050*** [70.182]	-0.091** [-2.373]	0.159*** [31.124]
Predicted profitability × High short-term liabilities				-67.956*** [-30.604]		-5.385*** [-15.600]		-0.436*** [-6.025]
Crisis	-0.189*** [-7.316]	-0.051** [-2.425]						
Log(sale)	0.039** [2.072]	0.038** [2.054]	1.590*** [48.460]	0.826*** [45.457]	0.098*** [15.048]	0.038*** [10.202]	0.010*** [10.127]	0.005*** [9.923]
Market-to-book	-0.026 [-1.090]	-0.037** [-2.040]	-0.323*** [-6.751]	-1.943*** [-21.952]	-0.037*** [-4.220]	-0.165*** [-11.021]	-0.003** [-2.233]	-0.013*** [-4.785]
Dividend ratio	0.217 [0.795]	0.089 [0.447]	8.628*** [6.920]	-6.651*** [-6.249]	0.680*** [3.507]	-0.531*** [-3.151]	0.045 [1.288]	-0.053* [-1.857]
Capital investment	-0.182 [-1.132]	0.057 [0.645]	4.908*** [8.049]	-1.617*** [-3.302]	0.505*** [3.758]	-0.012 [-0.111]	0.051*** [2.801]	0.009 [0.666]
Total liabilities	-0.035 [-1.591]	-0.022* [-1.663]	-0.403*** [-8.608]	-0.955*** [-15.523]	-0.034*** [-4.157]	-0.078*** [-7.430]	-0.003** [-2.228]	-0.006*** [-3.506]
N_obs	727	727	727	727	727	727	727	727
R^2	0.6439	0.6590	0.9052	0.9052	0.7196	0.7196	0.8845	0.8845
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y

Table 6. Bankruptcy

This table reports estimates from panel regressions on the relation between future bankruptcy and investment in risky financial assets around the 2014 oil price crisis. The sample consists of 131 oil and gas firms with nonmissing observations from 2014 to 2016, out of which 21 firms filed for bankruptcy during this period. *Bankruptcy* is an indicator that equals 1 if a firm went bankrupt from 2014-2016 based on bankruptcy data from the UCLA-LoPucki Bankruptcy Research Database. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Log risky financial assets			Risky financial assets/ financial assets			Risky financial assets/ book assets		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-1.428*** [-2.879]			0.177* [1.959]			0.100* [1.951]		
Bankruptcy	0.288 [1.071]	0.852* [1.785]	0.993** [2.189]	0.127* [1.936]	0.218* [1.691]	0.229* [1.793]	-0.015 [-1.409]	-0.001 [-0.125]	0.002 [0.116]
Log(sale)	0.751*** [10.661]	0.009 [0.030]	0.143 [0.451]	0.022** [2.099]	-0.061 [-1.321]	-0.049 [-1.014]	-0.007 [-1.096]	-0.007 [-0.767]	-0.006 [-0.591]
Profitability	-1.398*** [-3.160]	-0.307 [-0.570]	0.088 [0.145]	-0.086 [-0.986]	-0.052 [-0.548]	-0.023 [-0.236]	0.004 [0.222]	-0.007 [-0.510]	0.001 [0.083]
Market-to-book	-0.03 [-0.274]	-0.212 [-1.261]	-0.136 [-0.804]	-0.014 [-0.485]	-0.012 [-0.401]	-0.008 [-0.237]	-0.018 [-1.375]	-0.007 [-1.368]	-0.005 [-1.027]
Dividend ratio	7.72 [1.495]	-1.996 [-0.686]	-1.629 [-0.543]	2.285** [2.250]	-0.416 [-0.624]	-0.388 [-0.561]	1.312 [1.477]	0.064 [0.820]	0.07 [0.838]
Capital investment	-2.842*** [-2.670]	0.002 [0.002]	0.663 [0.678]	-0.139 [-0.728]	0.309 [1.142]	0.353 [1.302]	-0.088* [-1.801]	-0.003 [-0.152]	0.015 [0.654]
Total liabilities	-0.088 [-1.213]	0.03 [0.423]	-0.002 [-0.021]	-0.006 [-0.428]	-0.008 [-0.547]	-0.01 [-0.623]	0.002 [0.539]	0.001 [0.754]	<0.001 [0.119]
N_obs	357	355	355	357	355	355	357	355	355
R^2	0.4936	0.8957	0.8993	0.1035	0.7364	0.7379	0.1996	0.94	0.9408
Firm FE	N	Y	Y	N	Y	Y	N	Y	Y
Year FE	N	N	Y	N	N	Y	N	N	Y

Table 7. Univariate Evidence on the Role of Collateral

This table presents evidence on the relation between a firm's outstanding short-term liabilities and collateral at the onset of the oil price crisis (the second quarter of 2014) and its investments in risky financial assets. Firms in the bottom two terciles of short-term liabilities (measured in the second quarter of 2014) are classified as *Low short-term liabilities* (Panel A) and those in the top tercile are classified as *High short-term liabilities* (Panel B). The firms in each panel are then sorted based on their ratio of *Collateral/book assets*. The table reports annual average investments in risky financial assets for the sample period 2011-2016 as well as a difference-in-means test between 2013 and 2016. T-statistics are reported in brackets. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. All variable definitions appear in Appendix A.

Panel A. Low Short-term Liabilities

	Year	2011	2012	2013	2014	2015	2016	2016 - 2013
High collateral	Risky financial assets($millions)	87 [2.605]	99 [2.485]	111 [2.348]	109 [2.613]	96 [2.542]	106 [2.481]	-4 [-0.095]
	Risky financial assets/ financial assets	0.33 [6.150]	0.374 [7.161]	0.353 [6.432]	0.421 [7.256]	0.334 [6.175]	0.317 [6.297]	-0.036 [-0.495]
	Risky financial assets/ book assets	0.019 [3.016]	0.016 [2.861]	0.015 [2.858]	0.018 [3.041]	0.021 [2.913]	0.045 [1.980]	0.03 [1.403]
	N of Firms	41	41	41	41	41	41	41
Low collateral	Risky financial assets($millions)	1608 [1.698]	1481 [1.843]	1416 [1.682]	1534 [1.753]	1385 [1.678]	1036 [1.373]	-380 [-0.640]
	Risky financial assets/ financial assets	0.358 [6.667]	0.394 [7.236]	0.385 [7.151]	0.367 [6.676]	0.332 [5.812]	0.341 [6.58]	-0.044 [-0.845]
	Risky financial assets/ book assets	0.064 [2.439]	0.068 [2.568]	0.062 [2.373]	0.064 [2.498]	0.057 [2.233]	0.057 [2.271]	-0.005 [-0.418]
	N of Firms	40	40	40	40	40	40	40

Panel B. High Short-term Liabilities

	year	2011	2012	2013	2014	2015	2016	2016 - 2013
High collateral	Risky financial assets($millions)	259 [1.736]	230 [1.628]	225 [1.581]	255 [1.444]	189 [1.327]	226 [1.684]	1 [0.011]
	Risky financial assets/ financial assets	0.197 [3.341]	0.139 [2.560]	0.16 [3.051]	0.145 [3.189]	0.189 [3.581]	0.210 [3.704]	0.050 [0.730]
	Risky financial assets/ book assets	0.021 [2.532]	0.015 [1.864]	0.009 [3.016]	0.009 [2.700]	0.01 [2.767]	0.018 [3.487]	0.009 [2.007]
	N of Firms	21	21	21	21	21	21	21
Low collateral	Risky financial assets($millions)	591 [2.101]	663 [1.801]	756 [1.702]	797 [1.877]	1311 [1.754]	1259 [2.071]	501 [2.165]
	Risky financial assets/ financial assets	0.246 [3.967]	0.231 [3.864]	0.267 [5.011]	0.342 [5.304]	0.352 [4.889]	0.464 [8.599]	0.196 [3.628]
	Risky financial assets/ book assets	0.029 [3.461]	0.029 [3.636]	0.034 [4.270]	0.04 [4.643]	0.054 [3.332]	0.079 [5.647]	0.045 [3.141]
	N of Firms	20	20	20	20	20	20	20

Table 8. Regression Evidence on the Role of Collateral

This table reports estimates from triple-differences regressions on the relation between short-term liabilities, collateral, and investments in risky financial assets around the 2014 oil price crisis. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. *High short-term liabilities* is an indicator that equals 1 for firms in the top tercile of *Short-term liabilities/book assets* and 0 otherwise. *Low collateral* is an indicator that equals 1 for firms in the bottom half of *Collateral/book assets* and 0 otherwise. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Log risky financial assets			Risky financial assets/ financial assets			Risky financial assets/ book assets		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-1.240** [-2.360]			0.190** [2.560]			0.060* [1.770]		
High short-term liabilities	-0.570 [-1.140]			-0.160** [-2.480]			0.020 [1.330]		
Low collateral	-0.110 [-0.260]			-0.020 [-0.270]			0.040* [1.900]		
Crisis	-0.390** [-2.380]			-0.010 [-0.280]			0.010 [1.250]		
High short-term liabilities × Low collateral	-0.390 [-0.500]	-0.508** [-2.514]	-0.471** [-2.256]	<0.001 [0.020]	-0.080* [-1.786]	-0.077 [-1.591]	-0.040 [-1.550]	0.008 [0.816]	0.013 [1.004]
Low collateral × Crisis	0.300 [1.120]	0.193 [0.939]	0.203 [0.843]	0.010 [0.220]	0.028 [0.606]	0.020 [0.285]	-0.020 [-1.400]	-0.003 [-0.514]	-0.014 [-1.315]
High short-term liabilities × Crisis	0.090 [0.460]	-0.065 [-0.491]	-0.06 [-0.330]	-0.050 [-0.770]	-0.032 [-0.854]	-0.041 [-0.615]	-0.030* [-1.900]	-0.006 [-0.913]	-0.019 [-1.298]
High short-term liabilities × Crisis × Low collateral	0.460 [1.260]	0.641* [1.898]	0.598 [1.641]	0.150* [1.870]	0.135* [1.933]	0.142 [1.622]	0.050*** [2.740]	0.037*** [3.103]	0.047*** [2.878]
Log(sale)	0.770*** [10.040]	0.232** [2.230]	0.236** [2.148]	0.030*** [3.100]	-0.007 [-0.336]	-0.008 [-0.336]	-0.010 [-1.100]	-0.003 [-0.523]	-0.005 [-0.682]
Profitability	-1.600*** [-4.080]	-0.144 [-0.607]	0.123 [0.472]	-0.070 [-1.060]	-0.010 [-0.158]	0.005 [0.093]	0.010 [0.640]	-0.009 [-1.102]	0.009 [0.820]
Market-to-book	0.060 [0.740]	0.032 [0.651]	0.038 [0.623]	-0.010 [-0.500]	-0.010 [-0.748]	-0.011 [-0.850]	-0.010 [-1.470]	-0.002 [-0.746]	<0.001 [-0.178]
Dividend ratio	8.060* [1.890]	0.219 [0.129]	0.204 [0.111]	2.020** [2.350]	0.090 [0.273]	0.081 [0.238]	1.280 [1.580]	0.001 [0.021]	0.007 [0.146]
Capital investment	-2.560* [-1.940]	-0.175 [-0.295]	0.131 [0.217]	-0.050 [-0.280]	0.108 [0.586]	0.127 [0.645]	-0.090 [-1.480]	0.010 [0.270]	0.032 [0.689]
Total liabilities	-0.100 [-1.240]	-0.025 [-0.546]	-0.037 [-0.773]	<0.001 [0.150]	-0.012 [-1.025]	-0.012 [-0.999]	<0.001 [0.250]	0.001 [0.917]	<0.001 [-0.290]
N_obs	727	727	727	727	727	727	727	727	727
R^2	0.500	0.904	0.9057	0.138	0.661	0.662	0.232	0.878	0.881
Firm FE	N	Y	Y	N	Y	Y	N	Y	Y
Year FE	N	N	Y	N	N	Y	N	N	Y

Table 9. Univariate Evidence on the Role of Hedging

This table presents evidence on the relation between a firm's outstanding short-term liabilities and hedging at the onset of the oil price crisis (the second quarter of 2014) and its investments in risky financial assets. Firms in the bottom two terciles of short-term liabilities (measured in the second quarter of 2014) are classified as *Low short-term liabilities* (Panel A) and those in the top tercile are classified as *High short-term liabilities* (Panel B). The firms in each panel are then sorted based on whether or not they report using *oil derivatives* contracts in 2014. The table reports annual average investments in risky financial assets for the sample period 2011-2016 as well as a difference-in-means test between 2013 and 2016. T-statistics are reported in brackets. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. All variable definitions appear in Appendix A.

Panel A. Low Short-term Liabilities

	Year	2011	2012	2013	2014	2015	2016	2016 - 2013
Hedging	Risky financial assets($millions)	913 [1.264]	765 [1.448]	704 [1.318]	756 [1.387]	622 [1.350]	180 [2.609]	-523 [-0.973]
	Risky financial assets/ financial assets	0.410 [7.399]	0.446 [8.376]	0.408 [7.644]	0.459 [8.322]	0.377 [7.050]	0.380 [7.410]	-0.027 [-0.421]
	Risky financial assets/ book assets	0.022 [3.779]	0.022 [3.899]	0.020 [3.815]	0.026 [3.814]	0.023 [4.011]	0.042 [2.103]	0.023 [1.098]
	N of Firms	43	43	43	43	43	43	43
No Hedging	Risky financial assets($millions)	791 [1.320]	833 [1.340]	858 [1.283]	918 [1.301]	914 [1.283]	1048 [1.320]	190 [1.463]
	Risky financial assets/ financial assets	0.293 [6.103]	0.330 [6.602]	0.347 [6.490]	0.336 [6.095]	0.307 [5.443]	0.279 [5.768]	-0.068 [-1.112]
	Risky financial assets/ book assets	0.064 [2.324]	0.065 [2.335]	0.061 [2.228]	0.059 [2.176]	0.059 [2.179]	0.064 [2.254]	0.002 [0.174]
	N of Firms	38	38	38	38	38	38	38

Panel B. High Short-term Liabilities

	year	2011	2012	2013	2014	2015	2016	2016 - 2013
Hedging	Risky financial assets($millions)	232 [1.993]	156 [1.916]	130 [1.819]	166 [2.114]	275 [1.305]	338 [1.485]	208 [1.090]
	Risky financial assets/ financial assets	0.145 [3.003]	0.134 [2.682]	0.135 [2.889]	0.128 [3.304]	0.166 [3.305]	0.221 [4.191]	0.086 [1.669]
	Risky financial assets/ book assets	0.022 [2.458]	0.016 [2.122]	0.010 [3.154]	0.011 [2.984]	0.014 [2.538]	0.030 [3.105]	0.019 [2.132]
	N of Firms	22	22	22	22	22	22	22
No Hedging	Risky financial assets($millions)	564 [1.812]	704 [1.749]	806 [1.674]	843 [1.778]	1153 [1.492]	1188 [1.777]	382 [1.819]
	Risky financial assets/ financial assets	0.260 [3.754]	0.208 [3.251]	0.258 [4.427]	0.342 [4.835]	0.342 [4.494]	0.448 [6.969]	0.190 [2.678]
	Risky financial assets/ book assets	0.024 [3.273]	0.026 [2.988]	0.031 [3.609]	0.036 [3.930]	0.047 [2.748]	0.065 [4.554]	0.033 [2.599]
	N of Firms	19	19	19	19	19	19	19

Table 10. Regression Evidence on the Role of Hedging

This table reports estimates from triple-differences regressions on the relation between short-term liabilities, hedging, and investments in risky financial assets around the 2014 oil price crisis. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. *High short-term liabilities* is an indicator that equals 1 for firms in the top tercile of *Short-term liabilities/book assets* and 0 otherwise. *No hedging* is an indicator that equals 1 for firms that have no oil derivatives contracts in 2014 and 0 otherwise. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Log risky financial assets			Risky financial assets/ financial assets			Risky financial assets/ book assets		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-1.340** [-2.480]			0.240*** [3.010]			0.050* [1.670]		
High short-term liabilities	-0.160 [-0.380]			-0.110* [-1.680]			0.040* [1.690]		
No hedging	-0.840* [-1.950]			-0.260*** [-4.030]			0.010 [0.890]		
Crisis	-0.440*** [-2.980]			-0.040 [-1.060]			<0.001 [-0.040]		
High short-term liabilities × No hedging	-0.130 [-0.170]	-5.943*** [-6.190]	-5.776*** [-5.497]	0.180* [1.940]	-0.661*** [-3.495]	-0.655*** [-3.265]	-0.040 [-1.370]	-0.066* [-1.885]	-0.065 [-1.598]
No hedging × Crisis	0.190 [0.960]	-0.085 [-0.797]	-0.095 [-0.516]	0.020 [0.290]	-0.009 [-0.206]	-0.004 [-0.062]	-0.010 [-0.920]	-0.003 [-0.471]	-0.007 [-0.687]
High short-term liabilities × Crisis	0.300 [1.330]	0.297* [1.742]	0.280 [1.309]	0.050 [0.800]	0.059 [1.336]	0.064 [1.051]	-0.010 [-1.000]	0.003 [0.478]	<0.001 [-0.029]
High short-term liabilities × Crisis × No hedging	0.340 [0.890]	0.565* [1.663]	0.551 [1.472]	0.080 [1.070]	0.090 [1.279]	0.084 [1.001]	0.030* [1.900]	0.024* [1.833]	0.026* [1.814]
Log(sale)	0.760*** [10.79]	0.216** [2.094]	0.221** [1.999]	0.030*** [3.150]	-0.005 [-0.284]	-0.005 [-0.259]	<0.001 [-0.920]	-0.001 [-0.295]	-0.001 [-0.304]
Profitability	-1.450*** [-3.640]	-0.121 [-0.486]	0.130 [0.468]	-0.050 [-0.720]	-0.028 [-0.495]	-0.010 [-0.178]	0.020 [1.370]	-0.002 [-0.359]	0.004 [0.569]
Market-to-book	0.090 [1.050]	0.034 [0.623]	0.039 [0.576]	-0.010 [-0.510]	-0.011 [-0.972]	-0.011 [-0.959]	-0.010 [-1.430]	<0.001 [0.132]	0.001 [0.492]
Dividend ratio	7.890* [1.810]	-0.073 [-0.048]	-0.203 [-0.121]	1.990** [2.260]	-0.083 [-0.255]	-0.109 [-0.327]	1.300 [1.630]	-0.023 [-0.548]	-0.023 [-0.508]
Capital investment	-1.680 [-1.540]	0.255 [0.466]	0.683 [1.305]	0.020 [0.130]	0.342* [1.968]	0.363** [2.055]	-0.090* [-1.770]	-0.001 [-0.037]	0.015 [1.010]
Total liabilities	-0.110 [-1.130]	-0.039 [-0.744]	-0.052 [-0.983]	<0.001 [-0.260]	-0.018* [-1.705]	-0.018* [-1.765]	<0.001 [0.700]	-0.001 [-0.678]	-0.001 [-1.269]
N_obs	727	727	727	727	727	727	727	727	727
R^2	0.5033	0.9049	0.9068	0.1751	0.6694	0.6706	0.2623	0.9364	0.9373
Firm FE	N	Y	Y	N	Y	Y	N	Y	Y
Year FE	N	N	Y	N	N	Y	N	N	Y

Table 11. Univariate Evidence on Short-term Liabilities and Risky Real Investments

This table presents evidence on the relation between a firm's outstanding short-term liabilities at the onset of the oil price crisis (the second quarter of 2014) and its investments in risky real assets (exploratory wells). Firms in the bottom two terciles of short-term liabilities (measured in the second quarter of 2014) are classified as *Low short-term liabilities* (Panel A) and those in the top tercile are classified as *High short-term liabilities* (Panel B). The table reports annual average investments in risky real assets for the sample period 2011-2016 as well as a difference-in-means test between 2013 and 2016. T-statistics are reported in brackets. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. All variable definitions appear in Appendix A.

Panel A. Low Short-term Liabilities

Year	2011	2012	2013	2014	2015	2016	2016 - 2013
Risky real assets($millions)	2365 [1.158]	376 [3.488]	304 [4.755]	48 [4.606]	226 [4.000]	375 [2.986]	71 [0.619]
Risky real assets/ real assets	0.310 [7.754]	0.339 [8.339]	0.302 [7.587]	0.334 [7.696]	0.272 [6.567]	0.372 [7.584]	0.071 [1.521]
Risky real assets/ book assets	0.073 [5.514]	0.072 [7.721]	0.069 [6.075]	0.079 [6.032]	0.045 [4.964]	0.062 [4.714]	-0.007 [-0.474]
N of Firms	55	55	55	55	55	55	55

Panel B. High Short-term Liabilities

Year	2011	2012	2013	2014	2015	2016	2016 - 2013
Risky real assets($millions)	1154 [2.320]	1213 [2.505]	849 [2.924]	757 [3.095]	472 [2.799]	463 [3.322]	-386 [-1.842]
Risky real assets/ real assets	0.341 [5.637]	0.398 [6.664]	0.334 [6.549]	0.327 [6.045]	0.287 [4.949]	0.336 [5.945]	0.002 [0.039]
Risky real assets/ book assets	0.104 [4.228]	0.125 [4.066]	0.096 [4.786]	0.075 [4.578]	0.061 [4.299]	0.067 [3.327]	-0.030 [-1.209]
N of Firms	27	27	27	27	27	27	27

Table 12. Regression Evidence on Short-term Liabilities and Risky Real Investments

This table reports estimates from difference-in-differences regressions on the relation between short-term liabilities and investments in risky real assets around the 2014 oil price crisis. The sample consists of 82 oil and gas firms with nonmissing observations from 2011 to 2016. *High short-term liabilities* is an indicator that equals 1 for firms in the top tercile of Short-term liabilities/book assets and 0 otherwise. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Log risky real assets			Risky real assets/real assets			Risky real assets/book assets		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	-0.680 [-0.990]			0.300*** [2.990]			0.110* [1.810]		
High short-term liabilities	0.200 [0.430]			<0.001 [0.010]			0.070 [1.550]		
Crisis	-0.590*** [-3.980]	-0.169 [-1.242]		0.010 [0.310]	0.031 [0.963]		-0.010 [-0.730]	0.006 [0.593]	
High short-term liabilities × Crisis	-0.330 [-1.160]	-0.280 [-1.079]	-0.134 [-0.525]	-0.050 [-1.010]	-0.044 [-0.959]	-0.020 [-0.459]	-0.030 [-1.520]	-0.028 [-1.260]	-0.022 [-1.026]
Log(sale)	0.760*** [8.380]	-0.048 [-0.326]	0.073 [0.520]	<0.001 [-0.160]	-0.025 [-0.746]	-0.019 [-0.570]	-0.010 [-1.550]	-0.034* [-1.909]	-0.031* [-1.774]
Profitability	-0.330 [-0.550]	1.186*** [2.919]	0.661* [1.804]	-0.050 [-0.540]	0.044 [0.478]	0.125 [1.163]	<0.001 [0.050]	0.038 [1.253]	0.040 [1.031]
Market-to-book	0.260** [2.010]	0.219*** [2.775]	0.143* [1.855]	0.030 [1.170]	0.035** [2.016]	0.034* [1.969]	0.020 [1.100]	0.029 [1.354]	0.027 [1.286]
Dividend ratio	-15.530*** [-4.910]	-5.628* [-1.852]	-5.206 [-1.595]	-2.070*** [-3.460]	-1.917 [-1.451]	-2.002 [-1.512]	-0.620** [-2.320]	-0.475 [-1.141]	-0.486 [-1.144]
Capital investment	2.420** [2.080]	1.493 [1.434]	1.513 [1.534]	0.070 [0.390]	0.030 [0.191]	0.069 [0.420]	0.170** [2.500]	0.059 [0.839]	0.069 [0.895]
Total liabilities	-0.220*** [-2.890]	-0.109 [-0.928]	-0.080 [-0.717]	0.020 [0.740]	0.012 [0.779]	0.009 [0.605]	-0.010 [-1.080]	-0.008 [-1.168]	-0.008 [-1.161]
N_obs	492	492	492	492	492	492	492	492	492
R^2	0.472	0.835	0.846	0.065	0.679	0.686	0.130	0.698	0.700
Firm FE	N	Y	Y	N	Y	Y	N	Y	Y
Year FE	N	N	Y	N	N	Y	N	N	Y

Table 13. The Relation between Risky Financial Assets and Risky Real Assets

This table reports estimates from panel regressions on the relation between short-term liabilities, investments in risky real assets, and investments in risky financial assets around the 2014 oil price crisis. The sample consists of oil and gas firms with nonmissing observations from 2011 to 2016. Across the different columns, *High short-term liabilities* is an indicator that equals 1 for firms in the top tercile, quartile or quintile of Short-term liabilities/book assets, respectively, and 0 otherwise. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. *Risky financial assets* is a continuous variable. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Log risky real assets			Risky real assets/real assets			Risky real assets/book assets		
	Top Tercile	Top Quartile	Top Quintile	Top Tercile	Top Quartile	Top Quintile	Top Tercile	Top Quartile	Top Quintile
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Risky financial assets	-0.105	-0.035	-0.017	0.047	0.039	0.026	-0.313	-0.230	-0.188
	[-0.799]	[-0.300]	[-0.153]	[0.790]	[0.682]	[0.481]	[-1.381]	[-1.086]	[-0.884]
High short-term liabilities × Risky financial assets	0.335	0.367	0.449	-0.185*	-0.228**	-0.225	0.529**	0.680**	0.585*
	[1.584]	[1.620]	[1.650]	[-1.819]	[-2.033]	[-1.562]	[1.996]	[2.160]	[1.905]
High short-term liabilities × Crisis × Risky financial assets	-0.088	-0.140*	-0.197*	-0.135	-0.269**	-0.302**	-0.242*	-0.507**	-0.513**
	[-1.231]	[-1.737]	[-1.675]	[-1.526]	[-2.275]	[-2.142]	[-1.700]	[-2.336]	[-2.465]
Log(sale)	0.073	0.066	0.077	-0.026	-0.024	-0.028	-0.018	-0.017	-0.017
	[0.676]	[0.602]	[0.722]	[-0.932]	[-0.873]	[-0.995]	[-1.449]	[-1.390]	[-1.380]
Profitability	0.727**	0.729**	0.794***	0.136	0.144	0.14.	0.057	0.055	0.056
	[2.464]	[2.478]	[2.765]	[1.449]	[1.566]	[1.488]	[1.511]	[1.467]	[1.507]
Market-to-book	0.091	0.087	0.085	0.024	0.025*	0.026*	0.022	0.022	0.022
	[1.632]	[1.549]	[1.508]	[1.597]	[1.718]	[1.699]	[1.559]	[1.567]	[1.564]
Dividend ratio	-6.658**	-6.697**	-6.691***	-1.500	-1.529	-1.515	-0.564	-0.569	-0.567
	[-2.535]	[-2.621]	[-2.645]	[-1.249]	[-1.273]	[-1.271]	[-1.638]	[-1.646]	[-1.639]
Capital Investment	4.202***	4.132***	4.131***	-0.098	-0.084	-0.079	0.275***	0.272***	0.273***
	[5.370]	[5.301]	[5.351]	[-0.603]	[-0.515]	[-0.486]	[2.825]	[2.789]	[2.795]
Total liabilities	-0.087	-0.090	-0.084	0.002	0.003	0.002	-0.006	-0.006	-0.006
	[-1.139]	[-1.162]	[-1.069]	[0.107]	[0.217]	[0.177]	[-1.209]	[-1.226]	[-1.221]
N_obs	611	611	611	611	611	611	611	611	611
R^2	0.866	0.866	0.867	0.728	0.731	0.730	0.709	0.708	0.709
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Table 14. Overall Risk

This table provides evidence on the effect of investments in risky financial assets around the 2014 oil price crisis on the firm's overall risk, as measured by the volatility of cash flows (columns 1-3), profitability (columns 4-6), and asset growth (columns 7-9) over the 12 quarters following the onset of the crisis. For each measure of volatility, we report results from 3 separate regressions in which risky financial assets are measured as *Log risky financial assets*, *Risky financial assets/financial assets*, and *Risky financial assets/book assets*, respectively. The sample consists of 122 oil and gas firms with nonmissing observations from 2011 to 2016. *Crisis* is an indicator that equals 0 in 2011-2013 and equals 1 in 2014-2016. All variable definitions appear in Appendix A. The t-statistics [in brackets] are adjusted for heteroscedasticity and clustered by firm. Significance levels are indicated as follows: *=10%, **=5%, ***=1%.

Dependent Variable	Volatility of cash flow growth			Volatility of profitability			Volatility of asset growth rates		
Column	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Risky financial assets × Crisis	0.263**	8.494	3.780***	0.007***	0.092	0.076***	0.010**	-0.036	0.095*
	[2.231]	[1.443]	[2.822]	[3.082]	[0.793]	[3.192]	[2.161]	[-0.137]	[1.894]
Log(sale)	1.323*	1.398*	1.074	0.008	0.010	0.003	-0.019	-0.016	-0.023
	[1.880]	[1.920]	[1.657]	[0.505]	[0.637]	[0.230]	[-0.790]	[-0.679]	[-0.981]
Profitability	-0.344	-0.518	-0.405	0.052	0.048	0.050	-0.011	-0.017	-0.014
	[-0.208]	[-0.309]	[-0.249]	[1.448]	[1.376]	[1.383]	[-0.227]	[-0.351]	[-0.294]
Market-to-book	-0.251	-0.226	-0.149	-0.014	-0.013	-0.012	0.004	0.005	0.007
	[-0.779]	[-0.674]	[-0.456]	[-0.690]	[-0.642]	[-0.575]	[0.211]	[0.300]	[0.365]
Dividend ratio	0.722	1.878	2.072	-0.014	0.017	0.018	0.073	0.118	0.119
	[0.118]	[0.310]	[0.331]	[-0.081]	[0.105]	[0.106]	[0.228]	[0.402]	[0.378]
Capital investment	-4.753	-4.393	-4.441	0.038	0.049	0.047	0.038	0.055	0.053
	[-1.260]	[-1.169]	[-1.149]	[0.291]	[0.383]	[0.374]	[0.198]	[0.284]	[0.275]
Total liabilities	-0.006	-0.015	-0.062	0.006	0.005	0.005	-0.025	-0.026	-0.027
	[-0.023]	[-0.055]	[-0.233]	[0.269]	[0.246]	[0.216]	[-1.042]	[-1.050]	[-1.073]
N_obs	627	627	627	632	632	632	633	633	633
R^2	0.610	0.607	0.620	0.663	0.658	0.667	0.710	0.707	0.711
Firm FE	Y	Y	Y	Y	Y	Y	Y	Y	Y
Year FE	Y	Y	Y	Y	Y	Y	Y	Y	Y

Order Flows and Financial Investor Impacts in Commodity Futures Markets*

Mark Ready
University of Wisconsin-Madison

Robert Ready
University of Oregon

May 16, 2019

Abstract

We examine signed order flows and price formation in six commodity futures markets and find that trading in futures markets plays an important role in price discovery. We then use these results to investigate the impacts of financial investors in these markets. We find strong evidence of order flows and price impacts in agricultural futures markets associated with changes in the positions of index traders reported by the CFTC. These impacts are consistent in size with the magnitudes of the index positions, and are concentrated in the minutes just prior to daily futures settlement, when the price impact of trades is generally lowest. While we confirm the positive returns around the issuance of commodity-linked notes documented by Henderson, Pearson, and Wang (2015), we find no evidence that these returns are driven by abnormal order flows. We also find that these notes are too small for the price impacts of hedging trades to explain these returns. We are unable to replicate their finding of significant negative returns on the notes' determination dates.

JEL Codes: G12, G13
Keywords: Commodities, Futures, Order flow, Financialization

*We would like to thank Steffen Hitzemann (discussant) and participants at the 2017 Pacific Northwest Finance Conference, the 2018 Midwest Macroeconomics Meetings, and the 2019 Midwest Finance Association Meetings for their helpful comments.

1 Introduction

Increasing financial and retail investment in commodity futures over the last decade has generated substantial interest in the impact of this investment on commodity markets. While theoretical work predicts that the trading of uninformed financial investors can create price impacts in futures markets, the empirical evidence is mixed. Moreover, most prior empirical work focuses solely on daily or monthly futures returns and thus ignores a basic question: Are the trades of financial investors large enough to materially impact the futures market?

To help answer this question, we examine minute-by-minute signed trading volume (we refer to this as "order flow imbalance", or simply "imbalance") and returns for six major commodity futures markets from January 2007 to April of 2014: WTI crude oil, Brent crude oil, gold, copper, wheat, and corn. These data allow us estimate the price impact associated with a dollar value of buying or selling, which in turn allows us to estimate the magnitude of return we would expect to see in the futures market in response to a given amount of financial investment. Our intraday data also allow us to observe trading at specific times during the trading day, and thereby increase the power of our tests. In particular, we find that volume is highest and price impacts are lowest just prior to the daily futures settlement, so it would be reasonable to expect that uninformed investors may concentrate their trading in this period to reduce the effects of their trades.[1]

The increase in uninformed investment after 2004 is often referred to as the "financialization" of commodity markets, so we use the term "financial investment" as opposed to "retail investment" to refer to participants who are likely to be trading for portfolio reasons, as opposed to traders who may possess superior information. These may be retail investors, but for much of our analysis we are also thinking of institutional investors who are managing index strategies and are responding to the demands of their clients. We focus on two sources of financial investment studied in the literature: weekly changes in the positions of commodity-index traders from from the CFTC, and commodity-linked notes (CLNs) following Henderson, Pearson, and Wang (2015) (henceforth HPW).[2]

[1] This concentration is consistent with the theoretical predictions of Admati and Pfleiderer (1988).

[2] Other work (e.g. Irwin and Sanders (2012) and Bessembinder, Carrion, Tuttle, and Venkataraman (2016)) studies investment in the United States Oil Fund (USO), the largest energy ETF, but finds little evidence of impacts on the level of prices. We also study retail investment in the USO and find similar results, so we relegate this analysis to the Internet Appendix. See Section IA.5.

Changes in the positions of index traders are only available for agricultural futures, so we restrict this portion of our analysis to corn and wheat futures. Our tests are similar in spirit to those of Irwin and Sanders (2011) and Stoll and Whaley (2010), who find little evidence of significant price impacts using daily data. We find similarly insignificant price impact at the daily frequency, but we are able to extend our analysis using our intraday data. Even at the daily frequency, we find significant evidence of imbalance in futures associated with index traders. Additionally, when we examine imbalance and return near the futures settlement, we find strong evidence largely consistent with theoretical models of financialization. When the net long positions of index traders increase (decrease), we see large buying (selling) volume and positive (negative) returns.

The effect of index traders in corn and wheat is strongly statistically significant, but explains only a small amount of the overall price variation for these futures. As an illustration, we find that a one standard deviation increase in the positions of index traders is associated with a positive return of 45 basis points in wheat futures accumulating in the 30-minute intervals prior to the daily settlement. These variations in index flows can explain 10% of the weekly return variation in these intervals prior to settlement, but less than 1% of the overall weekly variation in wheat returns.

Though these trades are unlikely to explain a large amount of volatility, their effect could potentially accumulate over time to have a larger effect on the level of prices. For instance, from January 2009 to January 2010 the net long position in wheat futures held by index investors nearly doubled. Using observed returns around the settlement, we estimate that this buying by index traders would lead to a cumulative increase of approximately 10% in the price of wheat. While this illustrates the potential impact of these index investors, this estimate should be viewed cautiously, as it assumes no reversal of the price impact. While we do not see any price reversal, the insignificance of our estimates at the daily frequency implies that we have little statistical power to discern reversals if they occur slowly.

The order flows from index traders in wheat and corn are large relative to the overall size of the futures market, so it may not be surprising that we find significant price impacts. Again, using wheat as an illustration, we find that the standard deviation of weekly index flows is approximately \$140 million/week. For comparison, the standard deviation of a single minute's imbalance in wheat futures is approximately \$4 million, which rises to approximately \$20 million in the minute prior to futures settlement. This suggests that a \$140 million flow, even when spread over 5 days in the trading week, would need to be executed carefully if it was traded near the daily settlement. We find that the returns associated with these trades are roughly half the size of what we would predict if the entire trade was naively executed near the settlement. This suggests that our estimates provide an upper bound of potential impact, as sophisticated traders are able to minimize their effect on prices.

In our second set of tests, we follow HPW, who find positive average returns in commodity futures on the pricing days CLNs. Their primary analysis focuses on notes of at least $2 million of face value linked to a single commodity with pricing days outside of Goldman Roll Period of Mou (2010). For these notes they find an associated average return of approximately 30 basis points, which rises to 40 basis points when they restrict the sample to notes with a face value of at least $10 million. They attribute these positive returns to the price impact of hedging trades in futures markets made by the issuers of these notes, and suggest that this result supports the theory that uninformed investors can have an impact on commodity prices.

We follow their approach for identifying notes and collect 594 CLNs linked to a single commodity with a face value at least $2 million that were issued before February of 2014, the date when their sample ends. Consistent with HPW, we find a significantly positive average return on the pricing dates outside of the Goldman Roll, with very similar magnitudes. However, when we consider the size of the necessary hedging trades and our estimates of price impacts in futures markets, we find that the notes are far too small for their hedging trades to explain the positive average return. For instance, we have intraday futures data for approximately 80% of the notes, and for the subset of these notes with $10+ million of face value we observe an average return of approximately 33 basis points on the pricing dates outside of the Goldman Roll. However, using our measures of price impact, we would predict that the average price impact of naively executed hedging trades would be only three basis points.[3]

We also report additional results that suggest it is unlikely hedging trade price impact can explain the observed returns. Looking within the trading day, we find no evidence of abnormal order flow on these pricing dates, or of abnormal returns in the minutes around around the pricing

[3]We extend the analysis of HPW and explicitly calculate the size of the hedging trade accounting for the embedded optionality of the notes. See Section 4.1. We also use more readily available daily data to extrapolate our estimates of price impacts to commodities outside of our intraday sample, and consider both the universe of single-commodity CLNs and those for which we have intraday data. We find similar results.

of the notes. Moreover, nearly all of the positive return occurs in the first part of the trading day, with approximately half of the effect accumulating between the prior day's settlement and the open of the market.

In order to understand these results we examine the distribution of notes across the trading month. While HPW argue that notes priced during the Goldman Roll Period should be excluded, it is not clear why this should be the case. Roll trades primarily impact calendar spreads as opposed to the level of the front month future, and, as noted by Mou (2010) and Neuhierl and Thompson (2016), the predictable returns around the roll had disappeared by 2003, when the first notes were issued. Using our intraday data, we find no heterogeneity in the price impact of order flow during the roll period. Accordingly, if the issuance of CLNs causes prices to move as a result of the hedging trades, then there should be similar results during the roll period. Instead, we find that the average returns on CLN pricing dates are near zero during the roll periods.

Outside of the roll periods, we find that the average positive return on CLN pricing dates is entirely explained by notes priced in the five days prior to the last day of the trading month. In this period we also see a substantial increase in the frequency of issuance, particularly for large notes. We refer to this period, which is outside of the Goldman Roll Period, as the "Active Issuance Period". During this period, it is clear that either the demand for or supply of notes increases, so it is notable that these are the notes with the positive average pricing date returns.

Taken together, our results suggest that CLN issuance may be reacting to changes in prices, as opposed to causing them. CLN issuers have flexibility to determine the specific date that the notes are priced and issued, so the association between issue and return prior to settlement suggests that CLN issuers prefer days with rising prices, or that demand for these notes is high on days in which prices are rising. HPW acknowledge this potential endogeneity, and address it by looking at returns on the determination dates when the final payoffs of the notes are set and any hedging trades are unwound. They find a significantly negative average return, primarily on 42 determination dates for notes with at least $10 million of face value. The size of this average return is -42 basis points (t-stat of 2.50), roughly equal in magnitude to the pricing date effect. Since these days are set months or years in advance, this result is not subject to the endogeneity concerns of the pricing dates.

The notes often have complicated embedded optionality, (e.g. call provisions, caps, floors, knock-outs, and buffer regions). Accordingly, many of the notes are either called early or have no sensitivity to the commodity price on the determination date. HPW indicate that they only consider surviving notes with a positive sensitivity to the underlying commodity price on the determination date. We use the contractual terms of each note and the realized price path of the underlying commodity to identify these notes. However, when we attempt to replicate the original HPW result, we find 50 determination dates outside of the Goldman Roll with at least $10 million of face value prior to the end of their sample. On these days we find a statistically insignificant average return of negative nine basis points (t-stat of 0.45).[4] Moreover, we extend the analysis to include the Goldman Roll period, and extend the sample to include notes which matured after the original HPW sample. In no case do we find a significantly negative average determination date return.

1.1 Related Literature

To our knowledge, our paper is the first to systematically examine trade imbalances in several commodity futures markets, and thus the first to document price impacts of order flows, as well as to examine the intraday behavior of signed order flow across several markets.

The study of the impact of financial investors on commodity markets is motivated by a growing theoretical literature. Hamilton and Wu (2014), Acharya, Lochstoer, and Ramadorai (2013), Sockin and Xiong (2015), Baker (2014), Basak and Pavlova (2016), Goldstein and Yang (2017) and others derive theoretical models by which uninformed investors can create price impacts in commodity futures markets. In these models sophisticated investors have limited risk-bearing capacity, so investment flow from uninformed traders has impacts on commodity futures prices.

Our goal in this paper is testing this theoretical prediction. Similarly, some previous papers find evidence supporting the impacts of financial traders including Buyuksahin and Robe (2011), Tang and Xiong (2012), Singleton (2013), Cheng, Kirilenko, and Xiong (2014), and HPW while others find no evidence of impacts, including Stoll and Whaley (2010), Irwin and Sanders (2010), Irwin and Sanders (2011), Silvennoinen and Thorp (2013), Fattouh, Kilian, and Mahadeva (2013),

[4] We are unsure of the source of this discrepancy. The only difference we can discern in our methodology is that we exclude notes linked to multiple commodities in the same sector, but we find none of these notes where the total portion of the face value linked to a single commodity is greater than or equal to $10 million. Therefore, our understanding is that this result should be identical to HPW. We shared our set of notes with HPW in an attempt to resolve this difference on March 11, 2019, and they have not yet shared their data at the time of this writing.

Alquist and Gervais (2013), Hamilton and Wu (2015), and Chari and Christiano (2017).[5] More recently, Yan, Irwin, and Sanders (2019) examine index fund rebalancing and find temporary price impacts in futures markets consistent with the size of our price impact estimates.

While the above empirical work studies prices at daily or longer frequencies, there is a small set of papers that study intraday trading and liquidity in commodity markets. Bessembinder et al. (2016) study liquidity around the predictable rolling of index funds, and Bessembinder (2015) reviews the empirical and theoretical framework for understanding predictable roll trades. However, these papers are focused on predictable calendar spread trades, and are therefore distinct from the price level effects we study here. Raman, Robe, and Yadav (2017) examine price impacts and liquidity over a one-year period around the electronification of WTI oil futures markets in 2007, but do not examine price impacts of financial investors.

Other related work includes Elder, Miao, and Ramchander (2014), who study intraday price patterns in Brent and WTI futures, Marshall, Nguyen, and Visaltanachoti (2011), who study liquidity proxies in commodity prices, and Halova, Kurov, and Kucher (2014), who study price reactions to inventory announcements. However, these papers do not study signed volume and price formation in futures markets.

2 Data

Our data sources include:

- Intraday futures data from Thomson Reuters Tick History from January of 2007 through March of 2014 (we exclude data for the Brent contract prior to January 1, 2008 due to issues in the reported timing of trades).
- A sample of commodity-linked notes obtained from 424b filings obtained from the SEC's EDGAR database covering all notes linked to a single commodity issued prior to February 1, 2014.
- Positions of index traders in corn and wheat futures from the CFTC "Supplementary Positions of Traders" reports from January 2007 to March of 2014.

[5] See Cheng and Xiong (2014) for a review of this literature.

- Daily futures prices from the Commodity Research Bureau from January of 2003 to January of 2019.
- Various commodity-index values from Bloomberg from January of 2003 to January 2019.

We focus on the period through the first quarter of 2014 to be consistent with the sample of HPW. The one exception is that we collect more recent data to examine the determination date returns of notes which mature after their sample. Our intraday data cover six major exchange-traded futures contracts. We include two energy contracts, both the West Texas Intermediate (WTI) contract traded on the NYMEX (now owned by the CME) and the Brent contract traded on the ICE. We also include the gold, corn, wheat, and copper contracts from the CME. In terms of open interest and volume, these contracts are generally largest in their respective commodity classes. Moreover, the gold, corn and wheat contracts on the CME are the dominant futures markets for each commodity. The copper contract on the CME rivals the contract traded on the London Metal Exchange, but generally has slightly lower volume. Nevertheless, even in copper, we find that CME volume plays an important role in price discovery.

Our primary analysis uses 1-minute returns and order imbalance for the near-to-maturity high volume contracts in each market (defined below). As an illustration for how we construct these measures, we first describe them in detail for the WTI crude oil futures.

2.1 Volume Patterns for WTI

WTI futures contracts are available for every month going out five years and for June and December delivery months going out an additional four years. Unlike stock index futures, where nearly all of the trading is in the contract with the nearest delivery dates, there is substantial trading and open interest in longer-dated WTI futures contracts. However, most of this trading in the longer-dated contracts is through calendar spread trades, wherein traders agree to simultaneously buy one maturity and sell another. Most of the trading in a single contract is concentrated in the nearer months. We use data starting in January 2007 and we calculate our imbalance measures using trades and quotes from the Globex platform that are obtained from Thomson Reuters. The NYMEX adopted the CME Globex platform for electronic trading of the WTI contracts in June of 2006 (the CME announced its acquisition of the NYMEX in March of 2008). The Thompson

Reuters data include some floor trades over the earlier part of our sample, and evidently includes most or all of the floor trades starting in March of 2013. Starting in March of 2013, the data also include calendar spread trades. We are able to separately identify floor and calendar spread trades, and we exclude them from our imbalance measures. In order to illustrate the typical pattern in trading volumes, Table 1 shows the WTI contract volumes (in thousands of contracts, each for 1,000 barrels of oil) for the trading days in June 2013.

Table 1 shows that the July 2013 contract last traded on June 20, but most of the trading volume had moved to the August 2013 contract the day before that. The table also shows that calendar spread trading makes up a fairly substantial portion of the front and next month volume, and it constitutes the vast majority of trading in the remaining months. Finally, the table shows that floor trading volume is much smaller than Globex volume, which is a feature common to most futures contracts. In fact, the NYMEX suspended floor trading in WTI futures and many other futures products in July of 2015.

We exclude floor trades because they are executed manually, making it impossible to accurately align them in time with the GLOBEX quotes, and therefore impossible to assign trade direction. We also exclude calendar spread trades from our imbalance measure, motivated in part by results from supplemental tests where we found that the imbalance in the calendar spread trades has little impact on the level of front and next month futures prices.

We classify each Globex single-month trade as a buy or sell by comparing the price to the current quote for that contract, and we aggregate buying and selling volume by minute. We also measure the (logged) return over each minute using quote midpoints as of the end of each minute. Globex trading in WTI futures runs from Sunday night at 6:00 p.m. to Friday night at 5:00 p.m. with one-hour breaks at 5:00 p.m. each day. The bulk of the trading occurs during the day, so we limit our analysis to the time periods from 7:30 a.m. to 4:00 p.m. each day. This time window captures 88% of the total WTI volume in the front and next month contracts.

2.2 Definition of Near Month Imbalance

Most of the trading activity in the contracts that we consider takes place in contracts that have only a few months to expiration. Many users of commodity futures maintain positions in these high volume contracts and roll their positions into later contract months as their contracts near expiration. While this general description applies to all six of our commodities, the specific trading patterns differ.

The WTI and Brent contracts are the easiest to understand. Contracts are available for every calendar month out through 5 years. Trading in the WTI nearest month contract continues until three business days before the 25th calendar day of the month before the delivery month. As illustrated in Table 1, the nearest contract to expiration, which called the front month contract, has the highest trading volume until a few days prior to expiration. The contract expiring in the next calendar month has the next highest volume across all contracts, and it becomes the highest volume contract as the front month contract nears expiration.

The CME procedures for determining daily settlement prices begin by focusing the contract that generally has the highest volume. This is called the "Active Month" for WTI, gold and copper, and is called the "Lead Month" for corn and wheat. We measure returns using the quote midpoints for the Active/Lead Month contracts. We measure imbalances using the difference between buy and sell volume for trades in all months from the front month through the month that is currently the Active/Lead month or is within three weeks of becoming the Active/Lead month. Although we exclude trades that are part of explicit calendar spreads, we recognize that some traders may roll their position using separate individual trades in the two contract months. Our definition of imbalance effectively nets out any trades that are a result of a trader rolling between the nearest contract months. For example, if a WTI trader uses market orders to sell the front month and buy the next month (within three weeks of the front month expiration), our measure will reflect zero net imbalance for those trades.

The Active Month in the WTI futures is the nearest month contract, except for the last two trading days prior to expiration, at which point the next month contract becomes the Active Month. Thus, referring back to Table 1, our return data on June 18, 2013 use the July 2013 contract and our return data on June 19, 2013 use the August 2013 contract. Our imbalance data include both the July 2013 and August 2013 through June 20, 2013, and reflect just the August 2013 contract starting June 21, 2013.

The volume patterns in the other commodities are more complex. Gold futures contracts are available for the nearest three calendar months and for all even calendar months (February, April, June, etc.) for the next two years. Although some trading occurs in odd calendar months that are close to expiration, the volume in odd expiration months is much lower than in the nearby even calendar months. In addition, volume for October tends to be lower that for the other even months. The Active Months in gold are the even months, except for October. The current Active Month is the nearest of these contracts that is not in the final calendar month of trade. For example, on February 1 the April contract becomes the Active Month. The active months in copper are March, May, July, September and December, and the current active month works the same way it does in gold. So for example, on March 1 the May contract becomes the Active Month.

Corn and wheat futures contracts are available for expirations March, May, July, September and December. Trading occurs through the business day prior to the 15th calendar day of the expiration month. For wheat, each of these months is the Lead month until the 12th business day of the calendar month prior to expiration. For example, on the 12th business day of November, the lead month changes from December to March. Corn is very similar to wheat, except September is never considered the Lead month in corn.

3 The Price Impact of Order Flows

We first use our intraday data to estimate the price impact of order flow imbalance in these markets. One common approach is to use the VAR formulation of Hasbrouck (1991). We perform these VARs and report results in the Internet Appendix.[6] The primary takeaways are that the price impacts of both order flow and public return news are mostly permanent at one-minute horizons, and that most of the imbalance in each minute represents an unpredictable innovation. As a result, the VAR estimates of the long-run price impacts of innovations in imbalance are quite close to the coefficients from simple regressions of price change on current imbalance. Thus the results we report below use the univariate regressions.

3.1 Interpreting the Price Impact Measures

Our primary impact measure is the slope in a regression of one-minute returns on one-minute imbalance. In our exposition, we will often refer to the imbalance driving the return. However, it is also possible that imbalance is reacting to returns within each minute. Additionally, informed

[6] See section IA.1.

traders may use limit orders or complicated trading strategies to mask their order flow. Indeed, there is growing evidence that signed volume and its associated price impact is not an effective measure for understanding the impacts of informed traders (see O'Hara (2015) for a review of some these issues).

In this paper however, we are focused on the potential impacts of *uninformed* traders. We simply wish to understand how markets respond, on average, to a given amount of buying or selling. The values that we estimate should then provide an upper bound for the plausible impact of an uninformed trader acting in a rational manner. While it is possible that certain financial traders behave irrationally and trade in a way that greatly impacts markets, the sophistication of the institutions involved in index trading and the issuance of CLNs makes this seem unlikely.

3.2 Summary of Near Month Imbalance and Returns

Table 2 shows summary statistics for our six futures contracts. We measure returns in percent, and express both volumes and imbalances as the number of contracts and as millions of dollars of futures notional. Oil futures contracts are for 1,000 barrels, and the average oil price over our sample was approximately \$100 per barrel. Gold futures contracts are for 100 troy ounces and the average gold price was a bit more than \$1,000 per ounce. Copper futures are for 25,000 pounds and the average copper price was about \$3 per pound. Accordingly, for oil, gold and copper, a single contract roughly corresponds to \$100 thousand notional value (gold notional value a bit higher and copper notional value a bit lower).

Corn and wheat futures contracts are for 5,000 bushels. The average price of corn was around \$5 per bushel, and wheat was just a bit higher, so one contract corresponds to approximately \$25 thousand of notional value. As the table shows, trade volumes are large and, trade volumes, imbalances, and returns are quite volatile over the period. Average one-minute volume ranges from approximately \$32 million of notional for WTI to approximately \$3.7 million of notional for Copper. Average imbalances are near zero, but they are quite volatile with standard deviations between \$10 and \$15 million per minute for gold, Brent, and the WTI, and between \$2 and \$7 million per minute for copper, corn and wheat.

3.3 Price Impacts and Volumes Across the Trading Day

To estimate the price impact of imbalance, we first estimate a univariate regression of futures returns (measured in percentage) on imbalance (measured in millions of dollars). The left column for each commodity in Table 3 shows the results of these regressions. For each commodity, we find that imbalance in futures markets has significant explanatory power for futures prices, suggesting that these markets play an important role in price discovery.[7] The R^2 range from 33% for WTI to 12% for Brent. The slope estimates provide our measure of impact, and range from 0.0022 for gold to 0.0143 for wheat. The interpretation is that a one million dollar buy (sell) will lead to a return in gold markets of positive (negative) 0.0022%, or 0.22 basis points. In contrast, in the smaller wheat market, a one million dollar trade will lead to a return impact of 1.43 basis points.

These full sample regressions reveal differences in impacts across commodities, but they obscure substantial variation within commodities across the trading day. These intraday patterns are important to understanding how a sophisticated investor might implement hedging positions. Since many financial products are benchmarked to the daily futures settlement price, it is likely that the hedging trades would take place near the settlement, which occurs at specific times for the various contracts.

To examine how trading impacts change through the day we estimate our univariate regression for each minute of the trading day (there are approximately 1,800 trading days in the sample, so each regression has approximately 1,800 observations, and approximately 1,500 for Brent since we exclude data prior to 2008). For WTI, Brent, gold, and copper we consider the interval from 7:30 a.m. through 4:00 p.m. Corn and wheat have extremely low volume after their close of floor trading at 2:15 p.m., so we end the analysis here. Corn and wheat also had their settlements delayed to 3:00 PM New York time for the 11-month period from 5/22/2012 to 4/5/2013, so we omit this period for the analysis in Figure 1, and for subsequent figures that present results across the trading day. We include these data when presenting analysis related to periods prior to the daily settlement, and when reporting full day regression results in the tables.

Figure 1 shows the results for these regressions, along with average volume, for each of the six commodities. The first panel shows the minute-by-minute average volume and trade impacts

[7] Evans and Lyons (2002) find a similar result in currency markets.

throughout the trading day for WTI futures. The volume rises on the open of pit trading at 9 AM, and then spikes at times of various announcements, including the EIA's weekly energy outlook published each Wednesday at 10:30 AM. The largest spike however occurs at 2:30 PM in New York when the daily futures settlement price is set.

The fall in price impact immediately before the WTI settlement suggests these trades have lower information content. The average impact throughout the day is relatively stable around 0.3 basis points per million dollars of imbalance, but this drops drastically in the minutes just around the settlement to roughly 0.12 basis points per million dollars of imbalance.

The implication of this finding is that even large trades during this period are unlikely to have a large impact on the market. For instance, a $10 million hedging trade (roughly the size of our average CLN), would only have an impact of 1.2 basis points if traded with a market order in the last minute before settlement. Note that a trade of this size would be less than one third of the standard deviation of imbalance for the settlement minute and less than 10% of the average settlement minute volume (see Table 2).

This pattern is repeated for each of the six commodities. For all of the commodities volume spikes and trade impact falls around the futures settlement, which occurs at 2:30 PM, 1:00 PM, 1:30 PM, and 2:15 PM New York time for Brent, copper, gold, and both corn and wheat respectively. The reduction in price impact is most notable for the WTI and gold, but is apparent in all six commodities. The high volume and volatility of imbalance at the settlement means that the impacts in these minutes are estimated with high levels of statistical accuracy. The right hand columns of Table 3 illustrates this. In unreported pooled regressions with dummy variables for the settlement minute, we confirm that in all cases, the settlement minute has significantly lower price impact than the full sample estimate.[8]

3.4 Inferring Price Impacts from Daily Data for Other Commodities

Roughly 20% of our CLNs are linked to commodities for which we do not have intraday data. In order to estimate the potential price impacts for these commodities, we regress our observed price

[8]The regressions shown in Table 3 and Figure 1 assume a linear impact of imbalance on returns. The Figure IA.2 in the Internet Appendix examines the settlement minute imbalance and returns for evidence of a nonlinear relation. We find little evidence that the relation is non-linear, but we see some evidence that very large imbalance leads to lower price impact than predicted by our linear estimates.

impacts on data that is available at the daily frequency and that we can obtain for a larger set of commodities. Intuitively we find that markets with lower volumes and higher volatilities have higher estimates of price impact. To formalize this intuition, we first calculate univariate regressions following the specification in Table 3 for each calendar year and commodity in our intraday data. We then regress these estimates on the average daily volume in millions of dollars across all futures maturities for a commodity market, as well as the daily volatility of returns to the near month contract. All variables are in logs. The specification is therefore

$$\text{Log}(Impact_{com,yr}) = \alpha + \beta_1 \text{Log}(AverageDollarVolume_{com,yr}) + \beta_2 \text{Log}(DailyVolatility_{com,yr}) \quad (1)$$

The first column of Panel A in Table 4 shows the regression using impact over all minutes in the calendar year as the dependent variable. The second column shows the same specification but with the dependent impact measured in the settlement minute. The third column shows the results from a pooled specification which adds a dummy variable for the settlement impacts. As the table shows, even with the relatively small sample, both volume and volatility are highly significant predictors of impacts with the expected signs. Moreover, the fit of the regression is extremely strong, with R^2 near 90%. Figure 2 illustrates the fit from the three regression specifications. As the figure shows, the regression performs extremely well in predicting impacts both across commodities and across years.

We then use the pooled regression to estimate impacts for each commodity-year for all of the contracts in our sample of CLNs. Panel B of Table 4 shows the averages for estimates across all years.[9] We find that LME copper, due to its high volume and relatively low volatility has the lowest estimated impact. Palladium contracts on the CME have the highest estimated impact. This high impact is partly a result of their relatively recent introduction in 2003. By the later portion of the sample the volume had risen and the estimated impacts had fallen substantially. While these estimates are likely imperfect, the strong fit shown in Figure 2 suggests that they should provide reasonable estimates for impacts in commodities where we do not have intraday data.

[9]See Table IA.3 in the internet appendix for a complete list of all commodity-year estimates.

4 Impacts of Financial Investors in Commodity Futures Markets

In this section we investigate the futures market impacts of two sources of financial investor flows: futures trades by indexers and issuances of CLNs. We calculate the change in net long positions of index funds in corn and wheat futures using data from the CFTC's supplementary positions of traders report. These data, provided for agricultural futures only, are generally considered the best measure of uninformed holdings in commodity futures.

To construct our sample of CLNs, we follow the procedure of HPW and collect and search the universe of 424b filings for issuers of CLNs from the SEC's Edgar website to identify CLNs linked to a single commodity with face value of at least \$2 million.[10] We find 594 notes, of which approximately 80% are in the commodities for which we have intraday data. Our sample of notes appears to closely track the set captured by HPW in terms of number and size.[11]

We also extend the analysis of HPW by explicitly calculating the hypothetical notional value of the hedge for each note, both at initial issue and on the determination date. We refer to the ratio of this notional value to the face value of the note as delta.

4.1 Calculating CLN Deltas

As mentioned in the introduction, many of the CLNs have complicated features (e.g. call provisions, caps, floors, knock-outs, and buffer regions). In order to accommodate the various structures, we calculate the initial delta via Monte Carlo valuation with 10,000 sample paths of daily returns over the life of the note.[12]

Figure 3 uses a representative note from the sample to illustrate the calculation of the pricing

[10] HPW also include notes linked to multi-commodity indices if all the indices are in the same sector (ie. energy). We collected these notes but do not include them because they complicate some portions of the analysis. None of these notes have more than \$10 million linked to a single commodity, and so for each of our analyses we report results for this subset of notes. HPW likewise report their results for this subset.

[11] We find some notes that were created and then transferred to a subsidiary for later sale to investors. We do not include these notes in our sample. We also do not include exchange-traded notes following HPW.

[12] We assume that returns are log-normally distributed with daily standard deviation equal to the realized daily standard deviation of the underlying over the month prior to issuance. The simulated risk-neutral drift of the underlying depends on the type of index used to calculate the note payoff. In some cases, the underlying is an excess return index, so the risk-neutral proportional expected return is zero. If the note uses a total return index, then the risk-neutral expected return equals the risk free rate. In many of the notes the underlying is a spot price, in which case we set the drift so that the expected value on the determination date is the futures price for the contract whose maturity is closest to the determination date. Roughly 10% of the notes have a final payoff calculated based on the average price over multiple trading days. We take this into account when calculating our deltas, so the determination date delta for these notes will be smaller.

date and determination date deltas. The figure illustrates the \$51,437,000 Capped Market Plus Notes linked to the S&P GSCI® Crude Oil Excess Return Index that were issued on January 24, 2011 by Barclays Bank. This note is typical in that it has no payments prior to maturity and the ending return on the note is a piece-wise linear function of the return on the underlying. The note also has a path-dependent "knock-out", which is another common feature in our sample. The note "priced" based on the closing value of the index on January 14, 2011 and the 424B form was filed with the SEC on January 19, 2011. When calculating the initial delta we follow HPW and assume that the full value of the notes was committed and hedged on January 14.

The notes matured on February 1, 2012. The determination date was January 25, 2012, when the final value of the index was observed and payoff of the note was set. If the notes were hedged, then the hedge should have been removed on the determination date, so we calculate the ending delta on that date. These notes have a knock-out buffer, a contingent minimum return, and a maximum return. A knock-out occurs if the index value falls below 80% of the pricing date value on any day over the life of the notes, and if a knock-out occurs then the contingent minimum of 8% is removed.

Panel A shows the actual return path for the index and three hypothetical return paths.[13] The hypothetical return paths are shown in part to illustrate the 10,000 simulated paths that are used to value the note and calculate the delta on the pricing date, and they are also used to illustrate the possible determination date deltas in Panel B. The initial note value is the average of the risk-neutral present values of the ultimate payments to the note along each simulated path based on the specific terms of the note, including interim interest payments and early calls.[14] The initial delta is then calculated by revaluing the note with a small change in the initial value of the underlying. The pricing date delta for this particular note is 0.89, so the size of the delta hedging trade would be 89% of the face value. Because most of the notes have concave payoffs with maximum slopes less than or equal to 1.0, the average (median) delta for the notes in our sample is only 0.61 (0.63). Accordingly, the face value of a note generally overstates the amount of the hypothetical initial hedge.

[13] For all of the notes, we obtain the values of the specific index or spot price to calculate the actual path. These data are obtained either via Bloomberg or the Commodity Research Bureau.

[14] The underlying index is an excess return, so the simulated risk-neutral paths used in the valuation of this note have zero expected return.

As shown in Panel A, the realized path for the underlying index was below the knock-out level during the life of the note, so as shown in Panel B, the return on the note matched the realized return on the index. The final return to the notes was -1% giving a final value of $50.9 million. The payoff function had a slope of 1.0 on the determination date, so the final delta is equal to the final value of $50.9 million divided by the initial issue amount (0.99). The hypothetical price path A for the underlying ends with the same return as the actual price path, but it never falls into the knock-out region so the ending return on the notes would have been the contingent minimum of 8%. The slope of the payoff is zero in this case, so if this had been the actual path, the ending delta would have been zero. The hypothetical price path B has an ending return for both the notes and the underlying of 15% which would have meant an ending value of $59.2 million for the notes, and since the slope of the payoff is 1.0 the size of the delta hedge would also be $59.2 million, or 115% of the face value, so the delta would be 1.15. The hypothetical price path C has a 40% return on the underlying, which would mean the note return would have been the maximum return of 30.5%. The payoff function has a zero slope at this point, so hypothetical path C would have resulted in an ending delta of zero. In our full sample, 342 of the 594 notes have a delta of zero on the determination date.

4.2 Summary Data for Changes in Index Positions and CLNs

Table 5 presents the summary data for the two sources of financial investment flows. Panel A shows the summary statistics for changes in the positions of index traders for corn and wheat. Both of these flows are quite substantial in magnitude. The standard deviation of flows are $225 million and $140 million for corn and wheat respectively. Interestingly, even though many index funds hold both commodities, the correlation of flows is not high, at only 0.23. To estimate the predicted price impacts from these flows, we apply our regression from column (3) of Panel A in Table 4. We obtain the estimated price impact for trades made at the settlement minute in the relative commodity for the given year, and then multiple this estimate times the change in the positions of traders. This yields a time series of predicted impacts. As the right hand side of Panel A in Table 5 shows, the standard deviation of this impacts is approximately 1% for both commodities.

Panel B of Table 5 shows summary statistics for the pricing dates of the CLNs. We follow HPW and combine notes in the same commodity with the same pricing date, and report face value, the size of the delta hedging trades, and the predicted price impacts of these trades. Here, in contrast to the changes in positions of index traders, the notes are very small relative to the size of the futures markets. The average size of the delta hedging trades is approximately $10 million. This yields an average predicted price impact of approximately 4 basis points if the note was traded near the futures settlement. This is the first indication that these hedging trades may not be responsible for the observed price impacts documented by HPW.

Figure 4 shows these results visually. The left hand panels of the figure show the weekly changes in index fund positions and the delta hedging trades associated with the pricing dates in millions of dollars. As the figure shows, the changes to corn and wheat index positions are large in magnitude throughout the sample. Panel E shows the issuance of CLNs. These issuances pick up in both size and frequency around 2008 and remain high through the end of the sample. The right hand side panels plot the price impacts associated with these flows on a common scale. As the figure shows, predicted returns associated with the hedging trades of CLNs are much smaller in magnitude than the changes in positions of index traders.

Finally, Panel C of Table 5 reports summary statistics for the CLN determination date. We start with the sample of notes with determination dates prior to 2019. The next line removes the many notes that have zero delta at the determination. This occurs either because the note is called or because the underlying commodity price is in a region in which the note has no exposure. Then we remove the smaller notes to focus on the larger notes where HPW find their significant determination date results. Finally we restrict the sample to prior to February 2014 to obtain the notes that were available at the time of the original HPW study. While the predicted price impacts are again small, they are larger in magnitude than the pricing impacts as the deltas tend to be higher for those notes which still have exposure to the underlying at the determination date.

4.3 The Impact of Commodity-Index Traders on Futures Markets

Here we explore the impacts of commodity-index traders on futures markets using our intraday data. Since we cannot directly identify these traders in our high-frequency data, we instead investigate whether or not we can associate changes in index trader positions with changes in aggregate order flow, and whether or not this order flow is concentrated at any point of the day. We also investigate whether these changes in positions are associated with returns in the futures market. To this end, we estimate regressions of the form

$$FuturesImbalance_t = \alpha + \beta \Delta IndexPositions_t \tag{2}$$

$$FuturesReturn_t = \alpha + \beta \overline{\Delta IndexPositions_t} \tag{3}$$

When performing regressions of imbalance, we regress weekly imbalance on the total change in index trader positions, where both are measured in contracts. Therefore, the slope coefficient can be interpreted as the percentage of the change in index trader position reflected in abnormal trade imbalance. For the return regressions, we standardize the index flow so it has a standard deviation of one. Thus, in the return regressions the slope can be interpreted as the weekly return impact of a one standard deviation change in index trade positions. The overbar denotes this standardized variable.

The index trader positions are available weekly, so we sum the dependent variable across the trading days in a week to create each observation. To the extent that index contracts are tied to daily settlement prices, we might expect the impacts of changes in index positions to be concentrated near the daily settlement. Accordingly, we also estimate our regressions using later portions of the trading day (aggregated across days in the week) as our dependent variable. Table 6 shows the results.

Columns (1) and (4) of both panels in Table 6 show the results for total returns and imbalance summed across the week. The coefficients in column (1) of 0.371 for corn and 0.514 for wheat are strongly significant, and indicate that for a given weekly change in index positions we see corresponding weekly imbalances in the same direction that average 37.1% and 51.4% of the changes in index positions. While it may seem puzzling that our estimates are lower than 100% this not surprising. We would expect these traders to utilize both limit orders and floor orders to execute these trades, and if they do so it will not be reflected in our measure of imbalance.

Column (4) in both panels shows that the responses of futures returns across the full day to changes in index positions are not statistically significant at the 5% level. The point estimates are positive for both corn and wheat (0.62 and 0.49) with a magnitude of approximately one half of the predicted impact of 1%. This is what one might expect if the traders are able to execute their orders with some sophistication. However, consistent with the findings of Stoll and Whaley (2010) and Irwin and Sanders (2012), a return of this magnitude is not large enough to discern from the noise of daily price changes.

When we focus the analysis further and examine the period around the futures settlement, we see a striking result. Column (2) shows that in the 30 minutes prior to settlement we see imbalance equal to approximately 14% and 24% of the total change in index positions for corn and wheat respectively. Column (5) shows that these imbalances close to settlement are translating into a return impact. A one standard deviation increase in index traders' positions is translating into a 28.1 basis point price increase across the week over these minutes for corn, and a 48.1 basis point increase for wheat. Columns (3) and (6) show that a large portion of this impact is concentrated in the single minute prior to settlement. All of the results near the settlement have strong statistical significance (t-stats range from 3.5 to 5.6). This significance is not a result of larger point estimates, as these estimates are lower than the full day estimates, but instead is a result of the increased power from focusing on the period of the day when index traders are most likely to be trading.

To visualize these patterns, Panel A of Figure 5 shows slope estimates where the dependent variable is the cumulative return up to each minute in the trading day. For example, the 12:00 PM point on the plot shows the estimated slope and 95% confidence interval for a regression where the dependent variable is the cumulative return (including the overnight return from the previous days settlement at 2:15 PM) through 12:00 PM, summed across the days of the week. The Figure shows that returns and imbalances associated with changes in the positions of index fund traders increase slowly across the day, and then spike just before the settlement.[15]

Figure 6 focuses on the 30 minutes prior to the daily settlement. Again the plots show regression coefficients for expanding windows. For example, looking at the the 15-minute point on the plot, the dependent variable is the cumulative return from 30 minutes before to 15 minutes before the daily settle, summed across the trading days in the week. Here we see a much stronger statistical relation for both imbalance and returns, and again the large spike is evident at the closing minute.

These results suggest that index traders are taking positions just prior to the close. This potentially allows them to reduce tracking error if the fund is targeting daily changes in price,

[15] In the figure the full day regression result for returns is significant at the 95% level. This discrepancy with the full day estimates in Table 6 arises due to the fact that the figure excludes the period of time when settlement was delayed until 3:00 PM.

and also reduces the impact of trades. Although these results show price impacts of these trades near the settlement, it is difficult to tell from our data whether and to what degree these impacts are permanent. If the impacts were reversed by the next morning, then that might suggest that overnight returns would be negatively related to the average weekly index flow. The left most portions of Panels B and D of Figure 5 show that overnight returns are essentially unrelated to the weekly index flows, but the wide confidence intervals make it difficult to draw conclusions.[16]

To understand the economic magnitude of the return impacts prior to the close, one can look first at the R-squared values in the return regressions of Table 6. The R-squared in column (5) shows that this return impact explains roughly 6% and 10% of the price variation in the 30 minutes prior to close for corn and wheat respectively. Although the coefficients are similar in column (4), the R-squared falls to 1-2% when considering the full week's return, suggesting that index funds are not contributing a large portion of the weekly variance in futures prices.

Despite the fact that these daily impacts do not contribute significantly to the overall variance of prices, it is possible that cumulatively they could add up to larger distortions in the level of price. To illustrate this, Figure 7 plots cumulative changes in the positions of index traders and estimated impacts. For this analysis, we use the return impact coefficient from the 30 minutes prior to settlement (Table 6) as our measure of price impact for index traders, and assume that there is no reversal. This should therefore be viewed as an upper bound on the overall impact.

Panels A and B of Figure 7 show the positions of index traders in corn and wheat respectively over our sample. There are some large changes over the period. For both corn and wheat the positions fall by roughly 40% over 2008, while full rebounding to above previous levels in 2010. As shown in Panels C and D, these large changes in positions, when multiplied by our impact estimates, would lead to price impacts of roughly 6% for corn and 8% for wheat. Panel E and Panel F plot observed prices of corn and wheat, and the but-for price in the absence of the observed impacts. Note that this is not intended to be a true measure of a "fundamental" price, as we do not include changes in prices prior to 2007 due to the fact we do not have the data to estimate price impacts over this period. Instead, this is to illustrate that these changes, while potentially economically meaningful in level, are again small compared to the overall volatility in corn and wheat.

[16] In unreported results we directly test for reversal but find similarly inconclusive results.

4.4 Pricing Date Returns for Commodity-Linked Notes

HPW find that days with the creation of CLNs have significant positive average returns. They attribute this to the price impact of hedging trades made in the futures market. Because the exposure to the underlying commodity starts at the daily settlement on the pricing date, that is where we would expect to see the hedging trades and their associated price impact.[17]

When conducting their analysis, HPW exclude notes pricing during the Goldman Roll documented by Mou (2010). This period includes the 5th to 9th trading days of the month and the five previous business days. They do this to avoid potential returns coming from the price pressure associated with the roll trades. However, it is not clear that this is the correct choice. In particular, as documented by both Mou (2010) and Neuhierl and Thompson (2016), the predictable returns associated with the roll trades no longer occur after 2003, which is the period during which we see the issuance of CLNs. Moreover, we find no difference in the price impacts of order flow in the Goldman Roll period.[18] Accordingly, we would not expect to see differences in the price impacts of hedging trades for notes issued during this period.

Although HPW do not mention it, there are substantial differences in the frequency of CLN pricing days across the trading month. Panel A of Figure 8 plots the number of notes with pricing dates on each of the ten trading days at both the beginning and end of the month, with a single bar representing the total notes issued in the middle day of the month (months have from 20-22 trading days, so for some months there are no days in the middle and for others there are one or two). Panel B repeats this figure for notes with at least \$10 million of face value. As the figure shows, issuance, particularly of large notes, is much more common the five days prior to the end of the trading month than on other days.[19] We refer to this five day period as the "Active Issuance Period". This increased level of issuance activity implies a change in the demand or supply of large notes in this period. Perhaps this is due to performance targets for sellers of the notes that lead them to market more aggressively to retail clients, or perhaps there are institutional buyers that

[17] HPW also report results for returns on days after the pricing date and results using abnormal returns controlling for various systematic variables. Since their results are similar using raw returns, and generally the most significant on the actual pricing date, we focus on these specifications for parsimony.

[18] See Table IA.2 in the Internet Appendix.

[19] Unreported logit regressions show that days in this period are more than three times more likely to have an issuance of a note with \$10+ million of face value, and this difference in issuance frequency is highly statistically significant ($p < 0.001$).

have reasons to buy during this period. This period is entirely outside of the Goldman Roll and these notes are therefore included in the main analysis of HPW. For our analysis below, we report our findings for the different portions of the trading month. As we will show, it is only the notes in the Active Issuance Period that have positive average pricing date returns.

We now examine daily futures returns on the pricing dates of the notes. Table 7 shows the realized pricing date returns and the predicted impacts for different subsets of notes and different portions of the trading month. Panel A includes all notes. In column (1) we show the average pricing date return for all notes, and find only a marginal positively significant average return. Column (2) of Panel A replicates the finding of HPW for all notes outside of the Goldman Roll period. We find a nearly identical average return of 28 basis points on these days. As column (3) shows, this increase is due to the fact that the notes issued during the Goldman Roll have a slightly negative return. When we cut the sample down further and only include notes in the Active Issuance Period, we see that these are the notes that drive the entire positive pricing date result, as these 201 days (of 532 total) have a positive average return of 44 basis points, while the remaining 331 days have an average return of negative 5 basis points. For all of the subsets with significantly positive returns, the returns are much higher than the predicted impact given the size of the delta hedges.

Panel B repeats Panel A but restricting the sample to days with \$10+ million of face value. The findings are qualitatively the same, but the positive returns are even stronger. The notes issued during the Active Issuance Period account for the entire result, and have a positive return on average of 66 basis points (t-stat = [5.03]), while the predicted impact of their hedging trades is only 9 basis points. The average return is thus 7 times larger than would be expected if the delta hedges were naively executed in futures markets. Panels C and D repeat the analysis of Panels A and B for the approximately 80% of the notes for which we have intraday data. Here we find nearly identical patterns in average returns. We note that these commodities have large futures markets, so the predicted impacts are less than 10% of the observed pricing date returns both when excluding the Goldman Roll and when restricting to the Active Issuance Period.[20] Panel E shows

[20]HPW also measure of the face value of the note relative to the open interest of the two nearest-month futures and find that larger notes have larger returns. We perform a similar test, but regress pricing day returns on our measure of predicted impact from the hedging trade. We find some evidence that notes with a larger predicted impact have higher pricing day returns. However, the slope of the regression is near one, so the small value of the average predicted returns means that this can explain only a small portion of the total pricing date average return. See Section IA.6 in

similar analysis, but uses the full set of dates in each subset from column (1) in Panels A - D and regresses the return on two dummy variables corresponding to the note being in the Active Issuance Period or outside of the Goldman Roll. In all cases the Active Issuance Period Dummy is highly significant, and both the Non-Goldman Roll dummy and the intercept are insignificant at the 5% level. This again illustrates that the positive returns are only associated with the notes that price in the Active Issuance Period.

Our next set of tests focuses on the notes for which we have intraday data, and looks within the trading day to see if we observe any patterns similar to those we see in the analysis of commodity-index traders.[21] We focus on days greater with $10+ million of face value issued during the Active Issuance Period since this is where we see the largest daily effects. Table 8 and Figure 9 show the results. As both the Figure and the Table show, we see no focused pattern of purchases or positive returns around the futures settlement. In fact, looking at Panel A of the figure we see no evidence of any abnormal imbalance during the pricing date of the note. Instead, approximately half the of the return, as well as what positive imbalance there is, appears to occur prior to 9 AM on the day of the pricing. Nearly all of the return effect has accumulated by 12 PM. While there are reasons that a hedging trade may be delayed, there seems no reason a hedging trade would occur prior to the pricing of the note, which typically occurs at the futures settlement. One possible explanation is that many of the gold notes price on the London PM fix, which occurs at 10 AM or 11 AM in New York depending on the time of year, and most of the copper notes price on the close of the LME which is at 12 PM or 1 PM in New York. We therefore examine the returns and imbalance in a two hour window starting 30 minutes before the pricing of the note, and again find no effect.

Our findings on the pricing date can be summed up as follows:

1. The pricing date returns are between 6 and 10 times too large to be explained by the size of the delta hedging trades.
2. The average positive returns are only present in notes that price in the 5 days prior to the end of the month, a period during which CLN issuance frequency greatly increases.

the Internet Appendix.

[21] Many of our gold notes are linked to the London PM fix and nearly all of our copper notes are linked to the LME spot price. We do not have intraday data for these prices. However, using our intraday data we confirm that the futures prices are highly correlated with the London markets. We therefore use our intraday data for copper and gold when examining notes linked to these commodities.

3. There is no return or imbalance in the minutes near the futures settle or the pricing of the notes, and instead most of the positive return occurs overnight or early in the trading day.

Together, these results suggest that demand or supply of notes is responding to the changing price of the underlying commodity. HPW acknowledge this potential bias, and address this by examining returns on the determination dates when final payoff of the CLN is set. If returns are negative on this date, it would be strong evidence of price impact from the unwinding of hedging trades, and this evidence would not be subject to the concerns of selection bias or endogeneity inherent with the pricing date results.

4.5 Determination Date Returns for Commodity-Linked Notes

Table 9 shows our analysis of futures returns on days with CLN determination. Following HPW, we restrict our analysis to notes which still have positive exposure so the underlying commodity on the determination date. The table shows average determination date returns for various subsets of the sample. In particular, column (2) of Panel D is our attempt to replicate the main finding of HPW. This is the return on determination dates of notes that still have a positive delta on the day of determination, are outside of the Goldman Roll, have at least $10 million of face value, and have a maturity prior to February of 2014. As in HPW we combine notes with the same underlying commodity and the same determination date, but in spite of attempting to match their approach exactly, our sample size is larger (we have 50 observations and they have 42). More notably, while they report a significant average return of -42 basis points (t-stat of 2.50), we find an insignificant average return of only -9 basis points (t-stat of 0.45).[22] Looking at the table, we do not find a negative average return that is significant at the 5% level for any subset of the determination dates, either including the Goldman Roll or extending the sample of determination dates through 2018.

It is also interesting to note that the larger deltas of the notes on the determination dates imply that the predicted impacts are considerably larger than those on the pricing dates. The fact that we do not find negative returns may reflect an ability of sophisticated traders to make large trades

[22] HPW do not indicate how they handle the roughly 10% of notes whose final value depends on the average commodity price over multiple days. When calculating the average return we only use the return on the final determination date. There are very few of these notes with a positive determination date delta, and in particular there is only one note with greater than $10 million of face value prior to 02/2014. Removing or including this note has no significant impact on any of our results.

without impacting futures markets, but it also could reflect a lack of futures hedging in some cases. Particularly with larger notes, it may be the case that the CLN issuance is triggered when an institutional client enters an over-the-counter contract with the investment bank, who then sells the CLNs to hedge their exposure to that contract.

5 Conclusion

In this paper we construct trade imbalances for six major commodity futures markets. We find that order flows in these futures markets play a large role in price discovery. We also document substantial intraday variation in price impacts, with high volumes and low price impact around futures settlements. We use our findings on trade impacts to examine the potential impacts of financial investors in this market. We examine the impact from changes in the positions of commodity-index investors for corn and wheat futures from the CFTC. We find strong evidence for trade imbalances and price impacts associated with these positions, concentrated in the minute prior to the daily futures settlement.

We also find that the positive returns associated with the issuance commodity-linked notes documented by Henderson et al. (2015) are surprisingly large given the notes' size, occur primarily early in the trading day, are not associated with abnormal trade imbalance near the futures settlement or pricing of the note, and are only present in notes issued near the end of the month when issuance frequency is substantially higher. We also find no evidence of significant negative returns on CLN determination dates. These findings suggest that the positive returns are potentially the result of CLN issuers or purchasers favoring days with increasing commodity prices, rather than evidence of impacts from associated hedging trades.

References

Acharya, V. V., L. A. Lochstoer, and T. Ramadorai. 2013. Limits to arbitrage and hedging: Evidence from commodity markets. *Journal of Financial Economics* 109:441–465.

Admati, A. R., and P. Pfleiderer. 1988. A theory of intraday patterns: Volume and price variability. *The Review of Financial Studies* 1:3–40.

Alquist, R., and O. Gervais. 2013. The role of financial speculation in driving the price of crude oil. *The Energy Journal* pp. 35–54.

Baker, S. D. 2014. The financialization of storable commodities. *Finance Down Under 2014 Building on the Best from the Cellars of Finance* .

Basak, S., and A. Pavlova. 2016. A model of financialization of commodities. *The Journal of Finance* 71:1511–1556.

Bessembinder, H. 2015. Predictable ETF Order Flow and Market Quality. *Available at SSRN 2506322* .

Bessembinder, H., A. Carrion, L. Tuttle, and K. Venkataraman. 2016. Liquidity, resiliency and market quality around predictable trades: Theory and evidence. *Journal of Financial Economics* 121:142–166.

Buyuksahin, B., and M. A. Robe. 2011. Does paper oil matter? Energy markets' financialization and equity-commodity co-movements. *Available at SSRN 1855264* .

Chari, V. V., and L. Christiano. 2017. Financialization in commodity markets. Tech. rep., National Bureau of Economic Research.

Cheng, I.-H., A. Kirilenko, and W. Xiong. 2014. Convective risk flows in commodity futures markets. *Review of Finance* 19:1733–1781.

Cheng, I.-H., and W. Xiong. 2014. Financialization of commodity markets. *Annu. Rev. Financ. Econ.* 6:419–441.

Elder, J., H. Miao, and S. Ramchander. 2014. Price discovery in crude oil futures. *Energy Economics* 46:S18–S27.

Evans, M. D., and R. K. Lyons. 2002. Order flow and exchange rate dynamics. *Journal of political economy* 110:170–180.

Fattouh, B., L. Kilian, and L. Mahadeva. 2013. The role of speculation in oil markets: What have we learned so far? *The Energy Journal* pp. 7–33.

Goldstein, I., and L. Yang. 2017. Commodity financialization and information transmission. In *AFA 2016 Annual Meeting*.

Halova, M. W., A. Kurov, and O. Kucher. 2014. Noisy inventory announcements and energy prices. *Journal of Futures Markets* 34:911–933.

Hamilton, J. D., and J. C. Wu. 2014. Risk premia in crude oil futures prices. *Journal of International Money and Finance* 42:9–37.

Hamilton, J. D., and J. C. Wu. 2015. Effects of index-fund investing on commodity futures prices. *International economic review* 56:187–205.

Hasbrouck, J. 1991. The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies* 4:571–595.

Henderson, B. J., N. D. Pearson, and L. Wang. 2015. New evidence on the financialization of commodity markets. *The Review of Financial Studies* 28:1285–1311.

Irwin, S. H., and D. R. Sanders. 2010. The impact of index and swap funds on commodity futures markets .

Irwin, S. H., and D. R. Sanders. 2011. Index funds, financialization, and commodity futures markets. *Applied Economic Perspectives and Policy* 33:1–31.

Irwin, S. H., and D. R. Sanders. 2012. Testing the Masters Hypothesis in commodity futures markets. *Energy economics* 34:256–269.

Marshall, B. R., N. H. Nguyen, and N. Visaltanachoti. 2011. Commodity liquidity measurement and transaction costs. *The Review of Financial Studies* 25:599–638.

Mou, Y. 2010. Limits to arbitrage and commodity index investment: front-running the Goldman roll. *Available at SSRN 1716841* .

Neuhierl, A., and A. Thompson. 2016. Liquidity Timing in Commodity Markets and the Impact of Financialization. *Available at SSRN 2682698* .

O'Hara, M. 2015. High frequency market microstructure. *Journal of Financial Economics* 116:257–270.

Raman, V., M. A. Robe, and P. K. Yadav. 2017. The Third Dimension of Financialization: Electronification, Intraday Institutional Trading, and Commodity Market Quality. *WBS Finance Group Research Paper* .

Silvennoinen, A., and S. Thorp. 2013. Financialization, crisis and commodity correlation dynamics. *Journal of International Financial Markets, Institutions and Money* 24:42–65.

Singleton, K. J. 2013. Investor flows and the 2008 boom/bust in oil prices. *Management Science* 60:300–318.

Sockin, M., and W. Xiong. 2015. Informational frictions and commodity markets. *The Journal of Finance* 70:2063–2098.

Stoll, H. R., and R. E. Whaley. 2010. Commodity index investing and commodity futures prices. *Journal of Applied Finance (Formerly Financial Practice and Education)* 20.

Tang, K., and W. Xiong. 2012. Index investment and the financialization of commodities. *Financial Analysts Journal* 68:54–74.

Yan, L., S. H. Irwin, and D. R. Sanders. 2019. Is the Supply Curve for Commodity Futures Contracts Upward Sloping? *Available at SSRN 3360787* .

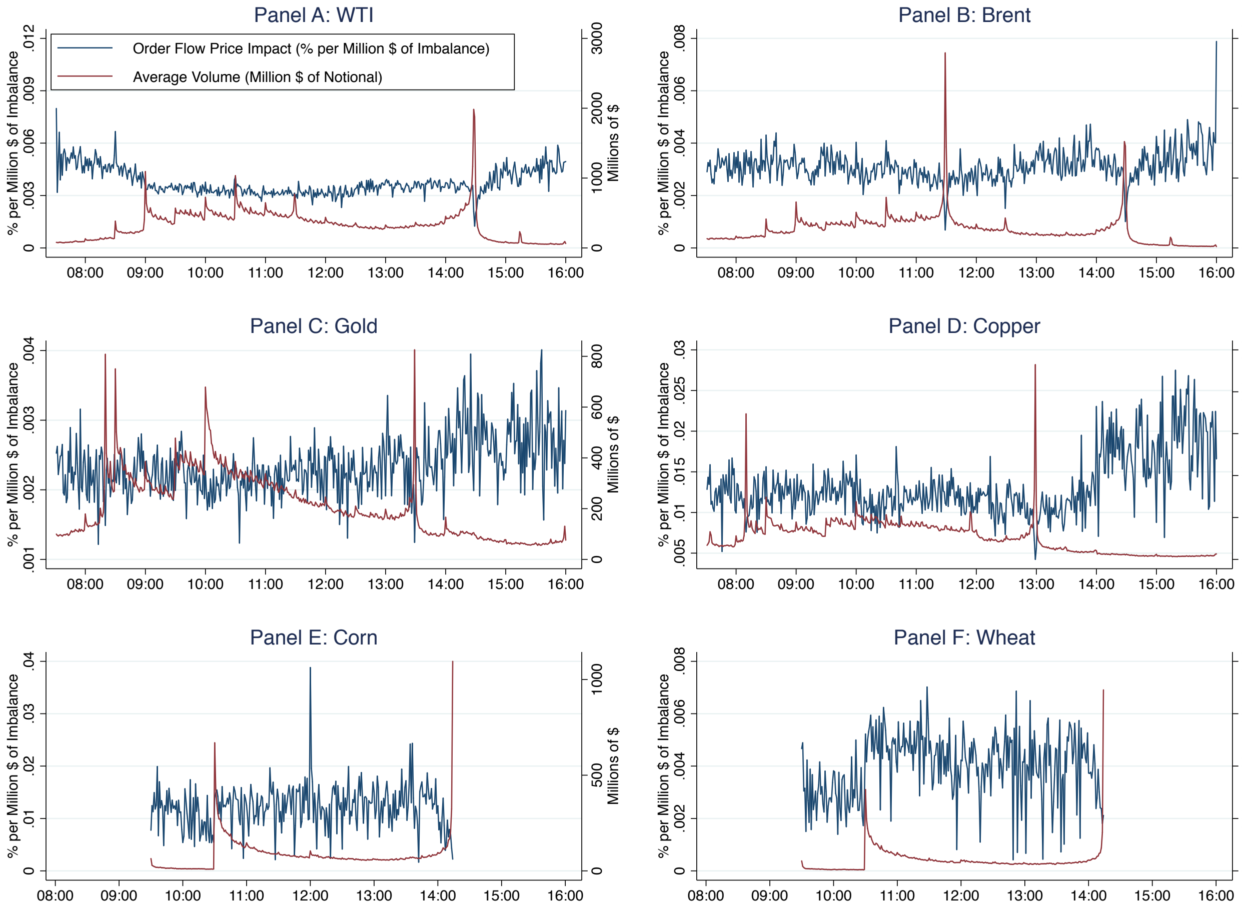


Figure 1. Volume and Price Impact of Order Flow Across the Trading Day
The figure shows the average intraday volume (in red) by minute for each commodity as well as the minute-by-minute trade impact (in blue). The trade impact is measured as the slope in a univariate regression of return (%) on trade imbalance (millions of $) estimated using imbalance and returns in each minute of the day. For instance, for the 12:00 average volume we calculate the total volume from 12:00:00 to 12:00:59 for each day, and take the average of this value across all trading days. Similarly, to calculate the 12:00 imbalance, we calculate the total return and imbalance from 12:00:00 to 12:00:59 for each day, and then run a univariate regression of return on imbalance for this minute across all trading days. The sample is 1/1/2007 to 4/1/2014. We exclude data prior to 1/1/2008 for Brent, and we exclude the period for Corn and Wheat in which the future settlement was delayed until 15:00 EST (5/22/2012 to 4/5/2013).

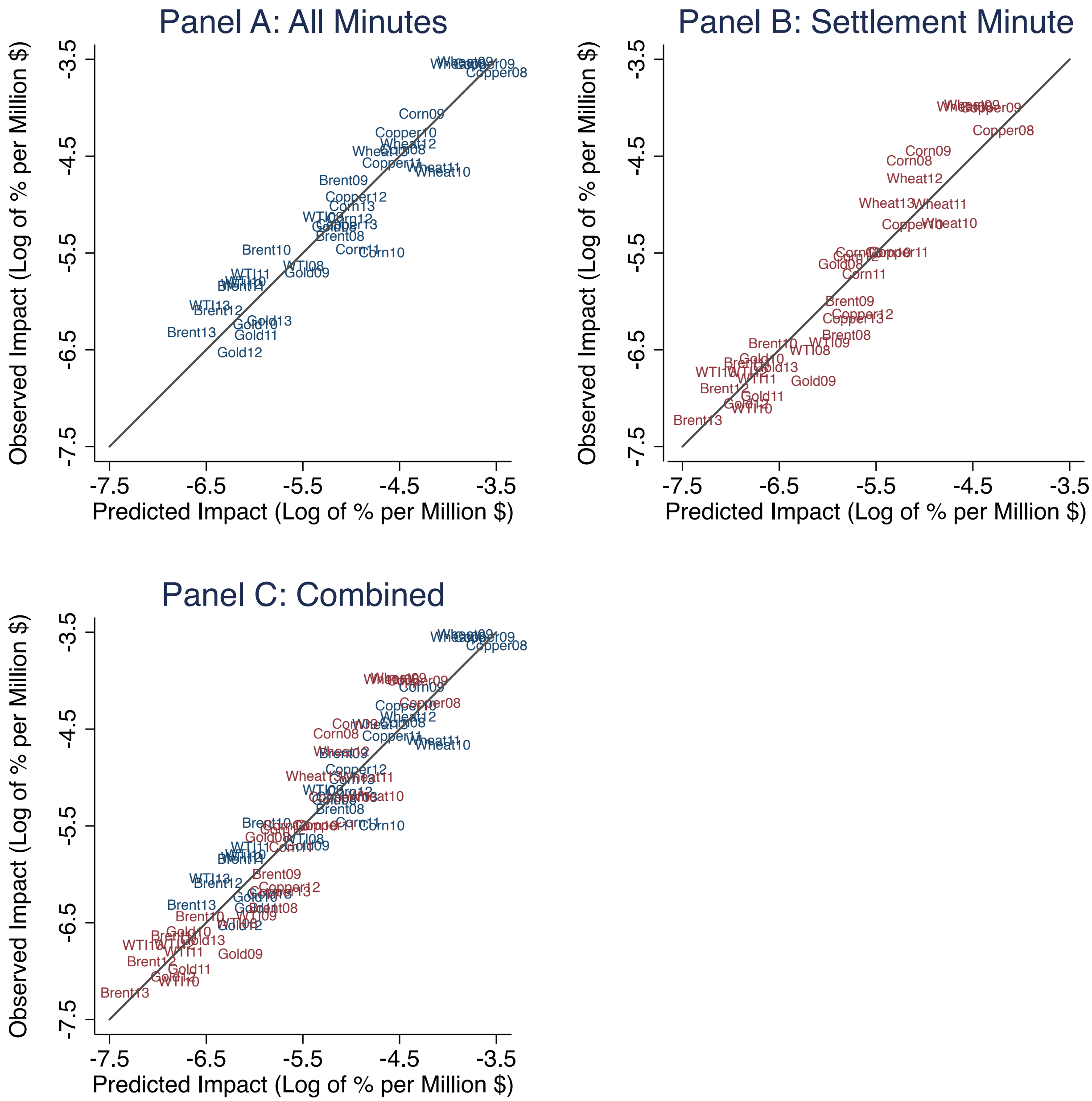


Figure 2. Regression fit for inferring order flow price impact from daily data
This figure plots the fit for regressions of the log of predicted price impacts for each commodity in a calendar year on the logs of average daily futures volume and daily futures volatility (See columns (1) - (3) of Panel A in Table 4).

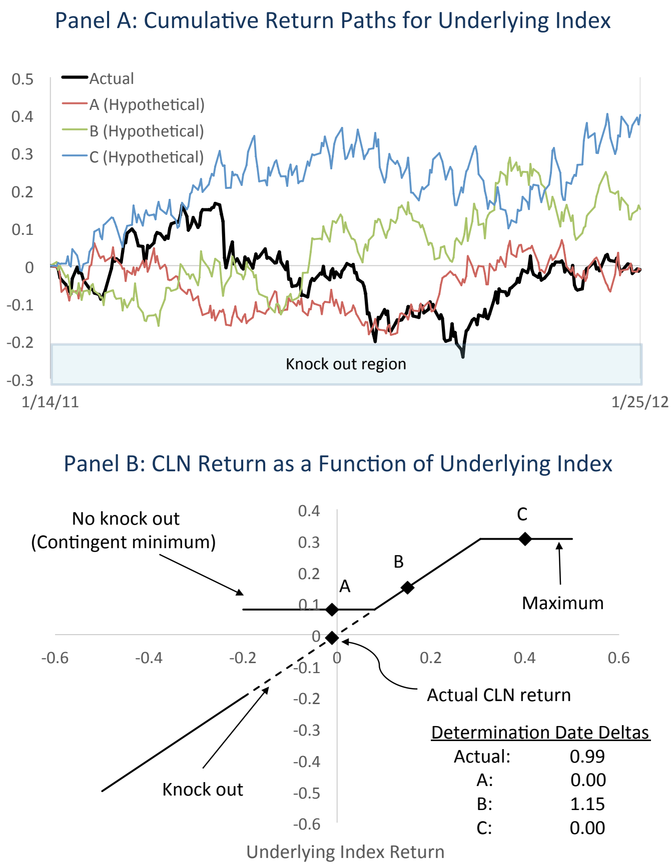


Figure 3. Return Paths and Determination Date Deltas for a Sample CLN
The figure illustrates the \$51,437,000 Capped Market Plus Notes linked to the S&P GSCI® Crude Oil Excess Return Index. These notes have a knock-out buffer, a contingent minimum return, and a maximum return. A knock-out occurs if the index value falls below 80% of the pricing date value on any day over the life of the notes, and if a knock-out occurs then the contingent minimum of 8% is removed. Panel A shows the actual return path for the index and three hypothetical return paths. Panel B shows the piecewise linear payoff structure across the ending cumulative returns of the underlying index along with the determination date delta for each path. The delta is calculated as (Ending Note Value × Slope of Payoff on Determination Date) / (Face Value of the Note). See section 4.1 for a detailed explanation of the four determination date delta values.

Figure 4. Sources of Order Flow from Financial Investors and Predicted Price Impacts The figure shows plots of sources of order flows and the predicted price impact from financial investors in commodity markets. Panel A and show weekly changes in position of commodity-index traders for corn and Panel B shows the predicted price impact from these trades using the estimates calculated using the regression specification in Table 4. Panel C and D repeats this analysis for wheat. Panel E shows the size of the delta hedging trades associated with CLN pricing and Panel F shows the predicted price impact of these trades.

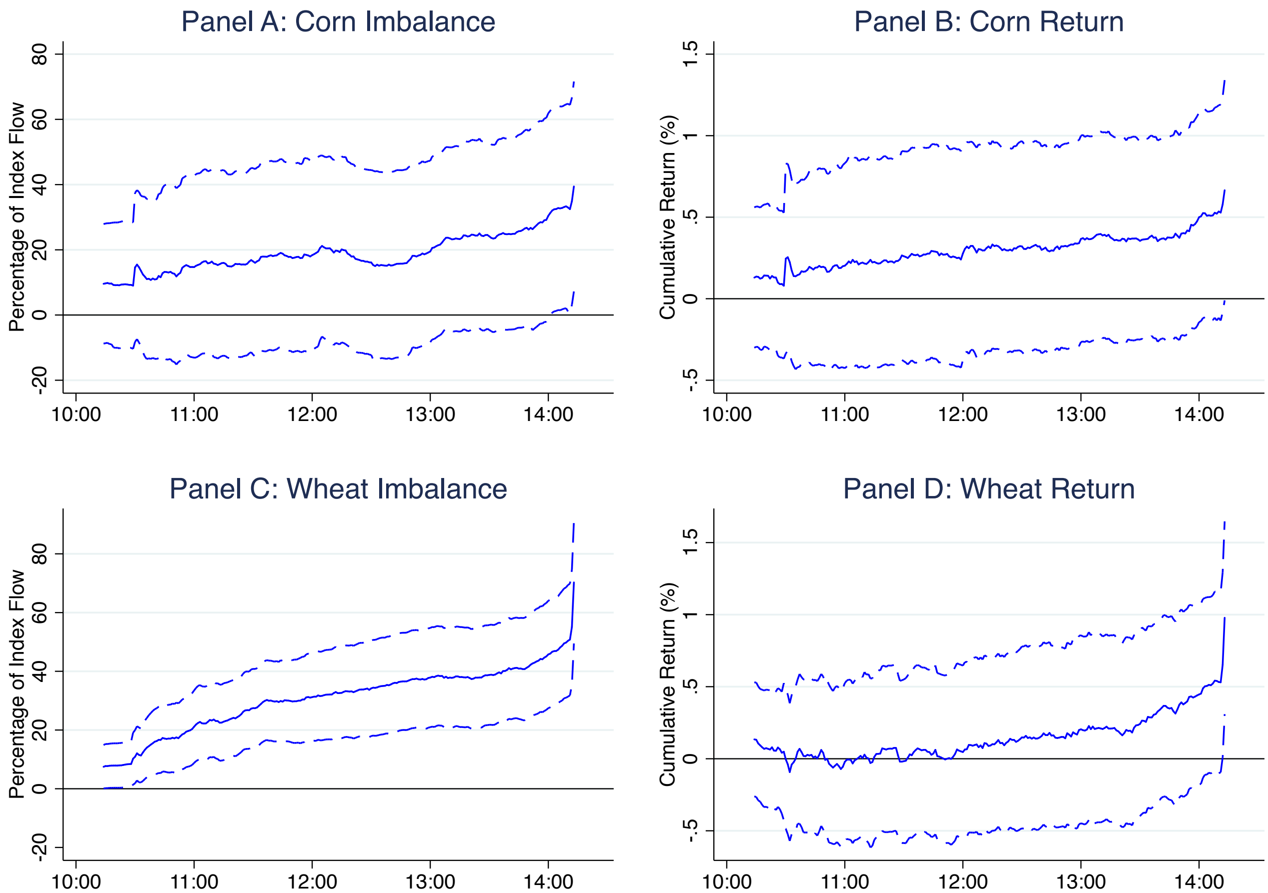


Figure 5. Intraday Impact of Changes in the Positions of Commodity-Index Traders: Full Day

The figure plots the slope coefficient and 95% confidence interval from regressions where the dependent variables are cumulative futures imbalances and returns for expanding windows across the trading day, and the independent variables are weekly changes in the positions of index traders for corn and wheat. In Panels A and C the independent and dependent variable are measured in number of contracts. In Panels B and D the dependent variable is returns in percent and the independent variable (index flows) is standardized to have a standard deviation of one. For each minute, the dependent variable is the cumulative return or imbalance measured from the previous day's settlement summed across the trading days in the week. Data are 1/1/2007 to 4/1/2014.

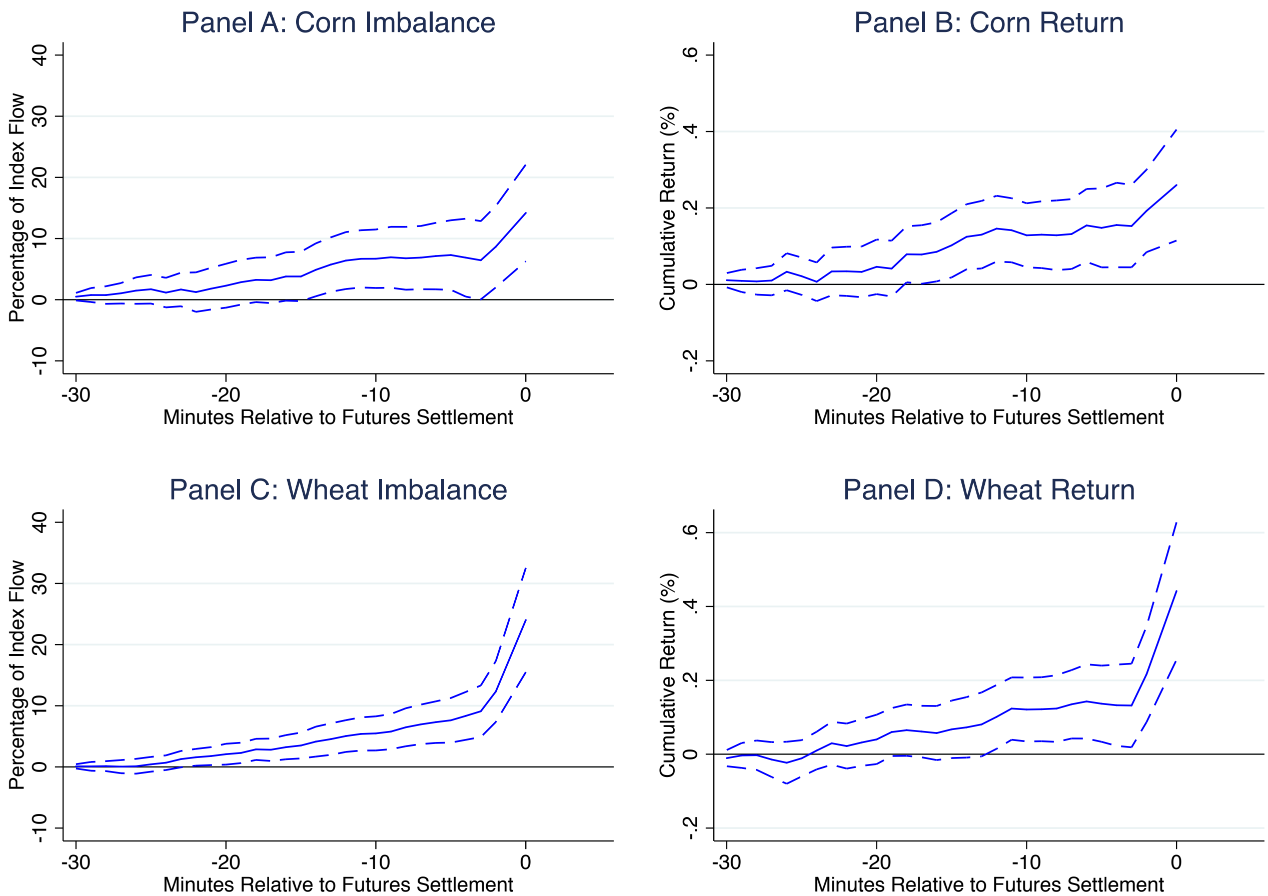


Figure 6. Intraday Impact of Changes in the Positions of Commodity-Index Traders: 30 Minutes Prior to Settlement

The figure plots the slope coefficient and 95% confidence interval from regressions where the dependent variables are cumulative futures imbalances and returns for expanding windows across the 30 minutes prior to futures settlement and the independent variables are weekly changes in the positions of index traders for corn and wheat. In Panels A and C the independent and dependent variable are measured in number of contracts. In Panels B and D the dependent variable is returns in percent and the independent variable (index flows) is standardized to have a standard deviation of one. For each minute, the dependent variable is the cumulative return or imbalance measured from 30 minutes prior to settlement summed across the trading days in the week. Data are 1/1/2007 to 4/1/2014.

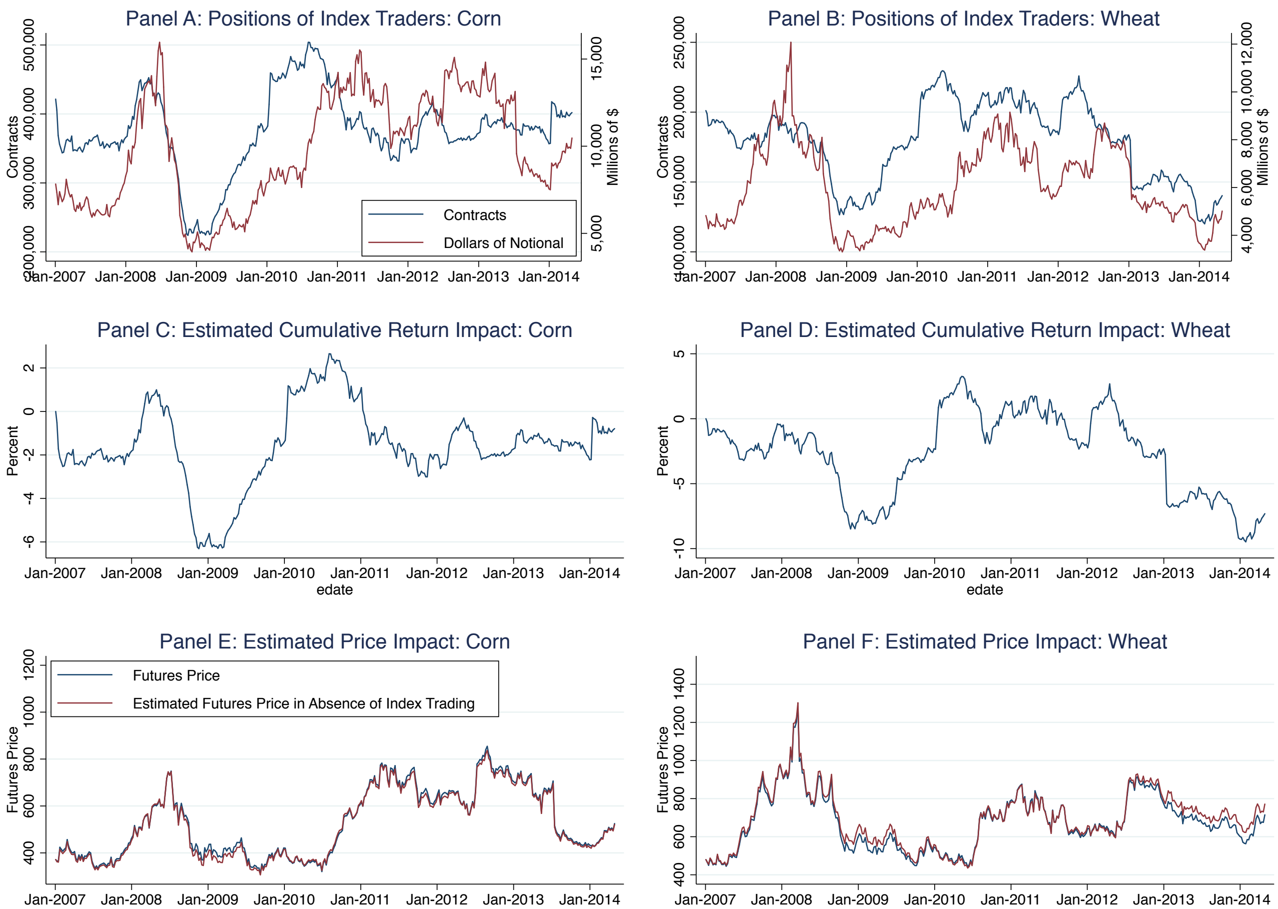


Figure 7. Estimated Cumulative Impact of Index Trader Flows
The figure shows the positions of index traders from the CFTC in corn and wheat along with the estimated price impacts of changes in these positions in using our estimates of price impact in the 30 minutes prior to futures settlement (Table 6 column (5)). Panels A and B show the positions of index traders in futures contracts and millions of dollars. Panels C and D show the cumulative sums of weekly impacts, which are calculated by multiplying each week's standardized change in index trader positions by the estimate of impact from Table 6. Panels E and F show the observed futures price and the futures price adjusting the cumulative return impact. This adjustment is done by multiplying the observed futures price by $(1 + CumulativeReturnImpact)$, where the $CumulativeReturnImpact$ is shown in panels C and D.

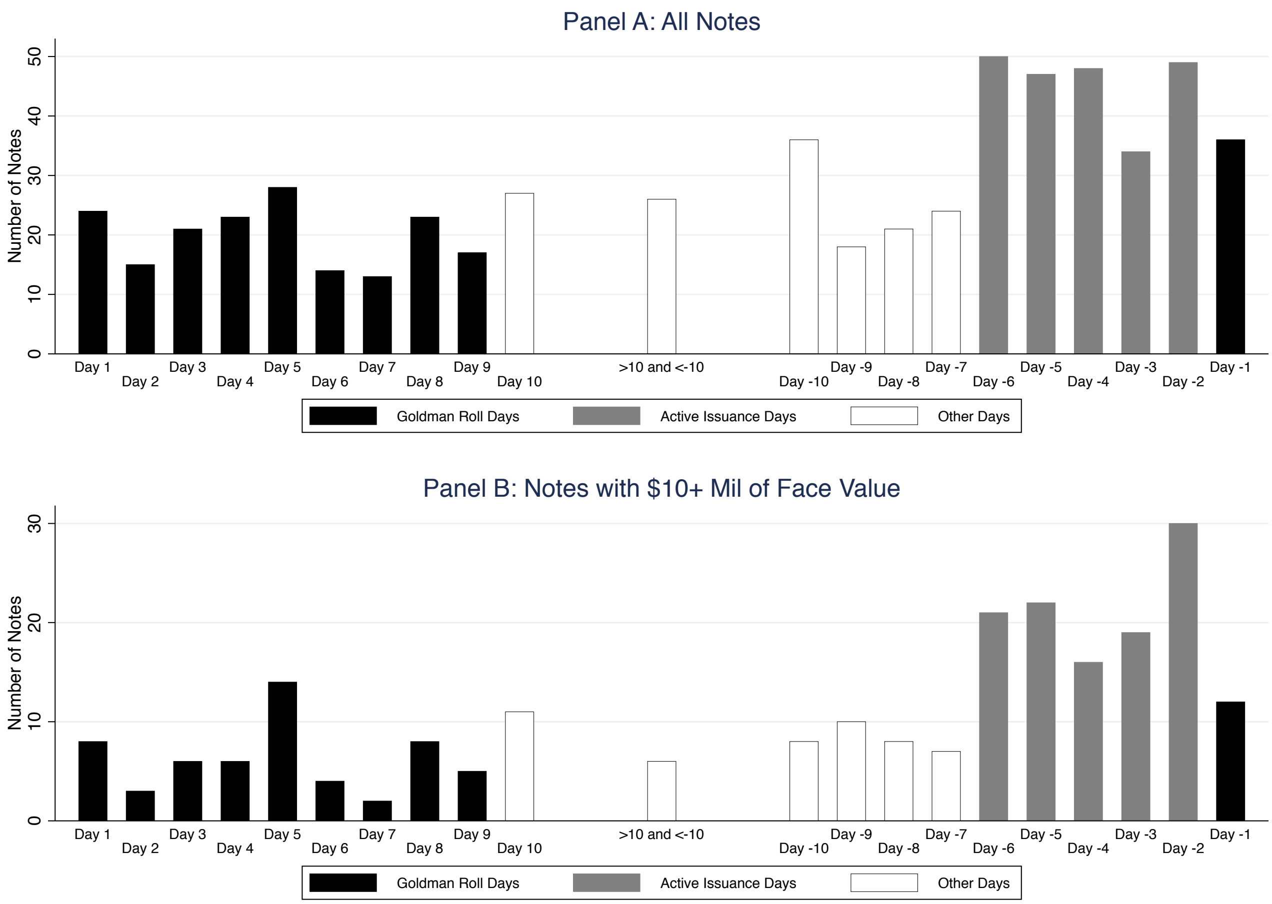


Figure 8. CLN Pricing Days Across the Trading Month

This figure shows the pricing frequency of CLNs across the trading month. The left hand bars represent the first 10 trading days of the the calendar month. The right hand bars represent the last 10 trading days of the calendar month. The central bar represents all notes issued more than 10 trading days from the start and end of the month (a period of 0-2 days depending on the month). The y-axis represents the number of notes issued on this trading day. The Goldman Roll period is defined as in HPW. We define the Active Issuance period as the 5-day period ending with the 2nd to last trading day of the month. As shown in Panel B, this is the period when the frequency of large CLN pricing is greatest.

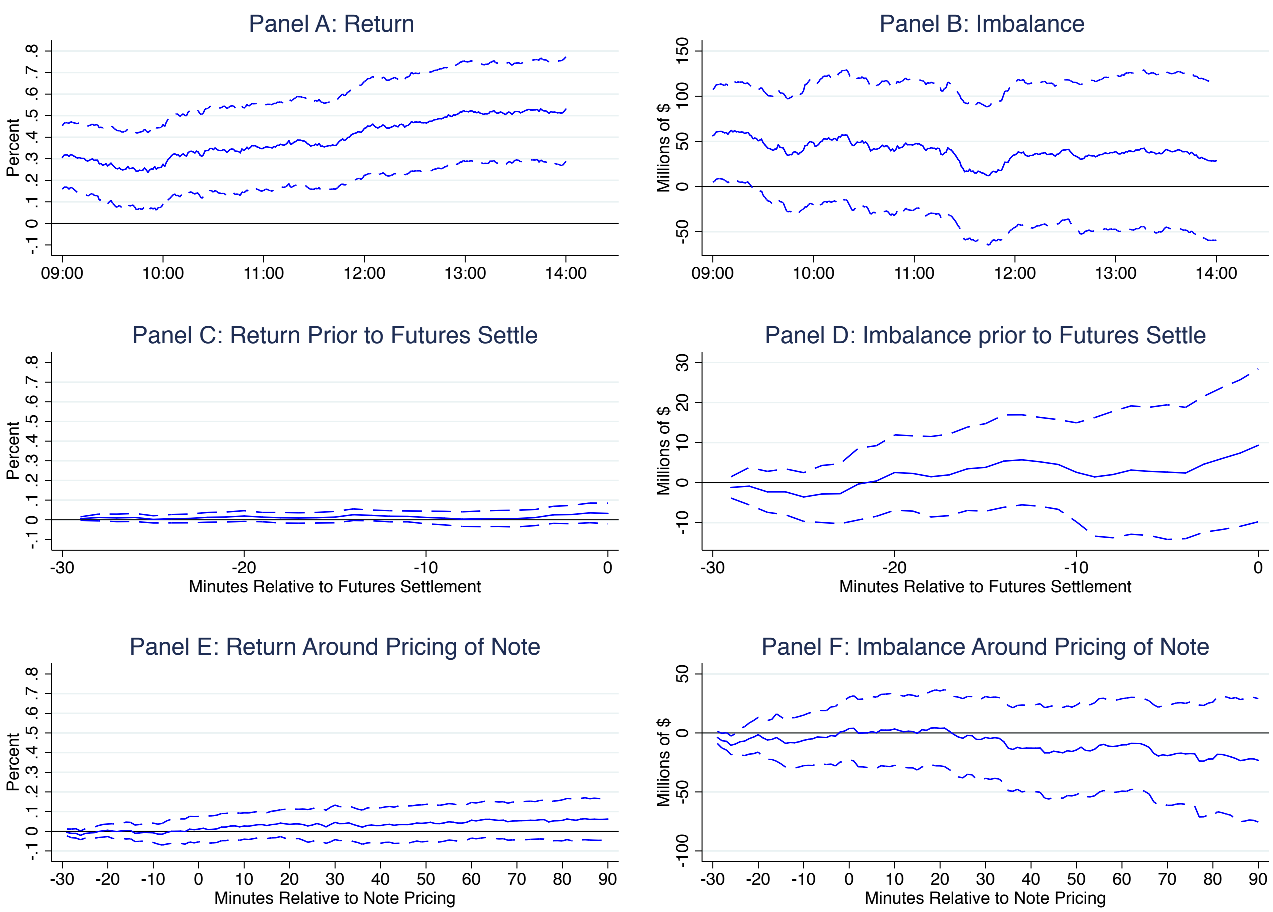


Figure 9. Intraday Returns and Imbalances on CLN pricing dates during Active Issuance Period

The figure shows the average returns and imbalances over expanding windows of the trading day for the underlying commodity on the pricing dates of CLNs with at least $10 million of face value issued in the Active Issuance Period for which we have intraday data (See Table 8). Returns are measured in percent and imbalance in millions of dollars. Panels A and B measure cumulative return and imbalance from the previous day's settlement through 2pm in New York. Panels C and D show cumulative return and imbalance in the 30 mintues prior to the daily settlement. Panels E and F show returns and imbalance in the 30 minutes prior to and 90 minutes after the pricing of the note.

Table 1. Daily WTI futures volumes for June 2013

The table shows volume for the days June of 2013 (in thousands of contracts) of the July 2013 and August 2013 delivery futures contracts.

	July 2013 Contract			August 2013 Contract			All other contracts		
	Globex			Globex			Globex		
Trade Date	Single Month	Cal. Spread	Floor	Single Month	Cal. Spread	Floor	Single Month	Cal. Spread	Floor
20130603	214.2	55.4	2	13.6	61.2	4.1	15.9	235.8	17.2
20130604	226.7	56.7	8.4	13.2	58.5	7.5	18.2	269.1	37.3
20130605	189.4	56.7	12.3	11.7	40.5	3	13.2	219.7	23
20130606	178.4	68.3	5.8	15.3	71.7	3.6	21.8	277.4	20.6
20130607	219.4	75.3	17.8	19.2	76.6	9.4	31.4	366.1	26.4
20130610	124.9	67.9	18.1	14.7	69.6	10.7	12.3	214.5	25.1
20130611	174	59.4	6.7	23.5	57.7	5.7	14.7	191	20.6
20130612	170	53.1	9.2	26.7	71.4	9.3	14.3	177.1	6.2
20130613	144.6	57.7	8.3	38.7	61.6	6	18	186.5	18.7
20130614	161.8	51.1	14.3	48.8	66.5	5.3	42.4	307.5	34
20130617	150.1	71.7	7	54	78.7	6.7	26.2	186.5	21.1
20130618	81.9	50.6	6.7	65.7	75.5	4.9	15.3	191.9	12.1
20130619	31.7	45.8	11.3	144.8	92.1	4	26.9	271.1	15.8
20130620	7.1	13.9	0.1	282.9	81.5	3.3	45.5	343.9	19
20130621	-	-	-	267.4	52.6	-	93.6	261.7	-
20130624	-	-	-	223.9	75.5	4.9	39.5	336.2	31.4
20130625	-	-	-	176.4	78.9	5.1	29.5	445.8	43.6
20130626	-	-	-	221.1	59.4	1.7	33	255.7	12.2
20130627	-	-	-	188.4	67.5	2.4	33.5	255.3	16
20130628	-	-	-	177.4	52.7	1.8	36.1	257.4	18.5

Table 2. Summary Data for Near Month Futures by Minute

The table shows means and standard deviations for minute-by-minute returns, trading volume, and signed trading volume (imbalance). Statistics for volume and imbalance are reported in both number of contracts and millions of dollars of notional value. The sample is January 1st, 2008 to April 1st 2014 for Brent Crude, and January 1st, 2007 to April 1st 2014, for all other commodities. The settlement minute is the minute prior to daily settlement. We exclude minutes before 7:30 AM or after 4:00 PM in New York.

CME WTI Crude Oil (All Minutes)						**CME WTI Crude Oil (Settlement Minute)**					
	Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)		Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)
Mean	0	38,351	-2.1	33.2	-0.2	Mean	-0.01	217,517	-50.9	187.8	-4.1
St. Dev.	0.09	46,415	169.4	42.1	14.9	St. Dev.	0.11	91,007	412.4	90	35.8
# of Min					921,522	# of Min					1,824

ICE Brent Crude Oil (All Minutes)						**ICE Brent Crude Oil (Settlement Minute)**					
	Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)		Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)
Mean	0	13,443	-0.4	14.3	0	Mean	-0.01	72,268	-2.1	79.3	-0.3
St. Dev.	0.09	21,784	101.9	23.2	10.4	St. Dev.	0.11	60,912	223.4	67.8	23.4
# of Min					793,281	# of Min					1,572

CME Gold (All Minutes)						**CME Gold (Settlement Minute)**					
	Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)		Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)
Mean	0	17,141	-1.3	21.6	-0.2	Mean	0	66,494	18	83.3	2
St. Dev.	0.05	26,047	92.6	36.1	12.1	St. Dev.	0.06	42,180	189.5	63.4	23.9
# of Min					918,897	# of Min					1,825

CME Copper (All Minutes)						**CME Copper (Settlement Minute)**					
	Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)		Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)
Mean	0	3,709	-0.1	2.8	0	Mean	0	32,386	10.2	23.2	0.7
St. Dev.	0.07	6,654	27.8	5.4	2.2	St. Dev.	0.12	27,388	116.9	24.5	9.2
# of Min					894,550	# of Min					1,867

CME Corn (All Minutes)						**CME Corn (Settlement Minute)**					
	Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)		Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)
Mean	0	33,550	-6.9	9.2	-0.2	Mean	0.01	434,837	119.7	120.2	4.3
St. Dev.	0.13	69,547	277.7	17.9	7.5	St. Dev.	0.25	298,308	1030.5	96.6	29.9
# of Min					510,149	# of Min					1,810

CME Wheat (All Minutes)						**CME Wheat (Settlement Minute)**					
	Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)		Ret. (%)	Vol. (# Contracts)	Imb. (# Contracts)	Vol. (Mil $)	Imb. (Mil $)
Mean	0	12,891	-2.6	4.5	-0.1	Mean	-0.04	214,641	-60.8	74.5	-1.9
St. Dev.	0.15	30,431	106.8	9.5	3.7	St. Dev.	0.36	167,621	583.6	64.3	20.8
# of Min					486,757	# of Min					1,808

Table 3. Price Impact Regressions

The table shows the results from univariate regressions where the dependent variable is one-minute returns and the independent variable is one-minute imbalance. Return is measured in percentage and imbalance is measured in millions of dollars, (ie. a coefficient of 0.01 represents a return response of 0.01% per million dollars of imbalance). The left column for each commodity shows the results using all minutes in the sample, while the right columns shows results using only returns and imbalances in the minute prior to futures settlement. Standard errors are shown in parentheses and T-statistics are shown in brackets. The sample is 1/1/2007 to 4/1/2014. We exclude data prior to 2008 for Brent and minutes prior to 7:30 AM or after 4:00 PM in New York.

	WTI Crude		Brent Crude		Gold	
	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute
Imbalance	0.0034***	0.0012***	0.0029***	0.0010***	0.0022***	0.0012***
	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)
	[275.1]	[19.0]	[155.5]	[12.8]	[99.8]	[11.1]
Constant	0.0006***	0	0.0001	-0.0052*	0.0003***	-0.0003
	(0.0001)	(0.0023)	(0.0001)	(0.0027)	(0.0000)	(0.0013)
	[7.8]	[0.0]	[0.6]	[-1.9]	[8.5]	[-0.2]
Observations	929,018	1,824	800,867	1,572	926,415	1,825
R-squared	0.328	0.172	0.124	0.046	0.305	0.224

	Copper		Corn		Wheat	
	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute
Imbalance	0.0113***	0.0042***	0.0064***	0.0045***	0.0143***	0.0086***
	(0.0001)	(0.0003)	(0.0002)	(0.0002)	(0.0005)	(0.0008)
	[112.7]	[12.3]	[35.4]	[19.1]	[30.5]	[11.1]
Constant	0.0003***	-0.0024	0.0011***	-0.0065	0.0010***	-0.0194**
	(0.0001)	(0.0026)	(0.0001)	(0.0050)	(0.0002)	(0.0077)
	[5.2]	[-0.9]	[8.1]	[-1.3]	[5.4]	[-2.5]
Observations	902,105	1,867	516,997	1,810	493,374	1,808
R-squared	0.147	0.113	0.188	0.297	0.148	0.254

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 4. Inferring the Price Impact of Order Flow from Daily Data

Panel A shows the results of regressions to estimate the price impact of order flow from daily data. The dependent variables are the log of the price impacts estimated for a single commodity in a calendar year. The independent variables are the log of daily futures return volatility and the log of the average daily volume (in millions of $ of futures notional) across all maturities of futures for a given commodity. Price impacts are defined as the slope of a regression of minute-by-minute returns (in %) on minute-by-minute imbalance (in millions of $) as in Table 3. All variables are obtained for Brent, WTI, Gold, Copper, Wheat, and Corn for each calendar year from 2008 to 2013. In column (1), the impacts are computed using all minutes. In column (2), the impacts are computed using only the minute prior to daily futures settlement. In column (3), the two sets of impacts are pooled in a single regression with a settlement minute dummy variable. Panel B uses the regression estimates of specification (3) in Panel A and estimates impacts for a broad set of commodity contracts. The contracts are sorted from lowest impact to highest. Estimates are calculated for the period 2003 to 2014 where data are available and averages for all years are reported (See Table IA.3 in internet appendix for all commodity-year impact estimates.)

Panel A: Regressions of Price Impacts on Daily Volatility and Volume

	Log(Order Flow Impact)		
	All Minutes (1)	Settlement Minutes (2)	Combined (3)
Log(Daily Volatility)	0.853***	0.747***	0.800***
	[5.506]	[4.408]	[6.864]
Log(Average Volume)	-0.532***	-0.671***	-0.601***
	[-11.471]	[-13.226]	[-17.240]
Settlement Dummy			-0.689***
			[-8.633]
Constant	3.257***	3.443***	3.694***
	[5.117]	[4.950]	[7.693]
Observations	36	36	72
R-squared	0.871	0.886	0.888

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Panel B: Average Predicted Impacts by Commodity

Contract	Daily Volatility (%)	Average Volume ($Mil/Day)	Estimated Impact All Min (%/$Mil)	Estimated Impact Settle Min (%/$Mil)
LME Copper	1.16	15,631	0.004	0.002
CME Crude Oil	2.12	37,551	0.004	0.002
LME Aluminum	0.99	7,694	0.005	0.003
ICE Brent Crude Oil	1.91	31,049	0.006	0.003
CME Gold	1.20	14,456	0.006	0.003
CME Soy	1.65	7,656	0.009	0.004
LME Zinc	1.30	3,087	0.012	0.006
CME Corn	1.86	5,002	0.012	0.006
CME Natural Gas	3.13	9,207	0.013	0.006
LME Nickel	1.51	2,849	0.015	0.007
CME Silver	2.12	4,159	0.019	0.010
CME Wheat	2.06	2,257	0.021	0.011
CME Copper	1.84	2,571	0.023	0.012
LME Lead	1.49	1,360	0.024	0.012
LME Tin	1.31	533	0.034	0.017
ICE Cotton	1.86	720	0.035	0.017
CME Platinum	1.37	383	0.065	0.033
CME RBOB Gasoline	2.28	9,663	0.088	0.044
CME Palladium	2.03	169	0.169	0.085

Table 5. Size and Predicted Price Impacts for Sources of Financial Investment

The table shows summary statistics two sources of financial investment in commodity markets. Panel A shows the weekly changes in position of commodity-index traders from the CFTC. Panel B shows the total face value and the calculated trade size necessary to delta hedge the notes on the pricing dates for each commodity. Panel C summarizes days with CLN determination for all commodities. All three panels also report the predicted price impact of order flow associated with each source. This is calculated as the size of the potential flow (The Change in Position for Panel A and Delta Hedge Size for Panels B and C) multiplied times the estimate of price impact per million dollars of imbalance for the applicable commodity-year combination (see Table 4).

Panel A: Changes in Positions of Index Traders

		Change in Position ($Mil)				Predicted Impact (%)			
	N	mean	stdev	min	max	mean	stdev	min	max
Corn	382	-8.3	225.8	-972.7	1278.9	-0.03	1.02	-3.90	4.73
Wheat	382	-7.8	140.4	-1254.0	414.0	-0.05	1.12	-6.00	3.52
					Correlation of Changes in Corn and Wheat:		0.23		

Panel B: Pricing Dates of Commodity-Linked Notes

		Notional ($Mil)			Δ Hedge Size ($Mil)			Predicted Impact(%)		
	N	mean	min	max	mean	min	max	mean	min	max
Gold	199	17.2	2.0	157.9	12.1	0.1	108.2	0.020	0.000	0.198
Brent Crude	114	12.5	2.0	103.4	7.5	0.4	56.0	0.010	0.000	0.310
WTI Crude	80	13.8	2.0	75.9	8.8	0.5	63.0	0.012	0.001	0.083
Palladium	41	13.3	2.3	80.2	7.7	0.7	43.5	0.194	0.010	1.034
LME Copper	34	16.0	2.1	155.5	13.8	0.4	172.4	0.025	0.001	0.312
Silver	25	23.2	2.0	84.9	15.9	0.4	54.9	0.071	0.001	0.267
Corn	18	18.7	2.0	59.9	15.3	1.2	51.1	0.088	0.004	0.404
Natural Gas	8	15.1	3.1	55.4	7.6	1.4	27.1	0.039	0.010	0.110
RBOB Gasoline	4	16.9	2.5	42.3	12.5	0.4	33.8	0.038	0.001	0.102
Cotton	2	7.5	5.0	10.0	4.3	3.2	5.5	0.045	0.033	0.057
Platinum	2	35.4	7.3	63.6	43.1	4.7	81.6	0.684	0.062	1.307
Lead	1	5.0	5.0	5.0	4.7	4.7	4.7	0.076	0.076	0.076
Zinc	1	11.0	11.0	11.0	10.3	10.3	10.3	0.072	0.072	0.072
Nickel	1	23.0	23.0	23.0	21.6	21.6	21.6	0.186	0.186	0.186
Aluminum	1	17.0	17.0	17.0	15.9	15.9	15.9	0.039	0.039	0.039
Tin	1	4.0	4.0	4.0	3.7	3.7	3.7	0.090	0.090	0.090
Total	532	15.6	2	157.9	10.7	0.1	172.4	0.039	0.000	1.307

Panel C: Determination Dates of Commodity-Linked Notes

		Notional Value ($Mil)			Δ Hedge Size ($Mil)			Predicted Impact (%)		
	N	mean	min	max	mean	min	max	mean	min	max
Prior to 2019/01	524	15.8	2.0	184.0	7.7	-23.3	228.0	-0.03	-2.76	0.02
...and $\Delta > 0$	214	18.0	2.0	156.0	18.7	0.0	228.0	-0.07	-2.76	0.00
...and $10+ Mil	92	35.2	10.0	156.0	36.8	0.9	228.0	-0.13	-2.76	0.00
Prior to 2014/02	414	16.3	2.0	184.0	8.4	-14.8	228.0	-0.03	-2.76	0.01
...and $\Delta > 0$	160	19.4	2.0	156.0	21.9	0.0	228.0	-0.08	-2.76	0.00
...and $10+ Mil	71	37.5	10.0	156.0	42.4	0.9	228.0	-0.15	-2.76	0.00

Table 6. Regressions of Return and Imbalance on Changes in the Positions of Commodity-Index Traders

The table shows the results from regressions of weekly futures imbalance and futures returns on changes in the positions of index traders as reported by the CFTC. For the imbalance regressions (columns (1) - (3)), futures imbalance and changes in index trader positions are measured in number of contracts. For the return regressions (columns (4) - (6)), futures returns are measured in percent and changes in index positions are standardized to have a standard deviation of one. In columns (1) and (4), the dependent variable is the imbalance or return for the entire trading day summed across the trading days in the week. For columns (2) and (5), the dependent variable is the total return or imbalance in the 30 minutes prior to futures settlement summed across the trading days in the week. In columns (3) and (6) the return and imbalance in the single settlement minute is summed across the trading days in the week. Data are 1/1/2007 to 4/1/2014.

Panel A: Changes in Corn Index Positions

Dependent Var:	Futures Imbalance				Futures Return		
	Full Day (1)	30 Minutes Prior to Settle (2)	Settle Minute (3)		Full Day (4)	30 Minutes Prior to Settle (5)	Settle Minute (6)
Δ Corn Index Positions	0.371**	0.137***	0.0433***	Standardized Δ Corn Index Positions	0.619*	0.281***	0.0811***
	[2.354]	[3.743]	[4.616]		[1.951]	[4.794]	[3.818]
Constant	-108.3***	-8.619***	0.25	Constant	0.315	0.068	0.0774***
	[-10.43]	[-3.014]	[0.380]		[1.318]	[1.135]	[3.806]
Obs	382	382	382	Obs	382	382	382
R-sq	0.024	0.043	0.078	R-sq	0.017	0.055	0.04

Panel B: Changes in Wheat Index Positions

Dependent Var:	Futures Imbalance				Futures Return		
	Full Day (1)	30 Minutes Prior to Settle (2)	Settle Minute (3)		Full Day (4)	30 Minutes Prior to Settle (5)	Settle Minute (6)
Δ Wheat Index Positions	0.514***	0.244***	0.113***	Standardized Δ Wheat Index Positions	0.492	0.481***	0.243***
	[4.466]	[5.596]	[3.776]		[1.531]	[5.047]	[3.864]
Constant	-37.73***	-5.966***	-2.693***	Constant	-0.0571	-0.273***	-0.169***
	[-9.907]	[-4.525]	[-4.010]		[-0.229]	[-3.582]	[-4.117]
Obs	382	382	382	Obs	382	382	382
R-sq	0.07	0.124	0.104	R-sq	0.01	0.095	0.085

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 7. Average Returns and Predicted Impacts of Delta Hedges on CLN Pricing Dates

Panels A-D show average futures returns and the average value of the predicted price impacts of delta hedging trades on days with CLN pricing. Panel A includes all dates with CLN pricing in the sample. Panel B includes only dates with more than $10 million of face value. Panels C and D are the same as A and B, but only consider the commodities (Brent crude oil, WTI crude oil, copper, gold, and corn) and periods where we have intraday data. In each panel, column (1) includes all dates, column (2) considers dates outside of the monthly Goldman Roll Period, column (3) considers dates during the Goldman Roll Period, column (4) considers dates in the Active Issuance Period, and column (5) considers days outside of the Active Issuance Period. (See Figure 8 for definition of the Goldman Roll and Active Issuance periods). Panel E regresses the pricing date returns on dummies indicating if the date is in the Active Issuance Period or outside of the Goldman Roll Period. The four columns correspond to the samples in column (1) of Panels A - D.

Panel A: All Dates

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg.	0.13*	0.28***	-0.12	0.44***	-0.05
tstat	[1.74]	[3.29]	[-0.85]	[4.32]	[-0.54]
Predicted Impact of Delta Hedges					
Avg.	0.04	0.04	0.03	0.06	0.03
Obs	532	338	194	201	331

Panel B: Dates w/ $10+ Million of Face Value

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg.	0.29***	0.42***	0.01	0.66***	-0.03
tstat	[2.84]	[3.55]	[0.07]	[5.03]	[-0.22]
Predicted Impact of Delta Hedges					
Avg.	0.08	0.08	0.07	0.09	0.06
Obs	220	153	67	104	116

Panel C: All Dates w/ Available Intraday Data

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg.	0.02	0.17**	-0.23	0.36***	-0.21**
tstat	[0.24]	[1.99]	[-1.55]	[3.53]	[-1.99]
Predicted Impact of Delta Hedges					
Avg.	0.02	0.02	0.01	0.02	0.01
Obs	426	264	162	171	255

Panel D: Dates w/ $10+ Mil. and Intraday Data

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg.	0.18*	0.33***	-0.15	0.57***	-0.20
tstat	[1.76]	[2.86]	[-0.73]	[4.68]	[-1.30]
Predicted Impact of Delta Hedges					
Avg.	0.03	0.03	0.03	0.03	0.03
Obs	171	118	53	84	87

Panel E: Regressions with Dummy Variables

	Pricing Day Return			
Sample:	Panel A	Panel B	Panel C	Panel D
Active Issuance Period Dummy	0.40**	0.76***	0.55***	0.84***
	[2.28]	[2.92]	[2.71]	[3.24]
Non-Goldman Roll Dummy	0.16	-0.11	0.04	-0.12
	[0.80]	[-0.37]	[0.20]	[-0.44]
Constant	-0.12	0.01	-0.23*	-0.15
	[-0.85]	[0.07]	[-1.84]	[-0.84]
Obs	532	220	426	171
R-sq	0.02	0.05	0.04	0.08

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 8. Intraday Futures Returns and Imbalance on CLN pricing dates in Active Issuance Period

Table shows the average returns (in %) and imbalance (in millions of $) over various intraday periods during the pricing dates of CLNs in the Active Issuance Period. The sample is restricted to the commodities (Brent crude oil, WTI crude oil, copper, gold, and corn) and periods where we have intraday data. Panel A shows returns and Panel B shows imbalance in the associated futures market. Panels C and D repeat the analysis but restricts the sample to days with at least $10 million of face value.

Panel A: Return on Pricing Dates in Active Issuance Period

	Full Day (1)	Prior to 9:00 AM (2)	30 Min Prior to Settle (3)	Settle Minute (4)	30 Min Prior to CLN Pricing (5)	One Hour Centered on CLN Pricing (6)	Pricing Minute (7)
Average	0.36***	0.14*	0.06**	-0.00	0.02	0.03	0.00
	[3.53]	[1.80]	[2.15]	[-0.41]	[0.66]	[0.78]	[0.84]
Observations	171	171	171	171	171	171	171

Panel B: Imbalance on Pricing Dates in Active Issuance Period

	Full Day (1)	Prior to 9:00 AM (2)	30 Min Prior to Settle (3)	Settle Minute (4)	30 Min Prior to CLN Pricing (5)	One Hour Centered on CLN Pricing (6)	Pricing Minute (7)
Average	5.43	11.63	12.18*	2.98*	-0.63	-9.59	0.49
	[0.18]	[0.68]	[1.88]	[1.69]	[-0.07]	[-0.81]	[0.32]
Observations	171	171	171	171	171	171	171

Panel C: Return on Pricing Dates in Active Issuance Period w/ $10+ Mil of Face Value

	Full Day (1)	Prior to 9:00 AM (2)	30 Min Prior to Settle (3)	Settle Minute (4)	30 Min Prior to CLN Pricing (5)	One Hour Centered on CLN Pricing (6)	Pricing Minute (7)
Average	0.57***	0.31***	0.03	-0.00	0.01	0.03	0.01
	[4.68]	[4.13]	[1.23]	[-0.43]	[0.33]	[0.80]	[1.06]
Observations	84	84	84	84	84	84	84

Panel D: Imbalance on Pricing Dates in Active Issuance Period w/ $10+ Mil of Face Value

	Full Day (1)	Prior to 9:00 AM (2)	30 Min Prior to Settle (3)	Settle Minute (4)	30 Min Prior to CLN Pricing (5)	One Hour Centered on CLN Pricing (6)	Pricing Minute (7)
Average	55.11	56.15**	9.34	1.95	3.38	-3.79	2.17
	[1.27]	[2.18]	[0.97]	[0.86]	[0.25]	[-0.22]	[1.28]
Observations	84	84	84	84	84	84	84

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table 9. Futures Returns on CLN Determination Dates

The table repeats the analysis in Table 7, but uses the average return on days with CLN determination rather than CLN pricing, and with predicted return calculated as the inverse of the determination date delta times the face value. Panels A uses all days with a note that has a delta greater than zero on the a determination date prior to 1/1/2019. Panel B restricts this to days with at least $10 million of Face Value. Panels C and D further restricts the sample to notes with determination dates before 2/1/2014 to replicate the determination date event study of HPW.

Panel A: All Dates

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg	0.03	0.12	-0.14	0.32**	-0.14
tstat	[0.32]	[0.95]	[-0.67]	[2.01]	[-1.00]
Predicted Impact of Unwinding Delta Hedges					
Avg	-0.07	-0.08	-0.04	-0.04	-0.08
Obs	214	147	67	81	133

Panel B: Dates w/ $10+ Mil Face Value

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg	-0.11	-0.04	-0.26	0.25	-0.46*
tstat	[-0.67]	[-0.27]	[-0.67]	[1.44]	[-1.76]
Predicted Impact of Unwinding Delta Hedges					
Avg	-0.13	-0.14	-0.08	-0.06	-0.19
Obs	92	64	28	45	47

Panel C: Dates prior to 2014/02

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg	0.09	0.08	0.36*	-0.05	-0.08
tstat	[0.65]	[0.35]	[1.94]	[-0.33]	(0.02)
Predicted Impact of Unwinding Delta Hedges					
Avg	-0.08	-0.09	-0.05	-0.05	-0.09
Obs	160	110	50	53	107

Panel D: Prior 2014/02 w/ $10+ Mil Face Value

	All (1)	Outside Goldman Roll (2)	During Goldman Roll (3)	During Active Issuance Period (4)	Outside Active Issuance Period (5)
Realized Daily Returns					
Avg	-0.00	-0.09	0.21	0.32	-0.25
tstat	[-0.00]	[-0.43]	[0.54]	[1.31]	[-0.97]
Predicted Impact of Unwinding Delta Hedges					
Avg	-0.15	-0.17	-0.09	-0.07	-0.21
Obs	71	50	21	31	40

t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Internet Appendix for Order Flows and Financial Investor Impacts in Commodity Futures Markets

Mark Ready
University of Wisconsin-Madison

Robert Ready
University of Oregon

May 16, 2019

IA.1 Vector Autoregressions as in Hasbrouck 1991

As a first step to understanding the impact of order flows in this market, we follow the Vector Autoregression approach developed in Hasbrouck (1991). Specifically, assume that the (log) quote midpoint for the commodity evolves according to:

$$q_t = m_t + s_t$$

Where m_t is the "efficient price" based on all relevant information, including public announcements and order flow up to time t, and the s_t component captures transient market microstructure effects. The efficient price evolves according to:

$$m_t = m_{t-1} + w_t$$

where the increments w_t are mean zero, have variance σ_w^2, and are serially independent at all lags. The s_t process has zero unconditional mean and is jointly covariance stationary with w_t. We observe the evolution of log quote midpoints, $r_t = q_t - q_{t-1}$, and the signed order flow x_t, and following Hasbrouck 1991 we assume these evolve according to the following VAR:

$$\begin{aligned} r_t &= a_1 r_{t-1} + a_2 r_{t-2} + \ldots + \quad b_0 x_t + \quad b_1 x_{t-1} + b_2 x_{t-2} + \ldots + v_{1,t} \\ x_t &= c r_{t-1} + c_2 r_{t-2} + \ldots + \quad d_1 x_{t-1} + d_2 x_{t-2} + \ldots + v_{2,t} \end{aligned}$$

In the above VAR, $v_{1,t}$ denotes the impact of public announcements in period t and $v_{2,t}$ denotes the surprise in current period order flow, and these have variances σ_1^2 and σ_2^2, respectively. The assumption that the current period order flow does not depend on the current period public announcement allows the above VAR to be recast in the following VMA representation:

$$\begin{aligned} r_t &= v_{1,t} + a_1^* v_{1,t-1} + a_2^* v_{1,t-2} + \ldots + b_0^* v_{2,t} + b_1^* v_{2,t-1} + b_2^* v_{2,t-2} + \ldots \\ x_t &= c_1^* v_{1,t-1} + c_2^* v_{1,t-2} + \ldots + v_{2,t} + d_1^* v_{2,t-1} + d_2^* v_{2,t-2} + \ldots \end{aligned}$$

The VAR is estimated using OLS, giving the coefficients as well as estimates for σ_1^2 and σ_2^2. Then a Cholesky decomposition recovers the coefficients. This VMA representation allows for the calculation of impulse response functions. Hasbrouck 1991 shows that the fraction of the variance of the efficient price innovations w_t that is due to the innovations in the order flow is given by:

$$R_w^2 = \frac{(\sum_{t=0}^{\infty} b_t^*)^2 \sigma_2^2}{(\sum_{t=0}^{\infty} b_t^*)^2 \sigma_2^2 + (1 + \sum_{t=1}^{\infty} a_t^*)^2 \sigma_1^2} \tag{1}$$

When examining equity data, Hasbrouck 1991 applies the approach to trade-by-trade data, although trades within 5 seconds of each other are aggregated into a single observation. In contrast, we aggregate data into one-minute time intervals. As in Hasbrouck 1991, we set the lagged values returns and imbalances to zero at the start of each trading day. We examine three primary dimensions of liquidity based on the VAR, including: b_0 and b_0^* , which are the initial impact of order flow and the initial price impact of the unpredictable portion of order flow. Higher values for these coefficients may suggest a higher fraction of trades come from the informed, or that the information held by informed traders is more valuable, or that the market is illiquid for other reasons. $\sum b_t^*$, the permanent price impact of an innovation in order flow. We illustrate this with impulse response functions to test if the impact of an innovation in order flow is reversed in subsequent minutes. R_w^2, the fraction of the efficient price variance explained by order flow innovations (as with b_0^*, a higher value implies more information coming from trades, but this measure is relative to the amount of information that arrives through public announcements).

Table IA.1 shows the results of the regressions shown in equation (1) for the full sample. Imbalance is measured in 100s of contracts, and return is expressed in percentage to facilitate interpretation. Again, for most of the commodities, 100 contracts translates into roughly \$10 million of notional (with the exception of Corn and Wheat, where 100 contracts translates into approximately \$2.5 million of notional over the sample).

The parameter b_0 from equation the VAR is shown in the first row of each of the return columns in Table IA.1. This is the estimated response of the futures price to the order imbalance in the current minute. When the regressions from Table IA.1 are converted to the VMA representation (results not shown), we find that the values of b_0 from the VAR are very close to the values of b_0^*. This is not surprising, because as shown in the remaining rows of Table IA.1, current returns are not sensitive to past imbalances and there is only modest persistence in imbalances. The low R^2 values in the imbalance regressions indicates that most of the current minute imbalance is unpredictable.

The b_0 value of 0.033 for WTI futures shows that a minute with 100 contracts (about \$10 million notional) of buy (sell) imbalance will create a same-minute price increase (decrease) of 3.3 basis points. A roughly \$10 million dollar flow yields an impact of approximately 3 basis points for gold, similar to the WTI, but a trade of \$10 million notional value trade moves copper and corn prices approximately 10 basis points (the coefficient for corn must be multiplied by four to adjust for the lower notional value per contract). For all four of these commodities, the R^2 of these return regressions is relatively large, and results in a correspondingly high value of R_w^2 from the VMA representation, both results suggesting that order flow imbalance in these markets is playing a major role in price discovery.

To ascertain whether or not these price impacts from order flow reverse in subsequent minutes. We use the VMA representation to calculate impulse response functions. The graphs of these functions are shown in Figure IA.1. This figure plots impulse response functions for returns in response to a one standard deviation innovation in order flow and in public price news for the six commodities. The primary takeaway from these plots is that the price impacts of both order flow and public return news are mostly permanent at seven-minute horizons. For oil, gold, and copper there is essentially no reversal or continued trend in prices. For corn, wheat, and Brent there is a small reversal after a movement in prices unrelated to order flow, but for a price move corresponding to order flow we see very little reversal.

IA.2 Nonlinearities in Settlement Minute Returns

Figure IA.2 shows scatter plots of imbalance and returns in the minute prior to futures settlement for each of the six commodities. Also presented are the linear regression line, and fitted non-parametric LOESS smoother. For all six of the commodities, large flows generally lead to smaller impacts per dollar.

IA.3 Price Impacts of Order Flow In Different Subperiods of the Trading Month

In this table we repeat the analysis of Table 3 in the main text, but add dummy variables for both the Goldman Roll Period and the Active Issuance Period (see Figure 7 in the main text). Table IA.2 shows the results. The coefficients on the interaction with the dummy variables tend to be very small relative to the baseline estimates of impact, and the coefficients on the interaction terms are not consistent in sign. We conclude that there is little different in price impacts for trades in the Goldman Roll or the Active Issuance Period.

IA.4 Predicted Price Impacts for Commodities and Calendar Years

Table IA.3 shows the daily average volume, daily volatility, and predicted impacts for a large set of commodities from 2003 to 2014 where we have data. The average of these estimates and the regression specification are reported in Table 4 in the main text.

IA.5 Retail Trading in the United States Oil Fund

In this section we examine flows to the United States Oil ETF (USO). This fund has been studied in several other papers including Bessembinder, Carrion, Tuttle, and Venkataraman (2016) who study the impact of the fund rolling its futures positions from the front month contract to the next month contract, and Irwin and Sanders (2012) who find no impact of fund share creation or redemption on oil futures returns. The USO is very liquid, and may be used by informed traders to trade on oil news, so we proceed by first isolating order flow imbalance from retail investors using the algorithm proposed by Boehmer, Jones, and Zhang (2017).

Panel A of Table IA.4 shows summary statistics for trading volume in the USO. There is substantial retail volume in the USO, we find that these volumes are small when compared to the volume in WTI futures. The one-minute standard deviation of daily imbalance from retail traders in the USO is \$0.2 million, compared to \$15.2 million for WTI futures across all minutes (\$36.6 million in the settlement minute) as reported in the main text. This suggests that futures trades driven by uninformed volume from USO retail traders should have a relatively small effect on futures markets.

In Panel B of Table IA.4 we test for impacts of this volume. We first test for price impact of this volume at a daily frequency, and in doing so we obtain a puzzling result. Days with buying (selling) by retail investors in the USO tend to be days with negative (positive) return (column (6))). However, when we examine the relation of retail imbalance and returns at the one-minute frequency, we see that this result is an artifact of retail investors pursuing contrarian strategies. Retail investors tend to buy after drops in prices, and aggregating up to daily frequencies leads to a spurious contemporaneous correlation. When we move to a one-minute frequency (column (7)), the finding reverses, and we find a small positive association between retail order flow and price changes, but this effect is very minor relative to the overall volatility of this market.

IA.6 Predicted Impacts and Actual Returns for CLN Pricing Dates

Figure IA.3 shows scatter plots (with regression lines and equations) of actual pricing date returns on predicted pricing date returns calculated as the size of the associated delta hedging trades times the commodity-year price impact estimates from Table IA.3. Panel D includes a dummy variable for whether or not the note was issued in the Active Issuance Period.

As we see from the figures, for most specifications there is a significant slope with a point estimate near one. This is broadly consistent with the idea that larger notes do in fact create some larger impact on prices. Even so, the predicted impacts of the notes are not large enough to explain the results. This is evident by the strongly significant intercept terms for the notes outside of the Goldman Roll or in the Active Issuance Period. Panel D shows all notes with a dummy variable for the active issuance period, and this dummy variable has a positive and highly significant coefficient of approximately 45 basis points. This suggests that it is whether or not the note is issued in this period, rather than the note's size, that associates it with a large positive pricing date return.

References

Hasbrouck, J. 1991. The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies* 4:571–595.

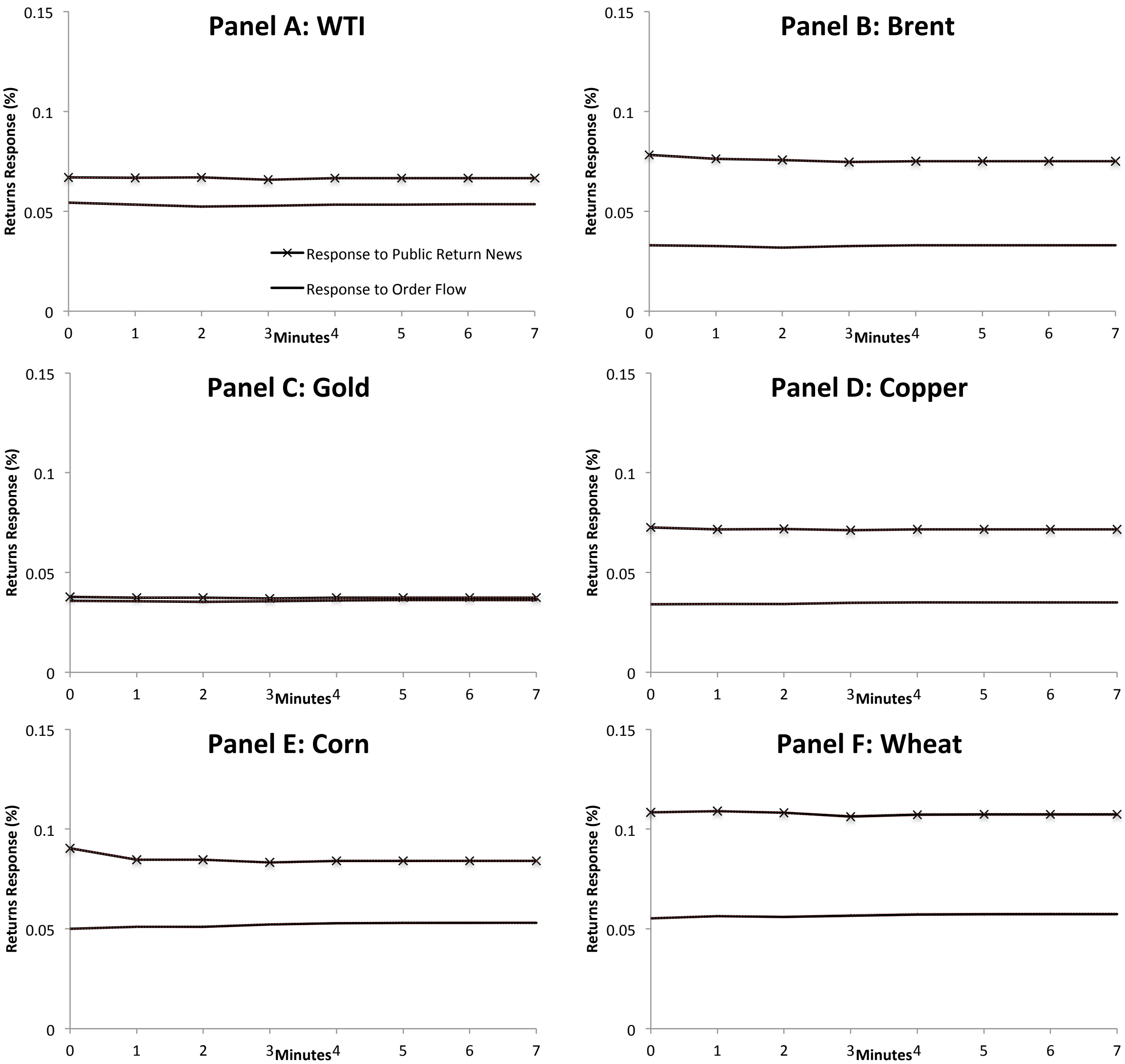


Figure IA.1. Return Impulse Response Functions for VARs

The figure shows the impulse response of returns to innovations in order flow and public news from the vector autoregression specification estimated in Table 3. Plots show return responses to one standard deviation innovations in public return news and unanticipated order flow.

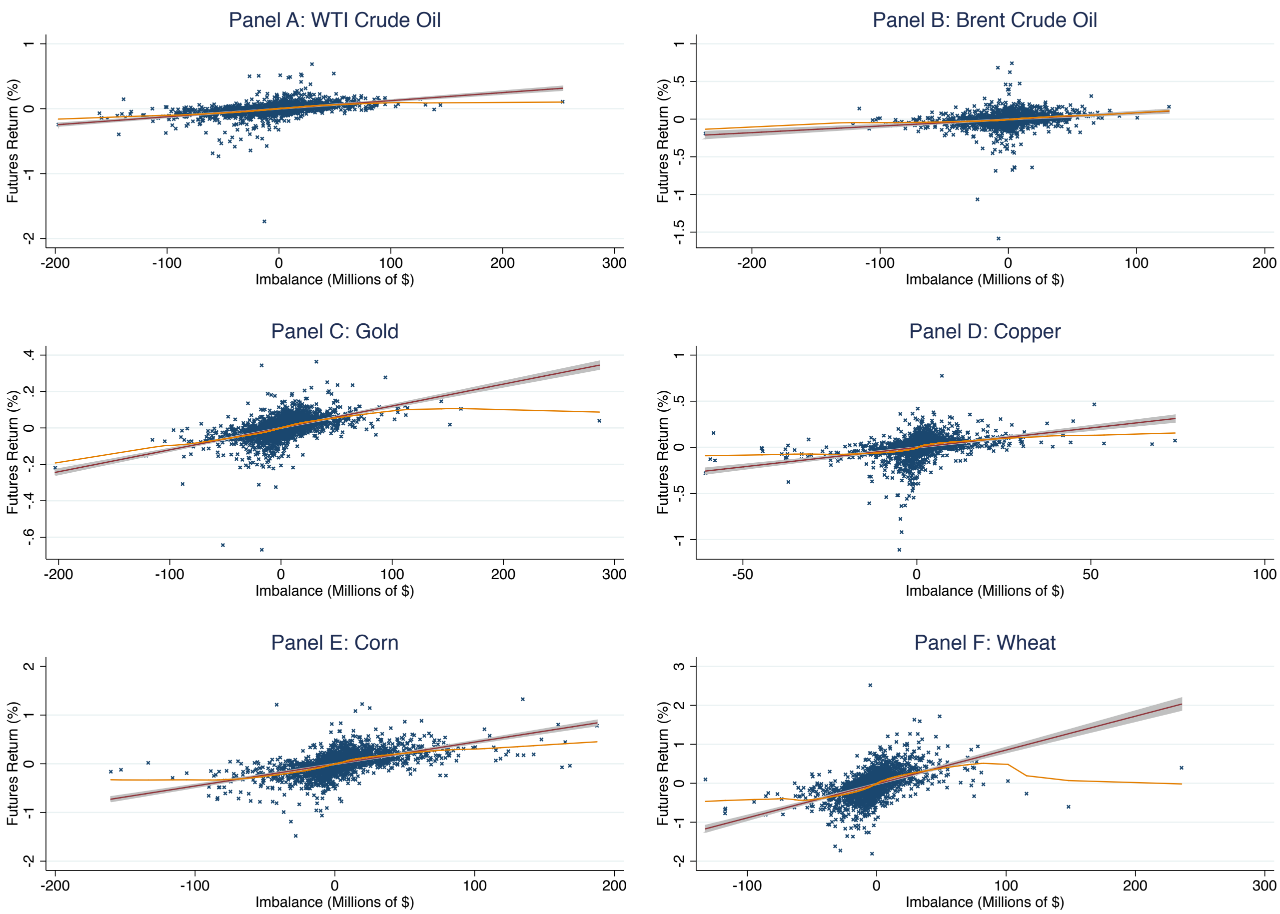


Figure IA.2. Nonlinearities in Imbalance and Return in Minute Prior to Futures Settlement

The figure shows scatter plots of order imbalance (in millions of \$) and return (in %) in the minute prior to settlement for each day across the sample. The shaded line shows linear fit and confidence interval. The single line shows a second-order LOESS smoother calculated using a tricube kernel with $\alpha = 0.8$. Data are 1/1/2007 to 4/1/2014. We exclude data prior to 1/1/2008 for Brent.

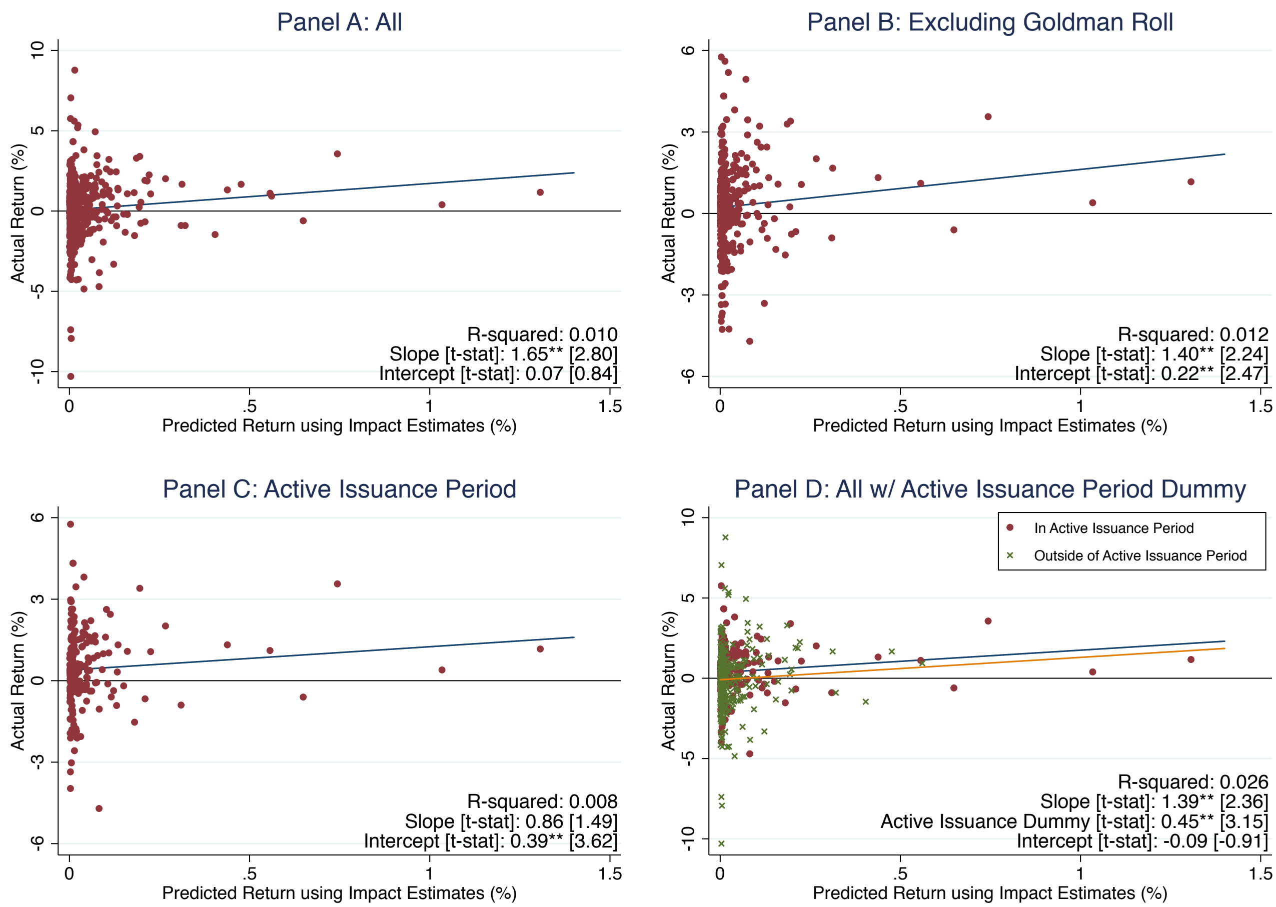


Figure IA.3. Actual Returns vs. Predicted Returns on CLN pricing dates
Figure shows scatter plots with regression lines of actual pricing date returns on predicted pricing date returns calculated as the size of the associated delta hedging trades times the commodity-year price impact estimates from Table IA.3. Panel D includes a dummy variable for whether or not the note was issued in the Active Issuance Period.

Table IA.1. Full Sample Price Impact VARs

The table shows the results from vector autoregressions of the form described in section IA.1. R_w^2 shown in the final row is the percentage of variation in returns explained by unexpected innovations in order flow, calculated from a vector moving average representation of the VAR. Return is measured in percent, while imbalance is measured in 100s of contracts. The sample is 1/1/2007 to 4/1/2014. We exclude data prior to 2008 for Brent and minutes prior to 7:30 AM or after 4:00 PM in New York.

	WTI Crude		Brent Crude		Gold		Copper		Corn		Wheat	
	Return	Imbalance	Return	Imbalance	Return	Imbalance	Return	Imbalance	Return	Imbalance	Return	Imbalance
Imb. (t)	0.033***		0.033***		0.031***		0.097***		0.020***		0.053***	
	[778.258]		[377.948]		[730.820]		[434.507]		[407.343]		[312.961]	
Imb (t-1)	-0.003***	0.089***	-0.004***	0.127***	-0.002***	0.062***	-0.006***	0.078***	-0.000***	0.103***	-0.001***	0.085***
	[-62.650]	[66.277]	[-41.088]	[104.247]	[-36.209]	[47.071]	[-22.881]	[65.939]	[-3.181]	[63.330]	[-6.285]	[53.497]
Imb (t-2)	-0.001***	0.029***	-0.002***	0.047***	-0.001***	0.030***	-0.005***	0.039***	-0.001***	0.051***	-0.002***	0.038***
	[-25.083]	[21.670]	[-19.935]	[38.760]	[-16.365]	[22.641]	[-18.549]	[32.783]	[-15.182]	[31.108]	[-9.215]	[23.055]
Imb (t-3)	-0.001***	0.029***	-0.002***	0.042***	-0.001***	0.028***	-0.003***	0.033***	-0.001***	0.035***	-0.002***	0.027***
	[-21.857]	[21.871]	[-16.525]	[34.798]	[-18.914]	[22.021]	[-13.401]	[28.185]	[-16.457]	[22.729]	[-8.448]	[16.793]
Ret (t-1)	-0.058***	1.650***	-0.037***	0.443***	-0.061***	1.900***	-0.034***	0.175***	-0.102***	1.589***	-0.063***	0.299***
	[-55.155]	[64.120]	[-33.273]	[30.564]	[-58.543]	[75.018]	[-31.644]	[34.466]	[-74.581]	[39.908]	[-43.874]	[24.441]
Ret (t-2)	-0.014***	0.440***	-0.012***	0.144***	-0.026***	0.681***	-0.001	0.082***	-0.026***	0.486***	-0.020***	0.109***
	[-13.515]	[17.065]	[-11.029]	[9.900]	[-24.994]	[26.798]	[-0.673]	[16.211]	[-18.938]	[12.230]	[-13.035]	[8.343]
Ret (t-3)	-0.007***	0.125***	-0.009***	0.060***	-0.012***	0.288***	-0.001	0.029***	-0.011***	0.127***	-0.010***	0.027**
	[-7.101]	[4.893]	[-8.099]	[4.145]	[-11.746]	[11.419]	[-0.786]	[5.709]	[-9.606]	[3.653]	[-6.954]	[2.194]
Cons	0.001***	-0.017***	0	-0.003***	0.000***	-0.012***	0.000***	-0.001***	0.001***	-0.060***	0.001***	-0.022***
	[8.280]	[-9.939]	[0.603]	[-2.817]	[9.294]	[-12.070]	[5.297]	[-3.266]	[9.829]	[-15.818]	[5.199]	[-15.132]
Obs	919,910	919,910	792,989	792,989	915,852	915,852	874,572	874,572	493,111	493,111	475,674	475,674
R^2	0.397	0.03	0.153	0.029	0.368	0.027	0.178	0.016	0.255	0.03	0.172	0.016
R_w^2	0.382		0.145		0.354		0.171		0.248		0.165	

t-stats in brackets
*** p<0.01, ** p<0.05, * p<0.1

Table IA.2. Price Impact of Order Flow in Goldman Roll and Active Issuance Periods

This table shows regressions as in Table 3 in the main text, with a dummy variable indicating if the trading day of the month falls in the Goldman Roll Period (Panel A) or the Active Issuance Period (Panel B) interacted with imbalance. See Figure 7 in the main text for a description of the Goldman Roll and the Active Issuance Period. Both the regression constant and dummy variable terms are supressed and only the interaction term is shown.

Panel A: Price Impact of Imbalance in Goldman Roll Period

	WTI Crude		Brent Crude		Gold		Copper		Corn		Wheat	
	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute
Imbalance	0.0034***	0.0012***	0.0029***	0.0010***	0.0021***	0.0010***	0.0106***	0.0038***	0.0068***	0.0041***	0.0165***	0.0098***
	(0.00005)	(0.00009)	(0.00006)	(0.00011)	(0.00006)	(0.00015)	(0.00031)	(0.00047)	(0.00029)	(0.00028)	(0.00088)	(0.00064)
	[63.0]	[12.9]	[47.7]	[8.9]	[36.1]	[6.8]	[34.1]	[8.2]	[23.6]	[14.5]	[18.8]	[15.4]
Period Dummy × Imbalance	-0.0001	0.0002	-0.0002*	0.0001	0.0002**	0.0004**	0.0016***	0.0008	-0.0006	0.0005	-0.0027*	-0.0013
	(0.00008)	(0.00013)	(0.00009)	(0.00015)	(0.00007)	(0.00020)	(0.00050)	(0.00066)	(0.00053)	(0.00048)	(0.00143)	(0.00137)
	[-0.8]	[1.3]	[-1.9]	[0.4]	[2.2]	[2.2]	[3.2]	[1.2]	[-1.1]	[0.9]	[-1.9]	[-0.9]
Obs	930,690	1,824	802,294	1,572	928,071	1,825	903,770	1,867	517,023	1,810	493,386	1,808
R sq	0.327	0.173	0.124	0.046	0.305	0.232	0.148	0.121	0.212	0.289	0.179	0.261

Panel B: Price Impact of Imbalance in Active Issuance Period

	WTI Crude		Brent Crude		Gold		Copper		Corn		Wheat	
	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute	All Minutes	Settle Minute
Imbalance	0.0033***	0.0012***	0.0029***	0.0010***	0.0022***	0.0013***	0.0115***	0.0041***	0.0065***	0.0045***	0.0145***	0.0090***
	(0.00004)	(0.00007)	(0.00005)	(0.00009)	(0.00004)	(0.00015)	(0.00029)	(0.00039)	(0.00033)	(0.00028)	(0.00088)	(0.00096)
	[77.7]	[16.8]	[57.3]	[11.2]	[55.1]	[8.8]	[39.9]	[10.6]	[19.9]	[16.0]	[16.4]	[9.4]
Period Dummy × Imbalance	0.0004***	0.0005***	-0.0002**	-0.0001	-0.0003***	-0.0002	-0.0007	0.0005	-0.0000	-0.0006	0.0027*	0.0002
	(0.00009)	(0.00017)	(0.00009)	(0.00017)	(0.00009)	(0.00020)	(0.00055)	(0.00078)	(0.00057)	(0.00052)	(0.00146)	(0.00131)
	[4.3]	[2.7]	[-2.6]	[-0.4]	[-2.8]	[-1.2]	[-1.3]	[0.7]	[-0.0]	[-1.1]	[1.8]	[0.2]
Obs	932,513	1,824	803,865	1,572	929,873	1,825	905,297	1,867	517,087	1,810	493,401	1,808
R sq	0.328	0.175	0.124	0.046	0.305	0.226	0.147	0.113	0.212	0.289	0.178	0.259

*** $p<0.01$, ** $p<0.05$, * $p<0.1$

Table IA.3. Estimated Impact of Order Flow by Commodity and Year

This table shows the predicted price impact of order flow for different commodity contracts. For each commodity-year the table shows the average daily volume across all futures maturities and the standard deviation of the nearest-to-maturity future return. Impact for all minutes and the settlement minute are then calculated using the regression specification of Table ?? in the main text.

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
LME Copper													
Millions $ Day					15,262	18,393	12,871	22,616	30,455	28,505	31,456	28,009	15,631
Stdev Daily Returns					0.020	0.028	0.024	0.017	0.018	0.012	0.011	0.008	0.012
Est. Impact All Min.					0.005	0.006	0.007	0.004	0.003	0.003	0.002	0.002	0.004
Est. Impact Settle Min.					0.003	0.003	0.003	0.002	0.002	0.001	0.001	0.001	0.002
NYMEX Crude Oil													
Millions $ Day	5,591	8,745	13,412	18,492	34,988	53,928	33,670	53,040	65,585	52,734	57,454	52,970	37,551
Stdev Daily Returns	0.023	0.022	0.020	0.017	0.019	0.037	0.034	0.017	0.022	0.016	0.012	0.015	0.021
Est. Impact All Min.	0.011	0.008	0.006	0.004	0.003	0.005	0.005	0.002	0.002	0.002	0.002	0.002	0.004
Est. Impact Settle Min.	0.005	0.004	0.003	0.002	0.001	0.002	0.003	0.001	0.001	0.001	0.001	0.001	0.002
LME Aluminum													
Millions $ Day					10,618	12,422	7,835	10,113	14,267	11,933	12,287	12,849	7,694
Stdev Daily Returns					0.013	0.019	0.021	0.017	0.013	0.012	0.011	0.011	0.010
Est. Impact All Min.					0.005	0.006	0.008	0.006	0.004	0.004	0.004	0.004	0.005
Est. Impact Settle Min.					0.002	0.003	0.004	0.003	0.002	0.002	0.002	0.002	0.003
ICE Brent													
Millions $ Day	2,702	3,816	6,697	11,426	16,817	26,136	18,150	31,930	58,080	65,145	68,566	63,124	31,049
Stdev Daily Returns	0.021	0.023	0.019	0.016	0.017	0.035	0.027	0.016	0.018	0.014	0.010	0.013	0.019
Est. Impact All Min.	0.015	0.013	0.008	0.005	0.004	0.007	0.006	0.003	0.002	0.002	0.001	0.002	0.006
Est. Impact Settle Min.	0.008	0.007	0.004	0.003	0.002	0.003	0.003	0.001	0.001	0.001	0.001	0.001	0.003

Continued on next page

Table IA.3. – continued from previous page

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
COMEX Gold													
Millions $ Day	1,766	2,464	2,804	3,841	7,034	13,276	13,799	21,745	30,858	29,056	26,522	20,304	14,456
Stdev Daily Returns	0.010	0.010	0.008	0.015	0.010	0.019	0.014	0.010	0.013	0.010	0.014	0.009	0.012
Est. Impact All Min.	0.013	0.011	0.009	0.010	0.006	0.005	0.004	0.003	0.003	0.002	0.003	0.003	0.006
Est. Impact Settle Min.	0.007	0.005	0.004	0.005	0.003	0.003	0.002	0.001	0.001	0.001	0.001	0.001	0.003
COMEX Soy													
Millions $ Day	2,252	2,785	2,431	2,684	5,534	9,107	7,324	7,825	11,773	15,177	13,010	11,975	7,656
Stdev Daily Returns	0.014	0.022	0.017	0.012	0.014	0.029	0.022	0.014	0.014	0.015	0.013	0.014	0.017
Est. Impact All Min.	0.013	0.016	0.014	0.011	0.008	0.010	0.009	0.006	0.005	0.004	0.004	0.005	0.009
Est. Impact Settle Min.	0.006	0.008	0.007	0.005	0.004	0.005	0.004	0.003	0.002	0.002	0.002	0.002	0.004
LME Zinc													
Millions $ Day					4,081	3,015	2,638	3,900	4,816	5,753	5,949	6,891	3,087
Stdev Daily Returns					0.025	0.029	0.026	0.022	0.017	0.014	0.011	0.012	0.013
Est. Impact All Min.					0.014	0.020	0.019	0.013	0.009	0.007	0.006	0.006	0.012
Est. Impact Settle Min.					0.007	0.010	0.009	0.007	0.005	0.004	0.003	0.003	0.006
COMEX Corn													
Millions $ Day	887	1,150	1,148	2,504	4,070	6,497	3,785	6,056	10,714	9,967	7,465	5,781	5,002
Stdev Daily Returns	0.013	0.015	0.015	0.018	0.020	0.027	0.023	0.020	0.022	0.019	0.016	0.014	0.019
Est. Impact All Min.	0.022	0.020	0.020	0.014	0.012	0.012	0.014	0.009	0.007	0.006	0.007	0.007	0.012
Est. Impact Settle Min.	0.011	0.010	0.010	0.007	0.006	0.006	0.007	0.004	0.004	0.003	0.003	0.004	0.006
NYMEX Natural Gas													
Millions $ Day	4,138	4,219	6,691	6,286	8,327	13,682	7,882	11,147	12,294	10,555	12,538	12,728	9,207
Stdev Daily Returns	0.043	0.034	0.031	0.039	0.030	0.030	0.042	0.027	0.021	0.031	0.019	0.030	0.031
Est. Impact All Min.	0.028	0.019	0.013	0.019	0.011	0.008	0.019	0.008	0.006	0.010	0.006	0.009	0.013
Est. Impact Settle Min.	0.014	0.010	0.007	0.009	0.005	0.004	0.009	0.004	0.003	0.005	0.003	0.004	0.006

Continued on next page

Table IA.3. – continued from previous page

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
LME Nickel													
Millions $ Day					3,384	2,626	2,370	3,834	4,414	4,696	5,031	7,838	2,849
Stdev Daily Returns					0.027	0.036	0.032	0.022	0.019	0.015	0.012	0.017	0.015
Est. Impact All Min.					0.017	0.028	0.026	0.013	0.011	0.009	0.007	0.007	0.015
Est. Impact Settle Min.					0.009	0.014	0.013	0.007	0.005	0.004	0.004	0.003	0.007
COMEX Silver													
Millions $ Day	404	674	808	1,252	1,811	2,729	2,397	5,405	14,150	8,278	6,807	5,193	4,159
Stdev Daily Returns	0.013	0.023	0.014	0.028	0.016	0.032	0.024	0.020	0.030	0.019	0.021	0.015	0.021
Est. Impact All Min.	0.035	0.038	0.024	0.032	0.016	0.024	0.018	0.010	0.008	0.007	0.009	0.008	0.019
Est. Impact Settle Min.	0.018	0.019	0.012	0.016	0.008	0.012	0.009	0.005	0.004	0.004	0.004	0.004	0.010
COMEX Wheat													
Millions $ Day	469	541	639	1,314	2,497	3,148	1,869	2,675	3,460	4,038	3,390	3,041	2,257
Stdev Daily Returns	0.018	0.018	0.016	0.019	0.021	0.032	0.024	0.023	0.025	0.020	0.012	0.017	0.021
Est. Impact All Min.	0.040	0.037	0.029	0.021	0.016	0.022	0.022	0.017	0.016	0.012	0.010	0.012	0.021
Est. Impact Settle Min.	0.020	0.018	0.015	0.011	0.008	0.011	0.011	0.009	0.008	0.006	0.005	0.006	0.011
COMEX Copper													
Millions $ Day	247	411	657	1,001	1,206	1,456	1,550	3,474	4,910	5,778	5,650	4,511	2,571
Stdev Daily Returns	0.013	0.020	0.015	0.024	0.021	0.030	0.027	0.018	0.019	0.014	0.012	0.010	0.018
Est. Impact All Min.	0.047	0.045	0.028	0.032	0.025	0.032	0.027	0.012	0.010	0.007	0.007	0.007	0.023
Est. Impact Settle Min.	0.024	0.023	0.014	0.016	0.012	0.016	0.013	0.006	0.005	0.004	0.003	0.004	0.012
LME Lead													
Millions $ Day					1,219	1,290	1,033	1,663	2,619	2,937	2,817	2,745	1,360
Stdev Daily Returns					0.028	0.037	0.031	0.024	0.021	0.016	0.012	0.010	0.015
Est. Impact All Min.					0.032	0.046	0.041	0.023	0.015	0.012	0.010	0.010	0.024
Est. Impact Settle Min.					0.016	0.023	0.021	0.012	0.008	0.006	0.005	0.005	0.012

Continued on next page

Table IA.3. – continued from previous page

	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	Average
LME Tin													
Millions $ Day					376	522	1,242	636	984	779	926	928	533
Stdev Daily Returns					0.020	0.031	0.025	0.019	0.023	0.018	0.013	0.010	0.013
Est. Impact All Min.					0.048	0.061	0.029	0.034	0.030	0.029	0.021	0.019	0.034
Est. Impact Settle Min.					0.024	0.030	0.014	0.017	0.015	0.014	0.011	0.009	0.017
COMEX Cotton													
Millions $ Day	372	360	385	468	723	792	375	1,005	1,437	918	965	844	720
Stdev Daily Returns	0.018	0.023	0.017	0.014	0.013	0.026	0.020	0.021	0.026	0.019	0.013	0.014	0.019
Est. Impact All Min.	0.045	0.056	0.043	0.033	0.025	0.039	0.049	0.027	0.027	0.027	0.021	0.023	0.035
Est. Impact Settle Min.	0.023	0.028	0.022	0.017	0.012	0.019	0.025	0.014	0.014	0.014	0.010	0.012	0.017
COMEX Platinum													
Millions $ Day	38	51	68	85	133	217	200	479	672	812	961	885	383
Stdev Daily Returns	0.012	0.014	0.008	0.015	0.010	0.027	0.017	0.013	0.013	0.013	0.013	0.009	0.014
Est. Impact All Min.	0.140	0.129	0.085	0.097	0.061	0.089	0.063	0.032	0.026	0.023	0.021	0.019	0.065
Est. Impact Settle Min.	0.070	0.065	0.043	0.049	0.031	0.045	0.032	0.016	0.013	0.011	0.010	0.009	0.033
COMEX RBOB Gasoline													
Millions $ Day			7	1,134	6,872	8,723	5,979	9,866	14,666	17,897	16,385	15,103	9,663
Stdev Daily Returns			0.030	0.023	0.021	0.036	0.029	0.017	0.020	0.015	0.014	0.014	0.023
Est. Impact All Min.			0.790	0.028	0.009	0.014	0.013	0.006	0.005	0.004	0.004	0.004	0.088
Est. Impact Settle Min.			0.397	0.014	0.004	0.007	0.007	0.003	0.003	0.002	0.002	0.002	0.044
COMEX Palladium													
Millions $ Day	8	25	26	49	57	74	44	193	331	287	429	503	169
Stdev Daily Returns	0.024	0.023	0.017	0.025	0.011	0.030	0.022	0.024	0.022	0.017	0.016	0.012	0.020
Est. Impact All Min.	0.594	0.276	0.210	0.201	0.108	0.192	0.191	0.085	0.058	0.051	0.038	0.030	0.169
Est. Impact Settle Min.	0.298	0.139	0.106	0.101	0.054	0.096	0.096	0.043	0.029	0.025	0.019	0.015	0.085

Table IA.4. Regressions of Return and Imbalance on USO Retail Imbalance

Panel A reports summary statistics for imbalance in the USO. Panel B shows regressions of return and imbalance in WTI futures on retail imbalance in the United States Oil Fund (USO). In columns (1)-(3) and (5)-(7) the independent variable is the sum USO retail imbalance for each trading day. In columns (1) and (5), the dependent variables are the sum of WTI imbalance and WTI returns for all minutes in each trading day. In columns (2) and (6) the dependent variables are the sum across the 30 minutes prior to futures settlement in each trading day. In columns (3) and (7) the dependent variables are the imbalance and return in the single minute prior to settlement. In columns (4) and (8) all variables are measured at the 1-minute frequency. Data are from 1/1/2007 to 4/1/2014.

Panel A: Summary Stats for USO Order Flow Imbalance

		Imbalance ($Mil)				Predicted Impact (%)			
	N	mean	stdev	min	max	mean	stdev	min	max
By Minute:									
All Trades	692,738	0.0	1.5	-348.5	299.1	0.00	0.00	-0.35	0.49
Retail Trades	692,738	0.0	0.2	-98.3	36.5	0.00	0.00	-0.23	0.06
By Day:									
All Trades	1,824	-2.0	35.7	-351.6	366.1	0.00	0.06	-0.41	0.87
Retail Trades	1,824	0.2	5.9	-92.4	74.0	0.00	0.01	-0.22	0.11

Panel B: Regressions of WTI Return and Imbalance on USO Retail Imbalance

	WTI Futures Imbalance				WTI Futures Return			
	Daily	30 Min Prior to Settle	Settle Minute	All Minutes	Daily	30 Min Prior to Settle	Settle Minute	All Minutes
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
USO Retail Imb. (Day)	-0.694	-1.850***	-0.089		-0.086***	-0.016***	0.000	
	[-0.362]	[-2.726]	[-0.546]		[-10.197]	[-6.151]	[0.212]	
USO Retail Imb. (Minute)				1.405***				0.003**
				[2.80]				[2.331]
Constant	-85.446***	11.386***	-4.110***	-0.167***	0.024	0.046***	-0.005**	0.000
	[-7.370]	[2.987]	[-4.866]	[8.90]	[0.471]	[2.936]	[-2.045]	[1.081]
Obs	1,824	1,824	1,824	692,738	1,824	1,824	1,824	692,738
R-sq	0.000	0.005	0.000	0.000	0.054	0.020	0.000	0.000

Robust t-statistics in brackets
*** p<0.01, ** p<0.05, * p<0.1

Risk Appetite and Intermediation by Swap Dealers*

Scott Mixon† Esen Onur‡

May 24, 2019

Preliminary

ABSTRACT

Using novel data on WTI crude oil swaps and futures positions of individual dealers, we relate dealer liquidity provision to changes in risk appetite over the 2007–2015 period. We find evidence consistent with a theoretical model of dealers who hedge bespoke contracts with standard, liquid instruments and face basis risk. Swap dealers provide less intermediation service for customers, and hedge the existing swaps more tightly, when risk appetite decreases. We also find evidence that dealers have larger single commodity swap books if they have a larger index book, suggesting that increased commodity index activity enhances liquidity provision to hedgers.

JEL classification:

Keywords: Dealers, Hedging, Risk Appetite, Liquidity Provision, Commodity Index

*The research presented in this paper was co-authored by Scott Mixon and Esen Onur, CFTC employees who wrote this paper in their official capacities. The analyses and conclusions expressed in this paper are those of the authors and do not reflect the views of other members of the Office of CFTC Chief Economist, other Commission staff, or the Commission itself.

†Commodity Futures Trading Commission; smixon@cftc.gov

‡Commodity Futures Trading Commission; eonur@cftc.gov

"We don't have a proprietary trading business in commodities, we have a client business that takes risk." – Bob Diamond, CEO of Barclays Capital[1]

I. Introduction

How do dealers manage a derivatives trading book? To what extent do dealers use derivatives contracts to increase their risk exposures or to hedge their business risks? Despite the fundamental nature of these questions, the existing evidence is sparse and mostly anecdotal, partly because of data limitations. Further, market participants claim that regulatory reform efforts since the financial crisis have reduced the overall intermediation capacity of dealers and altered the behavior of dealers. For example, a comment letter from the Coalition of Derivatives End-Users to the Board of Governors, OCC, FDIC, CFTC, and SEC on 17 October 2018 states that "Following the [Volcker] rule's implementation, however, we have been and are concerned by an apparent reduction in the availability of certain bespoke and less liquid derivative products." Have commercial end-users of derivatives, who often desire bespoke instruments in relatively illiquid products or tenors, faced increased difficulty in transacting post-Crisis?

Management of a dealer's commodity derivatives trading book provides a rich environment for examining risk management within the context of a financial institution. Anecdotally, dealer activity in commodity markets is often distinguished from activity in more liquid equity and fixed income markets. Whereas the archetypal market maker in equities participates in an active "flow" business, quoting two-sided markets and holding very transient positions, market making in commodities is described as much more sporadic in nature. Whereas financial market participants may have a wide latitude in

[1]Barclays Capital Investor Seminar Q&A, 17 June 2009, www.home.barclays/content/dam/barclayspublic/docs/InvestorRelations/IRNewsPresentations/2009Presentations/Barclays-Capital-Investor-Seminar-QandA.pdf

activities, commercial producers and consumers are more limited with respect to the products traded. Commercial users are beholden to the large volumes they would like to hedge over long time frames, while other participants prefer smaller volumes and shorter time frames. Dealers emphasize that their market making activities in commodities involves the warehousing of risks - holding positions for a period of time, rather than immediately laying off the risk. They also highlight the fact that their positions often involve basis risk, because they use standardized, liquid hedging instruments to hedge the customized products desired by customers. Dealers also readily engage in arbitrage activities across instruments.[2]

Dealers generally reduced their balance sheets as they re-evaluated their business models post-Crisis. Some commodity dealers are known to have exited the business during the years following the Crisis, and most are believed to have reduced their footprint in the space.[3] The regulatory environment also became more restrictive as regulatory reforms were implemented. Basel constraints limited balance sheet flexibility, the Volcker rule limited proprietary trading activity, the Federal Reserve proposed stricter rules on bank activity in physical commodities, and public reporting of swap transactions have all been cited as factors in reducing liquidity for commodity derivatives end-users. (See the discussions in Mixon (2018) and Commodity Futures Trading Commission (2018)). However, there is limited systematic evidence linking the balance sheets of individual dealers to their activities providing liquidity to commodity clients.

To address this lack of quantitative evidence in the literature, we provide an in-depth examination of swap dealing activity, pre- and post-Crisis, for a wide cross-section of

[2]See, for example, the discussion on p. 113 of Morgan Stanley's September 30, 2011 10-Q Quarterly Report.

[3]E.g., see "Deutsche Bank Quits Commodities Under Regulatory Pressure", David Sheppard and Ron Bousso, *Reuters*, 5 December 2013; "Credit Suisse to Exit Commodities Trading", Max Colchester and Sarah Kent, *Wall Street Journal*, 22 July 2014; "Barclays Exit from Energy Trading Stirs Concerns over Liquidity", Catherine Ngai and Olivia Oran, *Reuters*.

individual dealers. To our knowledge, this is the first time that such portfolio-level data has been used to gauge the provision of risk-bearing services ("liquidity") by OTC swap dealers and their related hedging activity in the listed derivatives market. We find evidence that dealer risk appetites persistently declined in the post-Crisis period, and that client-facing derivatives books declined along with risk appetite. This strong, supply-side relation holds even after controlling for demand-side shocks such as the growth in U.S. production of crude oil to be hedged.

Anecdotally, real-economy firms that use crude oil swaps to hedge prefer long-dated contracts in order to get net short exposure spread out over a period of months or years, while dealers tend to hedge with the most liquid, short-dated contracts. Dealers also face offsetting flows from index investors, who generally desire long exposure to near-dated contracts. We provide empirical evidence describing the interaction of these clienteles, as intermediated by swap dealers. We find that the majority of dealer futures hedging activity is in near-dated contracts when hedging commodity index exposures, but the hedging activity for single commodity swaps in WTI is more dispersed across the term structure. Finally, we also find evidence supporting the idea that dealers hedged a given swap exposure more tightly (i.e., took on less basis risk) during the post-Crisis part of the sample when balance sheet constraints were tighter. Overall, we find that dealers generally use futures contracts to mitigate their risks due to swap exposures, but they do not (or cannot) hedge all risks. Consistent with our theoretical model of optimizing dealers who provide swap exposure to hedgers but retain some risk (e.g., basis risk), equilibrium outcomes in the market are related to the risk appetite of dealers.

The reminder of the paper is organized as follows. Section II provides an overview of the related literature. Section III offers empirical facts motivating our modeling approach. Section IV introduces our theoretical model, while Section V gives the details of our data and summary statistics. Section VI provides the results of our empirical analysis and Section VII concludes.

II. Related Literature

Our paper is closest to the work of Naik and Yadav (2003*b*), who examine the trading behavior of individual dealers in the UK gilt market. Whereas they examine futures hedging of cash gilt positions over a one-year sample, we examine futures hedging of dealer swap positions over an eight year period encompassing major structural and regulatory changes. We also link the hedging behavior to the risk appetite of individual firms. Naik and Yadav (2003*a*) also provide insight into the internal structure of dealer portfolio behavior, providing evidence that trading desk activity within dealers is often decentralized and not coordinated. Begenau, Piazzesi and Schneider (2015) provide evidence that futures trading activity within dealers is often centralized, allowing dealers to provide more single commodity liquidity when it is offset with index activity within the same firm.

Also closely related is the literature focusing on dealer balance sheets and leverage representing a focal point of asset pricing. Adrian and Shin (2010) and Adrian, Boyarchenko and Shachar (2017) provide evidence that dealer risk appetite varies over time and can potentially rationalize movements in asset prices. Etula (2013), in particular, examines the impact of dealer leverage on commodity price changes and concludes that dealer balance sheets are important for explaining energy price changes.

Given the crude oil data used in the analysis, our work is also related to the literature specifically examining that market. Mixon, Onur and Riggs (2018) describe the aggregate positions taken by swap dealers and their counterparties in the WTI crude oil market over the 2014 – 2016 period, but they do not explore the behavior of individual dealers, nor do they test hypotheses explaining position changes over time. Irwin and

Sanders (2012) examine the aggregate, index-linked positions from the same data collection, but they do not examine the single commodity swap data or the activities of individual dealers.

Acharya, Lochstoer and Ramorai (2013) consider a model of risk averse hedgers and arbitrageurs and connect the model's implications to risk premia in the energy derivatives market. They find that the risk appetite of energy producers varies the intensity with which hedging occurs, and this is related to the risk bearing capacity of dealers. In contrast, we allow for risk averse dealers and hedgers in a world with swaps and futures markets, and we focus on the quantities traded more than the pricing implications.

Our paper is also linked to the work of Gârleanu, Pesersen and Poteshman (2009), who examine the pricing of derivatives when demand pressure matters. They focus on cases where the dealer is unable to hedge derivatives risks completely due to market frictions. In contrast, our theoretical motivation assumes that risk averse dealers optimally choose to underhedge, compared to the full hedging in the baseline, perfect markets case. Consistent with anecdotal dealer descriptions of their business, we find evidence that hedging of derivatives risk varies over time. This supports the view of Stulz (1996) that dealers should engage in selective risk taking as part of their business, but it is less in line with the view of Froot and Stein (1998) that intermediaries should hedge fully.

In the theoretical framework motivating our empirical analysis, we incorporate a two-tiered market similar to the frameworks of Vogler (1997) and Viswanathan and Wang (2004). In those papers, however, dealers interact with clients in a public market and then manage inventory risk in a second stage, dealer-only market. In our framework, we consider a dealer who first engages with a customer in the swaps market and then manages portfolio risk in a related, public futures market.

Our work differs from the aforementioned studies in that we examine portfolio-level data for individual dealers and relate their dealing and hedging activities to observable measures of risk appetite. We link balance sheet variables and trading VaR measures to dealer liquidity provision in the swaps market and liquidity taking in the futures market. The panel nature of the data, covering the pre- and post-Crisis period, provides a unique window into the business of managing a derivatives trading book in practice.

III. A First Look at the Data

In this section, we present several charts highlighting salient institutional features of the data that we will subsequently incorporate into the theoretical and empirical analysis that follows. Our focus is to use the futures and swap positions, aggregated across a broad sample of large dealers in WTI-related swaps, to illustrate broad co-movements of swap and futures positions, the relative magnitudes of commodity index-related positions and single commodity swap positions, and the change in the number of swap dealers engaged in this business over the sample period. In all cases, the positions are measured in terms of futures contract equivalents; options or swaps including optionality are delta-adjusted. Further, the futures positions include dealer holdings in the NYMEX WTI futures contracts, the ICE cash-settled WTI futures contract, and the NYMEX WTI calendar swap futures contract.

How do dealers manage the risk of a derivatives trading book? Figure 1 provides our first empirical evidence that variation in dealer futures risk appears to offset the variation in their swap risk. The solid black line shows the aggregate, net WTI-related swap position of dealers in our sample. Dealers were net short swap exposure during much of the 2008-2012 period, and they were net long during the remainder of the sample that ends in October 2015. The figure also displays the net WTI futures position of dealers. The data closely track the net futures positions reported in the CFTC's Commitments of Traders report.

The swap and futures positions have roughly the same magnitude at any given point in time, but with opposite signs. During the first three years of the sample, dealers were net short WTI exposure via swap and long a similar exposure via listed futures. For the next three years (2011–2013), the net swap position trended upward while the net futures position trended in the opposite direction. Finally, the two series trend in opposite directions during the final portion of the sample. Broadly speaking, the figure supports the idea that the dealer community used the WTI futures market to hedge their swap exposure and not to take strong directional bets on the WTI price.

Researchers have emphasized the lack of hedging activity for firms using derivatives. Chernenko and Faulkender (2011) conclude that a meaningful amount of interest rate swap activity by nonfinancial firms is not due to hedging, but to "speculation". Similarly, Begenau, Piazzesi and Schneider (2015) conclude that banks do not typically use interest rate swaps to hedge other businesses. Naik and Yadav (2003*b*) find that dealers use gilt futures to hedge cash gilt exposure locally, with cash and futures positions generally moving in opposite directions. However, they conclude that dealers were not targeting a fully hedged book but were targeting a net short duration position in both futures and cash during their sample. In contrast, the net WTI exposures described above do not appear to show a strong tendency for dealers to take, on aggregate, a net short or long position; the first impression is that dealers hedge a significant amount of swap exposure with futures.

We explore the net swap position of dealers in more detail in Figure 2. In aggregate, we find that dealers are net long WTI exposure via single commodity swaps (because hedgers are net short) and net short WTI exposure via commodity index swaps (because index investors have been net long). The aggregate swap dealer position therefore reflects the relative magnitudes of these two exposures, each of which is on the order of magnitude of hundreds of thousands of futures contracts. During the first several years of the sample, net WTI exposure due to index activity exceeded net WTI exposure to single commodity swap activity, resulting in a net short position of roughly 100,000 contracts. During the final years of the sample, commodity index activity declined in size and single commodity swap exosure increased, resulting in a dealer net long position on the order of 100,000 contracts.

Mixon, Onur and Riggs (2018) examine similar data for the final year and a half of the sample examined here and find that WTI positions due to index investing were smaller in size than the positions due to hedging activities of commercial end-users. Examination of the longer time series of data reveals that these relative sizes varied over time. Based on the evidence displayed in the chart, we conclude that a meaningful examination of dealer activity in the crude oil market must incorporate information on both the direct dealer exposure to single commodity swaps in WTI as well as the indirect exposure via commoidty index contracts. To date, the publicly available data on swaps (the CFTC's Index Investment Data report) incorporates only the index activity; therefore, the present study represents a signficant step forward in understanding the total financial activity of swap dealers in the commodity space.

The final chart in this section speaks to the comprehensiveness and variation through time in business activity. The aggregate sample includes 26 dealer firms; the CFTC initially identified dozens of entities likely to have large swap positions and requested information from them. Over time, some new firms were added, and some of the dealers left the sample because of bankruptcies or due to leaving the commodity swap business. In Figure 3, we display the three month average of the number of dealers engaging in the index swap dealing business and the number of dealers engaging in single commodity swaps dealing business for WTI. For this measure, we include dealers reporting more than a *de minimis* quantity (100 contracts) of WTI swap exposure. Examination of the chart suggests that the sample contains well over a dozen firms engaged in index swaps and a similar number engaged in single commodity swaps, with no obvious discontinuities in population coverage of dealers. Nonetheless, it is evident that the number of dealers engaged in the single commodity swaps business decreased dramatically after 2012, even as the number with an index swap book remained steady. We conclude that further analysis of this market should allow for a decline in the number of market participants, even as the size of the market increases.

In the sections that follow, we formulate a simple model that captures these initial observations, and we test the empirical predictions of the model on the data. We model swap dealers who intermediate different clienteles: index investors and hedgers. Commodity index investors are typically long and the indexes typically represent positions in liquid, nearby contracts, even though the swaps referencing the indexes might be for much longer tenors. Commercial hedgers in crude oil, such as exploration and production companies, are known to take positions over multiple maturities in order to hedge crude production or consumption over intervals that could span several years. Hence, dealers retain risk even after facing these two offsetting flows. We also allow dealers to choose an optimal hedge hedge in the futures market, based on their risk appetite.

Because the equilibrium prices and quantities depend crucially on dealer risk appetite, we carefully consider the comparative statics as dealer risk appetite varies. Such variation in risk appetite gives the static model a more dynamic flavor and could, in principle, generate a decline in dealer activity consistent with the evidence in in Figure 3. Note that we do not expliclty model the factors driving risk appetite (e.g., by modeling the effects of particular regulations), but we treat it as exogenous and use multiple proxies for it in our empirical analysis.

The observations also motivate us to split the sample into the 2007–2011 "Crisis and rule writing period" and the 2012–2015 "post-Crisis" period at several points in the analysis. As noted by Adrian, Boyarchenko and Shachar (2017), the nature of the rule implementation process led to a variety of measures occuring or being anticipated simultaneously. Given the difficulty of unraveling so many simultaneous factors, we view the splitting of the sample in this way as a transparent way to illustrate disparities across the two periods.

IV. Theory

The goal of our theoretical work is to provide a simple model that captures the key elements of a dealer's business transacting long-dated swap contracts to producers, and optimally hedging the risk in the liquid, short-dated futures market. Consider an economy with one commodity, "WTI crude oil". The producers of the commodity hedge production in the swap market by selling the commodity forward at the swap strike K. The swap dealer facilitates the swaps by going long and hedges by going short in the futures market. Mirroring actual market practice, swap dealers hedge swaps with the liquid, active futures contracts that are not perfect substitututes for these swap exposures, and hence they take on basis risk. Producers and swap dealers interact in the swaps market, while swap dealers interact with futures traders in the futures market. Futures traders do not participate in the swaps market, and producers do not enter into the futures market.

In abstract terms, we consider two correlated instruments that are traded, and for which we find the equilibrium prices and quantities traded. For our particular application, we will treat these two instruments as the short-dated futures contract and the long-dated swap contract. The prices of futures and swap markets are based on the same fundamental with different but correlated error terms. Liquidation values for the two instruments are given by $f = E[s] + \epsilon_f$ in the futures market and $k = E[s] + \epsilon_k$ in the swap market.

Error terms for the two instruments are mean zero noise ($E[\epsilon_f) = E[\epsilon_k] = 0$), have the same variance ($var(\epsilon_f) = var(\epsilon_k) = \sigma^2$), and have a correlation of ρ. All market participants have the same beliefs over these parameters and stochastic processes. Market prices for futures and swap instruments are determined endogeneously and are denoted by F and K, respectively. All market participants are mean-variance optimizers and their risk aversion parameters are denoted by γ_P for producers, γ_{SD} for swap dealers and γ_{FT} for futures traders. Finally, there is exogenous index investment in the economy, which is denoted by I. We model this index investment taking place directly in the futures market.

A. Equilibrium in the Futures Market

We begin by solving for the futures market equilibrium, conditional on the dealer's swap demand. In the futures market, swap dealers trade only with futures traders and optimally hedge their swap positions. Swap dealer demand for futures is denoted by Q^F_{SD}, and futures traders' demand for futures is denoted by Q^F_{FT}. As indicated above, the price F clears the market and comes from the equilibrium

$$Q^F_{SD} + Q^F_{FT} + I = 0 \tag{1}$$

The dealer has a portfolio of long-dated swaps and short-dates futures and maximizes

$$max\, E[Q^F_{SD}(f - F) + Q^S_{SD}(k - K)] - \frac{\gamma_{SD}}{2} var(Q^F_{SD}(f - F) + Q^S_{SD}(k - K)) \quad w.r.t. \quad Q^F_{SD} \tag{2}$$

The dealer's optimal futures demand is therefore

$$Q^F_{SD} = \frac{E[s] - F}{\gamma_{SD}\sigma^2} - \rho Q^S_{SD}, \tag{3}$$

where Q^S_{SD} is the dealer's swap demand.

The second source of demand in the futures market comes from the futures trader, who has no initial endowment and participates only in the futures market. The futures trader maximizes

$$max\, E[Q^F_{FT}(f - F)] - \frac{\gamma_{FT}}{2} var(Q^F_{FT}(f - F)) \quad w.r.t. \quad Q^F_{FT} \tag{4}$$

and her equilibrium demand is therefore given by

$$Q^F_{FT} = \frac{E[s] - F}{\gamma_{FT}\sigma^2} \tag{5}$$

Imposing the market clearing condition from Equation 1, the equilibrium futures price is

$$F^* = E[s] - (Q^S_{SD}\rho - I)\left[\frac{1}{\gamma_{SD}} + \frac{1}{\gamma_{FT}}\right]^{-1}\sigma^2 \tag{6}$$

Substituting this price into Equation 6, we find that the equilibrium futures demand by the swap dealer can be expressed as

$$Q^{F*}_{SD} = \rho Q^S_{SD}\,[R - 1] - IR, \tag{7}$$

where $R = \left[\frac{1}{\gamma_{SD}}\right]\left[\frac{1}{\gamma_{SD}} + \frac{1}{\gamma_{FT}}\right]^{-1}$. Note that for $\gamma_{FT}, \gamma_{SD} > 0$, $0 < R < 1$ and $(R-1) < 0$.

This equilibrium expression obviously comports with standard intuition. A dealer will short more futures if 1) the quantity of swaps Q^S_{SD} increases, 2) exogenous index investment I increases, or 3) if the futures provide a better hedge because ρ is larger and basis risk is lower. Further, our empirical tests also rely on intution from the partial derivative of Q^F_{SD} with respect to γ_{SD}:

$$\frac{\partial Q^F_{SD}}{\partial \gamma_{SD}} = (Q^S_{SD}\rho - I)\left[\frac{\partial R}{\partial \gamma_{SD}}\right] \tag{8}$$

where it can be shown that $\frac{\partial R}{\partial \gamma_{SD}} < 0$. As γ_{SD} increases and dealer risk appetite declines, the dealer hedge position in futures tends toward the more neutral position $-rhoQ^S_{SD}$ that obtains when $F = E[s]$. This neutral hedge would also obtain if there were no market impact ot the dealer's futures trading, or $\gamma_{FT} = 0$. In the model, $F > E[s]$ when index investment is large, increasing the market clearing price above the future, inducing dealers to "overhedge" their swap position and act as arbitrageurs. Technically, this occurs when $\rho Q^S_{SD} > I$. Similarly, dealers "underhedge" their swap position when the producer forward sales dominate and $F < E[s]$. In either case, an increase in dealer risk aversion attenuates the dealer position towards a more neutral stance. Our empirical tests on hedging behavior for dealers derive from these partial derivatives of this equation.

B. Equilibrium in the Swap Market

We also solve for the swap market equilibrium and derive empirical predictions relating the size of a dealer's single commodity swap book to risk aversion and other state variables. Only swap dealers and producers trade in the swap market. Swap dealer demand for swaps is denoted by Q^S_{SD} and producers have demand denoted by Q^S_P. The variable K is the swap strike, and we solve for the value that satisfies the market clearing equilibrium of $Q^S_{SD} + Q^S_P = 0$. Note that the market clearing solution is a function of Q^F_{SD}, which we solved for and presented in Equation 7.

Specifically, the swap dealer maximizes the following equation

$$max\, E[Q_{SD}^F(f-F)+Q_{SD}^S(k-K)]-\frac{\gamma_{SD}}{2}var(Q_{SD}^S(f-F)+Q_{SD}^S(k-K)) \quad w.r.t. \quad Q_{SD}^S \tag{9}$$

and the optimal swap market demand becomes

$$Q_{SD}^S = \frac{E[s]-K}{\gamma_{SD}\sigma^2} - \rho Q_{SD}^F* \tag{10}$$

where Q_{SD}^F* is as defined in 7. Substituting this back into 10 gives us

$Q_{SD}^S = \frac{E[s]-K}{\gamma_{SD}\sigma^2} - \rho Q_{SD}^S\,[R-1] - IR$ and solving for the equilibrium swap demand by the swap dealer, we get

$$Q_{SD}^{S*} = \left[\frac{E[s]-K}{\gamma_{SD}\sigma^2} + \rho IR\right]\left[1-\rho^2(1-R)\right]^{-1} \tag{11}$$

The other player in the swap market, the producer, also optimizes his swap demand by maximizing the following equation

$$max\, E[Q_0k + Q_P^S(k-K)] - \frac{\gamma_P}{2}var(Q_0k + Q_P^S(k-K)) \quad w.r.t. \quad Q_P^S \tag{12}$$

and the producer's demand in the swap market is therefore given by

$$Q_P^S = \frac{E[s]-K}{\gamma_{SD}\sigma^2} - Q_0 \tag{13}$$

Applying the market clearing condition, we solve for K^* using the market clearing condition

$$\left[\frac{E[s]-K}{\gamma_{SD}\sigma^2} + \rho IR\right]\left[1-\rho^2(1-R)\right]^{-1} + \frac{E[s]-K}{\gamma_{SD}\sigma^2} - Q_0 = 0 \tag{14}$$

and find

$$K^{*} = E[s] - \left[\left[\frac{1}{\gamma_{SD}}\right]\left[1-\rho^{2}(1-R)\right]^{-1} + \left[\frac{1}{\gamma_{P}}\right]\right]^{-1} \sigma^{2}\left[Q_{0} - \rho IR\left[1-\rho^{2}(1-R)\right]^{-1}\right] \tag{15}$$

Using this equilibrium price, we find that the equilibrium size of the producer's swap book by combining equations 13 and 15. Our testable predictions on the swap market are derived from the equilibrium given by

$$Q_{P}^{S*} = \frac{-Q_{0}\left[\frac{1}{\gamma_{SD}}\right] - \rho IR\left[\frac{1}{\gamma_{P}}\right]}{\left[\frac{1}{\gamma_{SD}}\right] + \left[\frac{1}{\gamma_{P}}\right]\left[1-\rho^{2}(1-R)\right]^{-1}} \tag{16}$$

Using equation 16, we can sign the partial derivatives of the producer's swap demand. We find that $\frac{\partial Q_{P}^{S*}}{\partial Q_{0}} < 0$, indicating that producers hedge more if they are endowed with more. Additionally, we also find $\frac{\partial Q_{P}^{S*}}{\partial I} < 0$, meaning producers hedge more if there is more index investment. The effect of swap dealer risk aversion on the producer's demand depends on the precise parameter values. Roughly speaking, however, we find that $\frac{\partial Q_{P}^{S*}}{\partial \gamma_{SD}} > 0$ if Q_0 is "large" compared to I. Under these conditions, we can state that producers would hedge less if swap dealers become more risk averse. The alternative scenario is that pricing is dominated by extremely large values of I, and the futures price far exceeds $E[s]$, dealers are incentivized to go short to benefit from this extreme imbalance. If dealer risk aversion increases, they want to decrease this short position, which would mean that producers would be hedging less. We generally consider the case where Q_0 is "large" compared to I as the more realistic case.

It is worth noting that because swap demand adds up to zero, $Q_{SD}^{S} + Q_{P}^{S} = 0$, the partial derivatives for the swap dealer demand are opposite of those for the producer. Namely, swap dealers have a larger swap book if producer endownment increases; if there is more exogenous index investment; and if dealer risk aversion decreases (again, assuming Q_0 is large compared to I). In the empirical analysis that follows, we examine the size of the swap book from the dealer's perspective.

V. Data and Summary Statistics

A. Description of the Data

We combine three main types of data for the analysis in this paper: swap positions, futures positions, and balance sheet/risk data. The final sample is monthly and spans the period from December 2007 to October 2015.

The swap data consists of end-of-month long and short positions held by dealers who received a special call from the CFTC to provide such data. In mid-2008, the Commission contacted 16 dealers known to have significant commodity index swap businesses and 13 other dealers having large commodity futures positions. The Commission also contacted 14 entities managing commodity index funds, including funds indexed to a single commodity. Respondents were required to provide position information related to commodity index transactions, starting from December 2007. The special call continued monthly until October 2015. The number of participants contacted varied over time as firms merged or entered/left the business.

Data were reported in notional terms and in the number of futures equivalent contracts (delta-adjusted). Dealers provided information on positions, broken down into the individual commodity exposure, resulting from commodity transactions including index swaps, single commodity swaps, and other products such as commodity index-linked notes or ETFs. Aggregated data on index investments were published by the CFTC in Commodity Futures Trading Commission (2008) and in a subsequent, periodical "Index

Investment Data" (IID) report. The aggregated data summed positions resulting from dealer index swaps and notes, as well as direct transactions in the futures market. The data has been used by researchers (e.g., Irwin and Sanders (2012)) to evaluate the effect of index investment on prices and volatility of commodity prices.

In this paper, we use the raw, firm-level data compiled for the IID report and focus our attention on the trading activity of individual dealers. Our primary measure of dealer positions is constructed from the single commodity swaps on WTI crude oil, which has not been previously reported publicly. Separately, we use the dealer-level WTI positions associated with index-linked swaps to measure the size of the dealer's index book. In both cases, we refer to the data as "swap data", although it includes other dealer transactions such as index-linked notes. In addition to using the size of the individual swap dealer's index book as a state variable, we use the size of non-swap index positions. These non-swap positions include direct futures market holdings by mutual funds, ETFs, or other funds. We also construct this variable using the raw data used to construct the IID report.

The futures data used in the paper comes from the daily futures and options on futures position data collected by the U.S. Commodity Futures Trading Commission as part of their Large Trader Reporting System. Data contain end-of-day long and short positions of different expirations of each contract per large trader, where a large trader is defined as having a position greater than some threshold number of contracts. Data also include the delta-adjusted options positions for each large trader.

The futures data we examine start from the end of 2007 and go until end of October 2015. We aggregate the net value of futures and delta-adjusted options in three contracts that have WTI crude oil underlying to gauge the WTI futures exposure of each dealer on a daily basis [4]. The futures data is then combined with the swaps data on certain

[4]The three contracts used in the analysis are NYMEX WTI crude oil contract, NYMEX WTI crude

observation dates we have the swaps data for to have a comprehensive data set.

The fundamental data for each dealer was collected through public sources: quarterly, semi-annual, and annual reports (as available), and investor presentations usually associated with earnings announcements. We collected the following balance sheet variables: Assets, Equity, Repo plus Short-Term Borrowing, Repo, Tier 1 Capital Ratio, and Physical Commodity holdings. We also collected Trading Value-at-Risk (VaR) figures for the aggregate trading portfolio, and its interest rate, credit, and commodity components, as available. The universe of firms includes entities that file under U.S. accounting practices and European practices, which required us to standardize some data for analysis. We follow standardizations employed by Bloomberg Markets for the balance sheet data. Because fundamentals vary in timing and frequency, we repeat variable values until the next observation is available.

VaR presentation varied over time across and within firms. Our target VaR measure is the 99%, one-day VaR, averaged over the trailing quarter. We used that measure when it was presented and used the best available proxies when it was not available. For example, we convert 95% VaR statistics to 99% by multiplying the reported value by the normal distribution function conversion factor of 1.41432, and we convert 10-day VaR statistics to 1-day statistics by dividing the reported value by the square root of 10. We used end-of-period values when average data was not reported, and we use data from multiple periods to compute quarterly VaR when required.

B. Summary Statistics

Table I presents descriptive information on the typical levels of major variables used in the analysis. We provide average values over the entire sample as well as over two subsamples. The subsamples each comprise approximately half of the total observa-

oil calendar contract, and ICE crude oil linked contract tions. Roughly speaking, the first subsample (December 2007 to December 2011) covers the financial crisis period and the rule-writing period, whereas the second subsample (January 2012 to October 2015) covers the rule implementation period and a period of substantially increased U.S. crude oil production due to new technology ("tight oil"). We briefly note some of the most interesting observations resulting from examination of the statistics. Panel A presents information on the aggregate dealer swap positions. As previously seen in Figure 1, the net WTI swap exposure of dealers due to their commodity index business became smaller during the sample, declining in magnitude by approximately one third (from 330,000 contracts to 230,000 contracts). At the same time, the aggregate WTI exposure from WTI-specific swaps increased by a similar magnitude of 130,000 contracts (increasing from 220,000 contracts to 350,000 contracts). Gross positions, which sum the absolute values of long and short positions, fell sharply from the first subsample to the second. Panel B displays information on WTI and market variables, highlighting that the second subsample generally featured increased U.S. crude oil production, which was associated with lower prices and lower volatility for crude oil, as well as more distress for producers, as measured by the Zmijewski z-score. There was was also a modest increase in commodity index investing activity not carried out via index swaps (i.e., through direct investment vehicles such as mutual funds). Finally, Panel C presents typical levels of the balance sheet and risk fundamentals for the universe of swap dealers. Broadly speaking, the statistics suggest that dealers were less levered, had more balance sheet equity and higher Tier 1 capital ratios, and had VaR levels that were roughly half as much during the second subsample, as compared to the first subsample. Further, dealers pursued less short-term borrowing, including repurchase transactions during the second subsample.

VI. Empirical Analysis

A. Risk Appetite Dynamics

One of the main findings from section IV is that the willingness of dealers to accommodate client hedging demand in the swap market, and their subsequent demand for hedge instruments in the futures market, depends crucially on risk appetite. We begin the empirical analysis by demonstrating the strong downward trend in dealer VaR over the sample, even after controlling for important state variables such as balance sheet equity and market volatility. Our interpretation of the results in table II is that the strong, residual, downward trend in VaR is striking empirical evidence that dealers fundamentally shifted their risk appetite over the sample period.

B. How Much Liquidity Provision to Customers?

In this section, we use panel regressions with dealer fixed effects to understand whether dealers with larger risk appetites have larger WTI swap books. Further, we test whether having a larger than average index swap book yields a larger WTI swap book. To answer these questions, we regress monthly net single commodity WTI swap positions on our proxies for dealer risk appetite, on individual dealer net index position and on other control variables as described below. Table III shows the results of our swap book size regression, which strongly confirm the theoretical predictions.

Each column in the table presents the results of a panel regression of dealer single commodity swap book size (net long position) explained by demand side variables (U.S. Crude Oil Production Forecast, Sector Z-Score for Oil Producers), market control variables (lagged price of one year ahead CL13 oil futures, the TED spread, WTI 3 month at-the money implied volatility index), supply side variables (equity for each dealer, net size of the dealer's WTI-index swap book, aggregate non-swap index investment in WTI futures), a time trend, and a proxy for dealer risk appetite. All variables are the same in each regression, except for the risk appetite proxies, which vary across regressions. T-statistics are in parentheses below coefficients; coefficients with absolute t-statistics larger than 1.96 are highlighted in gray.

The model predicts that increased production demand should lead to more hedging and larger swap books, and we find strong evidence to support this prediction. The production variable is the one-year ahead baseline forecast of U.S. crude oil production from the Energy Infomration Administration. The variable is highly significant, with t-statistics well above conventional significance levels in all specifications. We also control for producer risk aversion, proxied by producer distress. This is measured by the Zmijewski z-score of large producers in SIC code 1311, following the work of Acharya, Lochstoer and Ramorai (2013). The variable is signed correctly, and it is significant in two specifications.

We also control for the price of crude oil. We use the one-year ahead WTI crude oil futures price (denoted CL13) and lag it one month, and we generally find it is quite significant. We interpret this as a mechanical relation. Producers often sell a quantity of crude oil forward using collars (short calls and long puts). The data represent delta-adjusted futures equivalents, therefore, a change in crude price (holding the actual portfolios constant) would result in a change in the delta-adjusted position. If the price increases and the call options go into the money, this would increase the delta of the futures hedge.

Further, we find that an increase in a dealer's index investment book is associated with an increase in the size of the single commodity swap book. As the index position of investors increases, the dealer position gets shorter. A larger negative value multipled by a negative coefficient leads to a larger predicted single commodity swap book. However, we do not find a significant relation between other, direct index investment and the size of a dealer's single commodity swap book. If index investors take larger positions in mutual funds or ETFs indexed to commodity products, this does not appear to have a meaningful effect on the size of an individual dealer's swap book. We intend to explore this relation further to examine if there is an industry-wide impact of index investing.

We find that the risk appetite proxies are always highly related to the size of the dealer swap books. As dealers became more risk averse in the sample - whether measured by decreased leverage, decreased VaR to equity, less repo or short-term borrowing, or a higher Tier 1 capital ratio - the dealer single commodity swap book became smaller.

Finally, we also find that a trend variable remains highly signficant in most cases, even after controlling for the factors described above. The meaning of this finding is not immediately obvious, but it might suggest a role for the changing composition of WTI production in equilibrium hedging. As described in Commodity Futures Trading Commission (2018), some industry observers suggest that the shorter lifecycle of tight oil has led producers to shift some hedging from derivatives to operational hedging. We intend to explore this explanation in subsequent work.

C. Hedging Dynamics

We begin by providing descriptive information on dealer hedge portfolios. While it is suggested anecdotally that dealers prefer to hedge positions with the most liquid, near-dated futures contracts, there is little quantitative evidence to pin down the fact. in table IV, we break out dealer positions by tenor. Table IV shows net positions in Panel A and open positions in panel B. All the data is broken into 5 tenor buckets; 0 to 3 months bucket, 3 months to 12 months bucket, 12 months to 24 months bucket, 24 months to 36 months bucket, and longer than 36 months bucket. Additionally, the table shows dealers' futures and options positions separatley and together. Finally, data are presented for the pre-2012 (period 1) and post-2012 (period 2) periods separately, capturing the 2012 switch visible in figure 2.

We observe some simple patterns in Table IV. Starting with panel B, there is an overall drop in open positions held by dealers between period 1 and period 2. Just focusing on the change column, it is obvious that this drop is bigger in magnitude at the tail end of the tenor buckets. From panel A, we observe net futures and options positions to be positive in period 1 but they drop to a large negative value in period 2. Even with this drop in period 2, net exposure in the first tenor bucket is positive and positions in the rest of the tenor buckets are negative; a pattern that does not change across periods. Additionally, Comparing these numbers to the individual futures and options positions, we see that most of the changes are driven by futures, not by options.

We next provide analysis relating dealer hedge portolios, by tenor, to the swap book. We begin with regressions at the aggregate level, covering all 26 dealers in the sample. We regress change in dealers' futures, options and both futures and options positions on changes in index swap positions and changes in single commodity swaps. Motivated by our findings in table IV, we also run our regressions separately for the five different tenor buckets. More formally, we estimate Table V corraborates our earlier observations that hedging is mainly done with futures, not options. Additionally, when broken out into tenor buckets, hedging ratio coefficients are negative and significant for the shortest bucket, but either not significant or negative for longer tenor buckets for index swap position changes. For changes in single commodity swaps, hedging at the longer end of the tenor buckets seems to be statistically significant but coefficients are quite small.

Finally, we provide evidence on how individual dealers hedge their particular swap book. We find substantial variation across dealers, as do Naik and Yadav (2003*b*). Results are shown in VI. Panel A summarizes the coefficients of the baseline time series regression for each dealer $i = 1, ..., 26$:

$$\Delta F_{it} = \alpha_i + \beta \Delta S_{it}^{I} + \gamma \Delta S_{it}^{SCS} + \varepsilon_{it}, \tag{17}$$

where ΔS_{it}^{I} is the change in dealer i's net WTI swap exposure due to commodity index swaps, and Δ_{it}^{SCS} is the change in dealer i's net WTI swap exposure due to single commodity swaps on WTI. We present summary statistics for each coefficient. Overall, the hedging coefficients have the correct sign, and the magnitude of the coefficient is larger for index swaps (which have tenors more closely matching the futures hedge tenors) than for single commodity swaps (which have tenors much longer than the futures hedge). Coefficents are generally quite significant.

The results in Panel B allow the coefficents to change between the first and second subperiods. If dealer risk appetite declined during the second period and caused dealers to hedge the positions more tightly, as the model predicts, we would expect to see the coefficients cluster more tightly near a neutral hedge coefficient. The evidence is consistent with this prediction. Index hedge coefficients cluster much more tightly near a value of unity, and single commodity swaps also appear consistent with the prediction. Panel B summarizes the coefficients of the regression allowing the slopes to change for each of the 26 dealers:

$$\Delta F_{it} = \alpha_i + \beta_{i,1} \Delta S_{it}^{I} D_1 + \beta_{i,2} \Delta S_{it}^{I} D_2 + \gamma_{i,1}^{I} \Delta_{it}^{SCS} D_1 + \gamma_{i,2}^{SCS} \Delta_{it}^{SCS} D_2 + \varepsilon_{it},$$

where D_1 is a dummy variable taking the value 1 up to December 2011 and 0 afterwards and D_2 is a dummy variable taking the value 0 up to December 2011 and 1 afterwards.

Panel C reports the results of the cross-sectional regression of the change in dealer i's estimated slope coefficient between periods 1 and 2 on the estimated value for period

1. This regression provides a simple description of how an individual dealer's hedge coefficients varied over the subsamples. For dealers with index coefficients that were greater than approimately 0.90 in the first subsample, the regression predicts that the coefficient in the second subsample is lower and closer to 0.9. Similarly, dealers with low index hedge coefficients in the first subsample are predicted to have an index coefficient closer to 0.9 in the second subsample. For single commodity swap coefficients, dealer coefficients similarly present a story of reversion. Dealers with low values of the hedge coefficent are predicted to have increased values for the second subsample, and vice versa. We view these results as quite consistent with the model predictions of hedge ratios.

VII. Conclusion

Despite the size and importance of swap markets, relatively little systematic evidence is available on them. Dealers participate in the swap market and in listed derivatives markets and provide intermediation services to customers. However, the changes in the regulatory environment since the crisis have led to substantial changes in the business activities of dealers. End-users of derivatives have complained that competition in swap markets has declined, along with liquidity. We provide evidence that links the risk appetite of individual dealers to their intermediation business for customers. We use observable balance sheet measures to control for changes in risk appetite, and we use dealer-level data on their provision of swaps to clients and the futures hedges against those swaps. We find strong evidence that the decline in dealer risk appetite occurred alongside a decline in the provision of liquidity to customers. This link persists even after controlling for shifts in user demand for swaps. Further, consistent with a theoretical model of dealer activity, we find that dealers hedged these swaps more tightly as their risk appetite declined. That is, dealers were less willing to take on basis risk on behalf of customers when risk appetite declined. Dealers provided fewer swaps to clients and provided worse pricing on those swaps.

References

Acharya, Viral, Lars Lochstoer, and Tarun Ramorai. 2013. "Limits to Arbitrage and Hedging: Evidence from Commodity Markets." *Journal of Financial Economics*, 109(2): 441–465.

Adrian, Tobias, and Hyun Song Shin. 2010. "Liquidity and Leverage." *Journal of Financial Intermediation*, 19: 418–437.

Adrian, Tobias, Nina Boyarchenko, and Or Shachar. 2017. "Dealer Balance Sheets and Bond Liquidity Provision." *Journal of Monetary Economics*.

Begenau, Juliane, Monika Piazzesi, and Martin Schneider. 2015. "Banks' Risk Exposures."

Chernenko, Sergey, and Michael Faulkender. 2011. "The Two Sides of Derivatives Usage: Hedging and Speculating with Interest Rate Swaps." *Journal of Financial and Quantitative Analysis*, 46: 1727–54.

Commodity Futures Trading Commission. 2008. "Staff Report on Commodity Swap Dealers and Index Traders with Commission Recommendations."

Commodity Futures Trading Commission. 2018. "Impact of Tight Oil on NYMEX WTI Futures."

Etula, Erko. 2013. "Broker-Dealer Risk Appetite and Commodity Returns." *Journal of Financial Econometrics*, 11(3): 486–521.

Froot, Kenneth A., and Jeremy Stein. 1998. "Risk Management, Capital Budgeting and Capital Structure Policy for Financial Institutions: An Integrated Approach." *Journal of Financial Economics*, 47: 55–82.

Gârleanu, Nicolae, Lasse Heje Pesersen, and Allen M. Poteshman. 2009. "Demand-Based Option Pricing." *Review of Financial Studies*, 22(10): 4259–4299.

Irwin, Scott H., and Dwight R. Sanders. 2012. "Testing the Masters Hypothesis in Commodity Futures Markets." *Energy Economics*, 34(1): 256–269.

Mixon, Scott. 2018. "Research Issues in Commodities and Commodity Derivatives." In *Commodities: Markets, Performance, and Strategies.* , ed. H. Kent Baker, Greg Filbeck and Jeffrey H. Harris, 534–554. New York, NY:Oxford University Press.

Mixon, Scott, Esen Onur, and Lynn Riggs. 2018. "Integrating Swaps and Futures: A New Direction for Commodity Research." *Journal of Commodity Markets*, 10(1): 3–21.

Naik, Narayan Y., and Pradeep K. Yadav. 2003*a*. "Do Dealer Firms Manage Inventory on a Stock-by-Stock or a Portfolio Basis?" *Journal of Financial Economics*, 69(2): 325–353.

Naik, Narayan Y., and Pradeep Yadav. 2003*b*. "Risk Management with Derivatives by Dealers and Market Quality in Government Bond Markets." *Journal of Finance*, 58(5): 1873–1904.

Stulz, René. 1996. "Rethinking Risk Management." *Journal of Applied Corporate Finance*, 9: 8–24.

Viswanathan, S., and James D. Wang. 2004. "Inter-Dealer Trading in Financial Markets." *Journal of Business*, 77(4): 987–1040.

Vogler, Karl-Hubert. 1997. "Risk Allocation and Interdealer Trading." *Review of Financial Studies*, 41: 1615–1634.

Table I: Summary Statistics - Average Levels

Aggregate Dealer swap positions are displayed in Panel A; values are in thousands of delta-adjusted futures equivalents. Panel B contains WTI average prices, the 3 month at-the-money implied volatility, EIA baseline forecast for 1 year ahead U.S. crude production (millions of barrels per year), the sector average Zmijewski (1984) default score for SIC code 1311, and aggregate non-swap commodity index investing for WTI (contracts). Panel C presents average values for dealer fundamental variables; VaR levels, Equity, Short-Term Borrowing, and Repo values are in millions of US dollars. Data are monthly from December 2007 to October 2015.

Panel A: Aggregate Dealer Swap Positions (Thousands of Contracts)

Sample Period	**Net Index**	**Net WTI**	**Net Swap**	**Gross Index**	**Gross WTI**	**Gross Swap**
2007/12 - 2011/12	-326.7	222.3	-104.4	622.3	5,964.9	6,587.1
2012/1 - 2015/10	-231.8	346.5	114.7	474.6	2,040.0	2,514.7
Full Sample	-279.8	283.7	3.9	549.2	4,023.6	4,572.8

Panel B: WTI and Market Variables

Sample Period	**CL13 Price**	**Implied Vol**	**Forecast Production**	**Producer Z-Score**	**Non-Swap Index**
2007/12 - 2011/12	87.40	39.66	1,977	-2.73	118,067
2012/1 - 2015/10	84.34	27.35	2,940	-2.37	134,721
Full Sample	85.88	33.57	2,453	-2.55	126,305

Panel C: Individual Dealer Fundamentals

Sample Period	**Leverage**	**VaR**	**Commodity VaR**	**Equity**	**ST Borrow**	**Repo**	**Tier 1 Ratio**
2007/12 - 2011/12	23.94	97.41	16.52	62,146	281,127	108,321	11.5
2012/1 - 2015/10	18.98	45.07	8.55	77,311	209,051	88,842	13.7
Full Sample	21.49	71.52	12.58	69,647	245,476	98,686	12.6

Table II: Dealer Trading VaR Explained by Trend and Control Variables

This table displays a panel regression of individual dealer trading VaR on a trend, equity level of that dealer, and implied volatility indices. Regression results are also displayed for specifications with the lagged VaR statistic included. Dealer fixed effects are included in all specifications. T-statistics are in parentheses below coefficients; coefficients with absolute t-statistics larger than 1.96 are highlighted in gray. Data are monthly from December 2007 to October 2015.

	Dependent Variable					
	Aggregate VaR		Interest Rate VaR		Commodity VaR	
Trend	-1.017	-0.076	-0.690	-0.025	-0.150	-0.017
	(-22.22)	(-3.43)	(-15.59)	(-1.25)	(-17.75)	(-4.77)
Equity	0.254	0.018	-0.123	-0.020	-0.055	-0.004
	(5.03)	(0.86)	(-2.59)	(-1.06)	(-7.08)	(-1.29)
VIX	0.613	0.184				
	(4.25)	(2.92)				
MOVE			0.072	0.054		
			(1.74)	(2.93)		
WTI Implied Vol					0.007	-0.010
					(0.40)	(-1.58)
Lagged VaR		0.916		0.924		0.910
		(55.51)		(44.40)		(45.71)
Dealer Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Adj R^2 (%)	74.8	96.4	75.7	96.9	82.9	97.1
# obs	1719	1700	1572	1553	1534	1516
# Dealers	19	19	19	19	18	18

Table III: Relating Risk Appetite to Single Commodity Swap Book Size

The table presents estimated coefficients for a panel regression of the net size of dealer WTI single commodity swap books on demand side variables (U.S. Crude Oil Production Forecast, Sector Z-Score for Oil Producers), market control variables (lagged price of one year ahead CL13 oil futures, the TED spread, WTI 3 month at-the money implied volatility index), supply side variables (equity for each dealer, net size of the dealer's WTI-index swap book, aggregate non-swap index investment in WTI futures), a time trend, and a proxy for dealer risk appetite. T-statistics are in parentheses below coefficients; coefficients with absolute t-statistics larger than 1.96 are highlighted in gray. Data are monthly and span the period December 2007 to October 2015.

Independent Variables	**Risk Appetite Proxy**					
	Assets/ Equity	VaR/ Equity	Commodity VaR / Equity	ST Borrow/ Equity	Repo/ Equity	Tier 1 Ratio
Production Forecast	16.12	17.27	11.47	14.12	17.76	12.12
	(4.30)	(4.53)	(3.02)	(3.79)	(4.24)	(3.19)
Producer Z-Score	4,067.78	3,191.95	3,341.54	4,717.35	5,600.89	5,484.44
	(1.60)	(1.24)	(1.22)	(1.86)	(2.05)	(2.10)
CL13 Lagged Price	128.81	136.94	53.26	140.66	147.04	202.82
	(2.19)	(2.25)	(0.81)	(2.34)	(2.19)	(3.24)
TED Spread	-110.99	1041.70	-1257.53	-395.27	247.88	-1925.92
	(-0.07)	(0.64)	(-0.74)	(-0.25)	(0.14)	(-1.23)
WTI Implied Vol	-123.34	-191.30	-143.43	-59.30	-81.75	95.46
	(-1.28)	(-1.88)	(-1.38)	(-0.64)	(-0.75)	(1.14)
Equity	0.26	0.25	0.30	0.27	0.28	0.26
	(12.33)	(11.91)	(10.05)	(13.12)	(10.63)	(8.24)
Net Index Swap Book	-0.45	-0.33	-0.12	-0.45	-0.44	-0.34
	(-5.62)	(-4.09)	(-1.11)	(-5.63)	(-5.28)	(-4.40)
Non-Swap Index	0.00	0.01	-0.01	-0.01	-0.01	-0.04
	(-0.04)	(0.30)	(-0.29)	(-0.22)	(-0.27)	(-1.03)
Trend	-230.41	-195.25	-125.30	-185.53	-288.47	-121.27
	(-2.39)	(-2.05)	(-1.29)	(-1.93)	(-2.58)	(-1.24)
Risk Appetite Proxy	508.08	5.80	44.95	1.42	2.94	-479.89
	(6.00)	(7.72)	(6.15)	(5.17)	(4.27)	(-2.41)
Adj. R^2 (%)	68.5	68.8	71.2	68.4	67.4	71.8
# obs	1611	1550	1386	1599	1447	1529
# Dealers	20	19	18	20	18	19

Table IV: Average Futures Positions of Dealers, by Tenor

This table displays futures and delta-adjusted option positions for the 26 dealers in the sample. Values are in thousands of futures equivalent contracts and include the NYMEX WTI contract, ICE WTI contract, and NYMEX WTI Calendar swap contract. Data are monthly from December 2007 to October 2015.

Panel A: Net Positions - thousands of futures-equivalent contracts

Instrument	**Sample Period**	**0 - 3m**	**3m - 1yr**	**1yr - 2yr**	**2yr - 3yr**	**3yr+**	**Total**
Futures + Options	2007/12 - 2011/12	169.8	-37.7	-46.1	-17.3	-15.7	53.0
Futures + Options	2012/1 - 2015/10	52.7	-165.3	-113.7	-24.5	-5.7	-256.5
Futures + Options	Full Sample	113.1	-99.5	-78.8	-20.8	-10.8	-96.9
Futures	2007/12 - 2011/12	184.9	-13.8	-38.8	-17.4	-19.1	95.8
Futures	2012/1 - 2015/10	62.0	-151.0	-108.3	-21.7	-2.5	-221.5
Futures	Full Sample	125.4	-80.2	-72.5	-19.5	-11.1	-57.9
Options	2007/12 - 2011/12	-15.1	-23.9	-7.3	0.1	3.5	-42.7
Options	2012/1 - 2015/10	-9.3	-14.3	-5.4	-2.8	-3.2	-35.0
Options	Full Sample	-12.3	-19.3	-6.4	-1.3	0.2	-39.0

Panel B: Open Positions

Instrument	**Sample Period**	**0 - 3m**	**3m - 1yr**	**1yr - 2yr**	**2yr - 3yr**	**3yr+**	**Total**
Futures + Options	2007/12 - 2011/12	700.9	855.5	510.0	261.8	243.0	2,571.1
Futures + Options	2012/1 - 2015/10	528.5	705.7	404.3	140.3	100.7	1,879.6
Futures + Options	Full Sample	617.4	783.0	458.8	203.0	174.1	2,236.3
Change from Period 1 to 2		-24.6%	-17.5%	-20.7%	-46.4%	-58.6%	-26.9%

Table V: Hedging Regressions, by Tenor

The table presents coefficients for the aggregate time series regression

$$\Delta F_t^M = \alpha + \beta \Delta S_t^I + \gamma \Delta S_t^{SCS} + \varepsilon_t,$$

where ΔF_t^M is the change in net positions in futures portfolio M for Dealers, ΔS_t^I is the change in net WTI swap exposure due to commodity index swaps, and Δ_t^{SCS} is the change in net WTI swap exposure due to single commodity swaps on WTI. T-statistics are in parentheses below coefficients; coefficients with absolute t-statistics larger than 1.96 are highlighted in gray. All regressions incorporate the same aggregate swap exposures as independent variables, and the cases $M = 1, ..., 8$ reflect Dealer futures portfolios for varying instruments and tenors as the dependent variable. Regressions are estimated using positions for all tenors in Cases 1,2, and 3; regressions 4-8 use only contracts expiring during the timeframe specified for that regression. Case 1 and Cases 4-8 incorporate futures and delta-adjusted options, while Cases 2 and 3 break out futures and options, respectively. Data are monthly and span the period December 2007 to October 2015.

	Dependent Variable		Independent Variables		Adj. R^2 (%)
Case	**Expiries**	**Instrument**	**Δ Index Swaps**	**Δ WTI Swaps**	
1	All	Futures + Options	-0.96	-0.44	43.42
			(-6.40)	(-5.15)	
2	All	Futures only	-0.90	-0.40	36.46
			(-4.38)	(-4.54)	
3	All	Options only	-0.06	-0.04	1.55
			(-0.63)	(-1.92)	
4	0 - 3m	Futures + Options	-0.78	-0.15	32.62
			(-5.53)	(-3.38)	
5	3m - 1yr	Futures + Options	-0.26	-0.07	6.69
			(-2.19)	(-1.61)	
6	1yr - 2yr	Futures + Options	0.00	-0.16	15.26
			(-0.03)	(-3.59)	
7	2yr - 3yr	Futures + Options	0.05	-0.04	2.09
			(1.03)	(-2.48)	
8	3 yr+	Futures + Options	0.02	-0.02	1.21
			(0.69)	(-2.22)	

Table VI: Hedging Regression Coefficients for Individual Dealers

Panel A summarizes the coefficients of the baseline time series regression for each dealer $i = 1, ..., 26$:

$$\Delta F_{it} = \alpha_i + \beta \Delta S_{it}^I + \gamma \Delta S_{it}^{SCS} + \varepsilon_{it},$$

where ΔS_{it}^I is the change in dealer i's net WTI swap exposure due to commodity index swaps, and Δ_{it}^{SCS} is the change in dealer i's net WTI swap exposure due to single commodity swaps on WTI.

Panel B summarizes the coefficients of the regression allowing the slopes to change for each of the 26 dealers:

$$\Delta F_{it} = \alpha_i + \beta_{i,1} \Delta S_{it}^I D_1 + \beta_{i,2} \Delta S_{it}^I D_2 + \gamma_{i,1} \Delta_{it}^{SCS} D_1 + \gamma_{i,2} \Delta_{it}^{SCS} D_2 + \varepsilon_{it},$$

where D_1 is a dummy variable taking the value 1 up to December 2011 and 0 afterwards and D_2 is a dummy variable taking the value 0 up to December 2011 and 1 afterwards.

Panel C reports the results of the cross-sectional regression of the change in dealer i's estimated slope coefficient between periods 1 and 2 on the estimated value for period 1. Data are monthly and span the period December 2007 to October 2015.

Panel A: Baseline Hedging Regressions

		Δ Index Swaps		Δ WTI Swaps		Adj. R^2 (%)
Percentile	Intercept	Coefficient	T-Stat	Coefficient	T-Stat	
25%	-113.14	-0.98	-4.98	-0.83	-6.59	15.6
50%	-10.76	-0.78	-2.63	-0.52	-4.47	31.2
75%	67.13	-0.37	-1.40	-0.23	-1.65	50.0

Panel B: Hedging Regressions with Subsamples

		Δ Index Swaps		Δ WTI Swaps		Adj. R^2 (%)
Percentile	Intercept	Period 1 Coefficient	Period 2 Coefficient	Period 1 Coefficient	Period 2 Coefficient	
25%	-113.61	-1.00	-1.01	-0.83	-0.93	20.5
50%	-12.65	-0.82	-0.97	-0.48	-0.68	35.7
75%	94.44	-0.29	-0.74	-0.15	-0.28	51.8
# Dealers	26	22	21	21	19	

Panel C: Cross-Sectional Regression on Period 1 Slope Coefficients

Dependent Variable	Intercept	T-stat	Slope	T-stat	R^2 (%)
$\beta_{i,2}$ - $\beta_{i,1}$	-1.00	-4.98	-1.10	-5.68	65.5
$\gamma_{i,2}$ - $\gamma_{i,1}$	-0.43	-2.54	-0.76	-2.85	33.6

Figure 1: Dealer Net Positions in WTI Swaps and Futures, 2007-2015

The figure displays the net WTI swap exposure and net WTI futures and options positions, both aggregated across the 26 Dealers in the sample. Exposures are measured in delta-adjusted futures equivalent contracts. Swap exposure includes both implied WTI exposure via commodity index swaps and direct WTI exposure via single commodity swaps. Futures positions include the NYMEX WTI contract, ICE WTI contract, and NYMEX WTI Calendar swap contract. Data are monthly from December 2007 to October 2015.

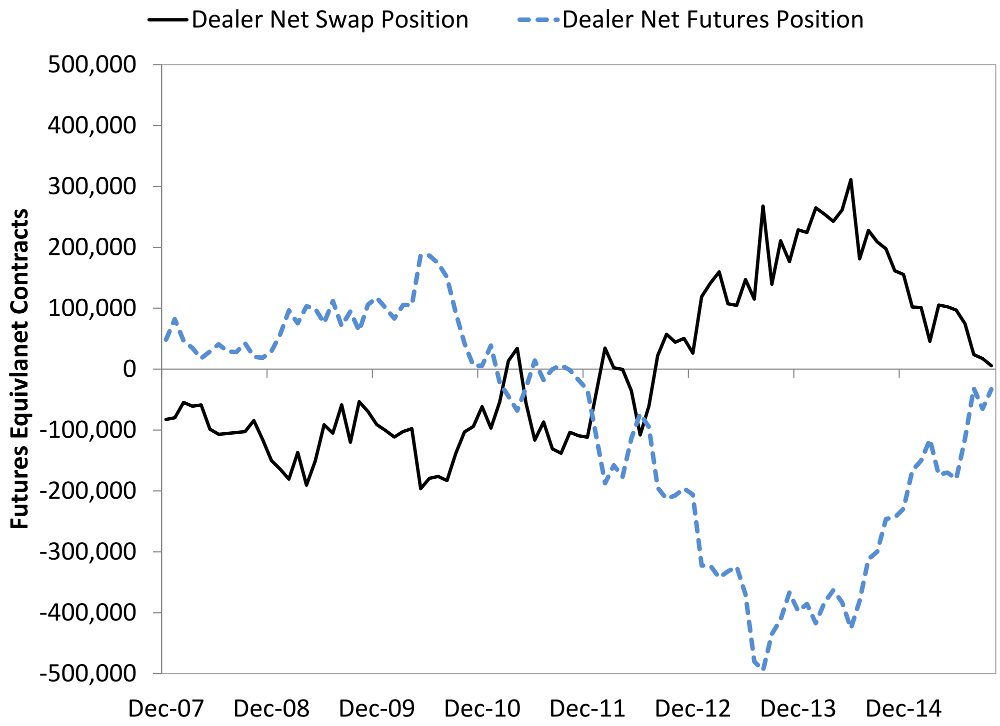


Figure 2: Dealer WTI Exposure due to Index and Single Commodity Swaps

The figure displays dealer net WTI exposure due to index swaps, net WTI exposure due to WTI single commodity swaps, and net WTI swap exposure. Values are aggregated across the 26 Dealers in the sample. Data are monthly from December 2007 to October 2015.

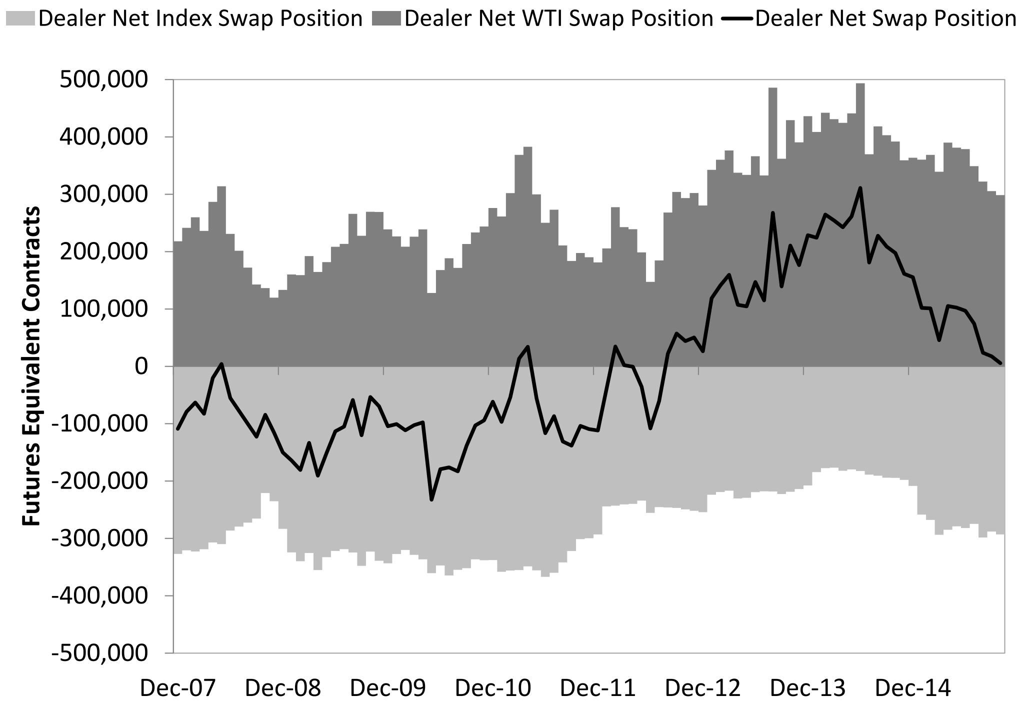
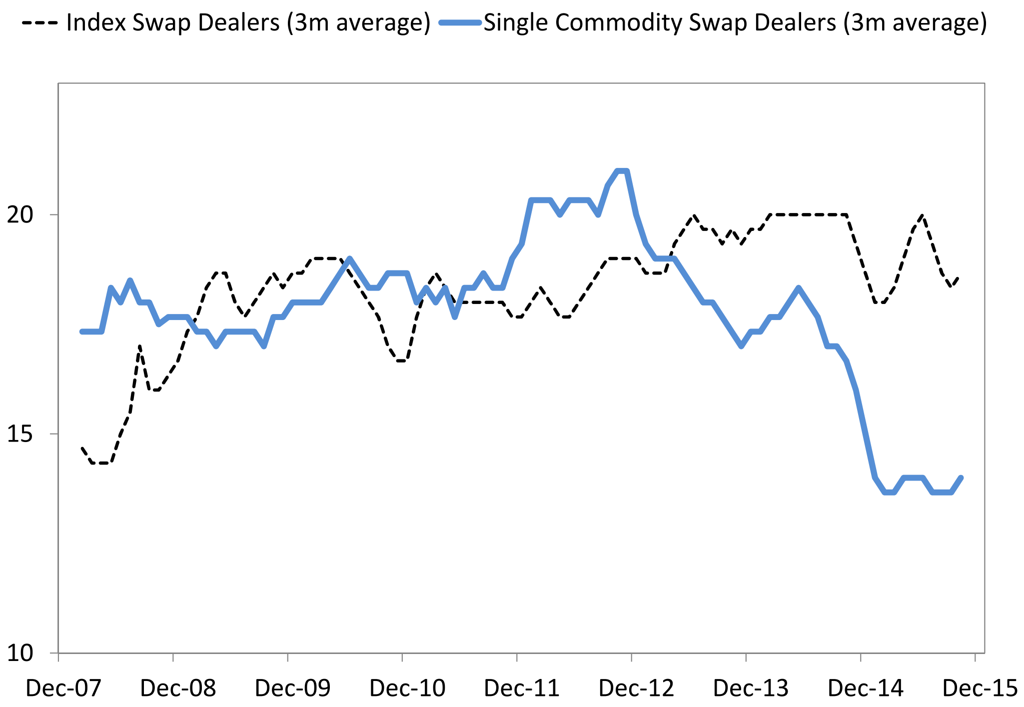


Figure 3: Number of Swap Dealers

The figure displays three-month moving averages of the count of swap dealers reporting non-zero net positions in commodity index or WTI swaps. Dealers are excluded from the count if they hold net swap positions less than 100 contracts.

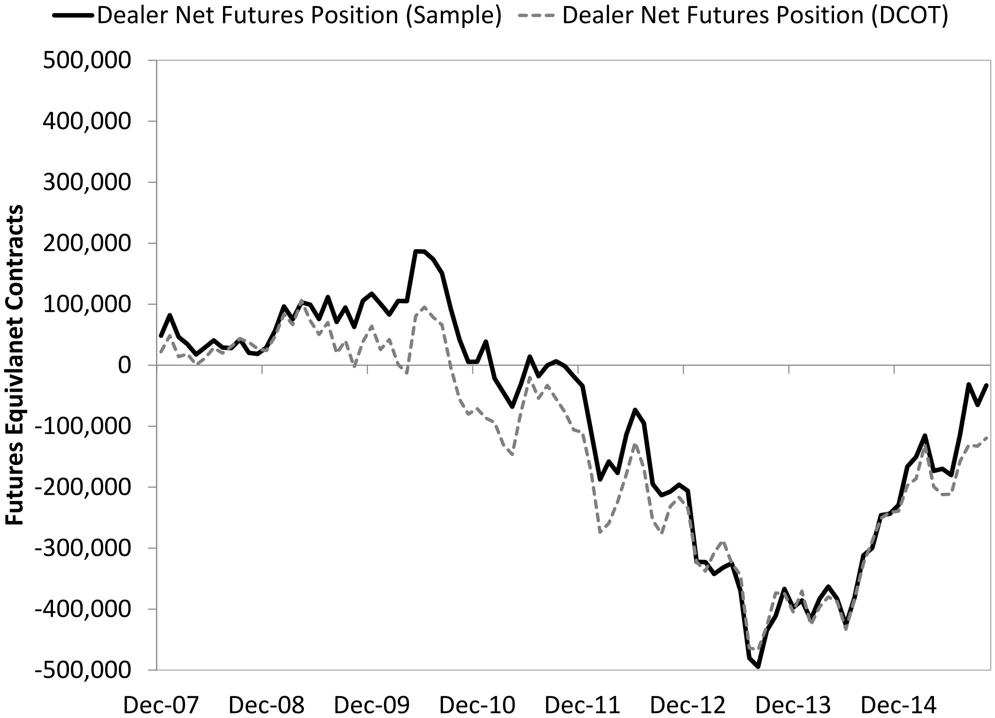


A. Appendix

Figure A1 compares the net futures position of the dealers in our sample with the dealers' net positions from the publicly available DCOT dataset.

Figure A1: Futures Positions of Dealers–Sample Data vs. DCOT

The figure displays both the net futures positions of Dealers utilized in this paper, compared with the net "Swap Dealer" futures and option position in NYMEX WTI futures from the CFTC's publicly available Disaggregated Committments of Traders (DCOT) report.



Hedging Input Commodity Price Risk: An Equilibrium View

Hamed Ghoddusi, Sheridan Titman, Stathis Tompaidis [1,2]

Abstract

We present a model of a commodity processing chain with endogenously determined input and output prices and characterize the effectiveness of hedging policies that employ forward contracts on the price of the input. The model illustrates that the variance minimizing hedge ratio depends on operational characteristics, such as the convexity of the production function, and on economic variables – the elasticity of supply and demand and the relative magnitude of input and output risks; i.e., the size of supply and demand shocks. We find that the optimal hedge ratio can change over time as capacity utilization in the industry changes. To gauge the quantitative importance of our implications we estimate a parametrized version of our model for the crude oil to refined products supply chain using the simulated method of moments.

Keywords: Supply Chain, Spread Dynamics, Optimal Hedging, Simulated Method of Moments (SMM)

[1] Ghoddusi is with the School of Business, Stevens Institute of Technology, `hghoddus@stevens.edu`. Titman is with the McCombs School of Business, University of Texas at Austin, Department of Finance, `Sheridan.Titman@mccombs.utexas.edu`. Tompaidis is with the McCombs School of Business, University of Texas at Austin, Department of Information, Risk and Operations Management and Department of Finance, `Stathis.Tompaidis@mccombs.utexas.edu`.

[2] We thank participants at seminars at HEC Montreal, Oxford Institute for Energy Studies, UT Austin, UC Berkeley, Humboldt University, Rutgers, Alberta Finance Institute Conference, BI Norwegian Business School, Athens University of Economics and Business, Sharif University of Technology, UC Davis, HEC Paris, University of Vienna, UQAM Montreal, the conference on Commodities and Financialization at the Institute for Pure and Applied Mathematics, the 2012 INFORMS meeting, the Commodity Futures Trading Commission, and the Supply Chain Finance and Risk Management Workshop for comments. All remaining errors are ours.

I. Introduction

A typical commodity processing chain consists of three components: an upstream primary input – such as crude oil – a downstream output product – for example refined products like gasoline and heating oil – and a capital asset that turns the raw commodity into the processed commodity – for example a refinery. Commodity supply chains are ubiquitous – smelters transform bauxite into aluminum; power plants convert natural gas to electricity; tomato processing plants transform tomatoes into ketchup, airlines transform jet fuel into air travel; etc.

Hedging input (and output) commodity prices is a major issue for processors. Power plants are concerned with the rising cost of natural gas; airlines use variety of methods to mitigate jet fuel price risk; and food processors try to protect themselves against agriculture product price shocks. Among various risk-management strategies, financial hedging (e.g., the use of futures, options, and swaps) is the most popular one in practice. The existing literature typically treats the price of input and output commodities as exogenous variables for the commodity processing industry. However, the equilibrium prices of input and output commodities are formed as a result of endogenous decisions of both processors and consumers.

This paper develops an equilibrium model of prices and quantities in a commodity processing chain with uncertainties regarding supply and demand conditions. The model describes how characteristics of the input and output markets, such as supply and demand elasticities, and the degree of competition, properties of the production function, e.g., its convexity, and the relative magnitude of supply and demand shocks, influence input and output prices, the spread between the two, and optimal hedging policies.

We use the model to explore the endogenous dynamics of profitability and the effectiveness of hedging in a market with financial contracts only on inputs or output but not on both. For example, airlines can hedge the cost of jet fuel but not their revenues; food processors can hedge commodity inputs (e.g. coffee beans, corn, wheat, even eggs) but typically cannot hedge their outputs (e.g. instant coffee, bread, corn syrup, sauces). The opposite is also true for certain firms: futures contracts exist for base metals such as copper, zinc, and aluminum (the output of a smelting unit), but not for metal ores and chemicals used as inputs and producers of major crops can hedge their output price risks but not their inputs (e.g. water, labor, fertilizers or pesticides). For the application that we study, the oil refinery business, there are in fact forward contracts for both the input (oil) and some of the outputs, e.g., heating oil and gasoline. However, although the market for oil is quite liquid, the derivative markets for refined products is very illiquid for contracts exceeding two years.

The conventional wisdom is that the price risk associated with inputs can be hedged by buying the inputs forward. For example, a number of airlines make forward purchases of either fuel oil or oil to hedge their fuel price exposure. As our model illustrates, this view is largely based on the case where input price changes arise because of exogenous supply shocks; e.g., a rise in oil prices due to a conflict in the Middle East. When this is the case, an increase in the cost of production caused by a shock to the supply of the input decreases the spread between input and output prices and hence, the profitability of the producer. Our model illustrates that if supply shocks are the only source of uncertainty, the conventional view of hedging holds and the variance-minimizing hedge ratio increases when the convexity of the production function increases, and decreases when the elasticity of the demand for the output increases.

An analysis of the oil to refined petroleum products supply chain reveals that this conventional wisdom does not always hold. Figure 1 plots the correlation between the spread between refined petroleum products and oil prices (the crack spread) from 1985 to 2017. The correlations plotted at each point in time are calculated from the following 60 months. As the figure illustrates, the correlations start out negative, which is consistent with the conventional wisdom, and then become positive. In recent years the correlations are close to zero, indicating that producers could not have benefited from hedging input risk.

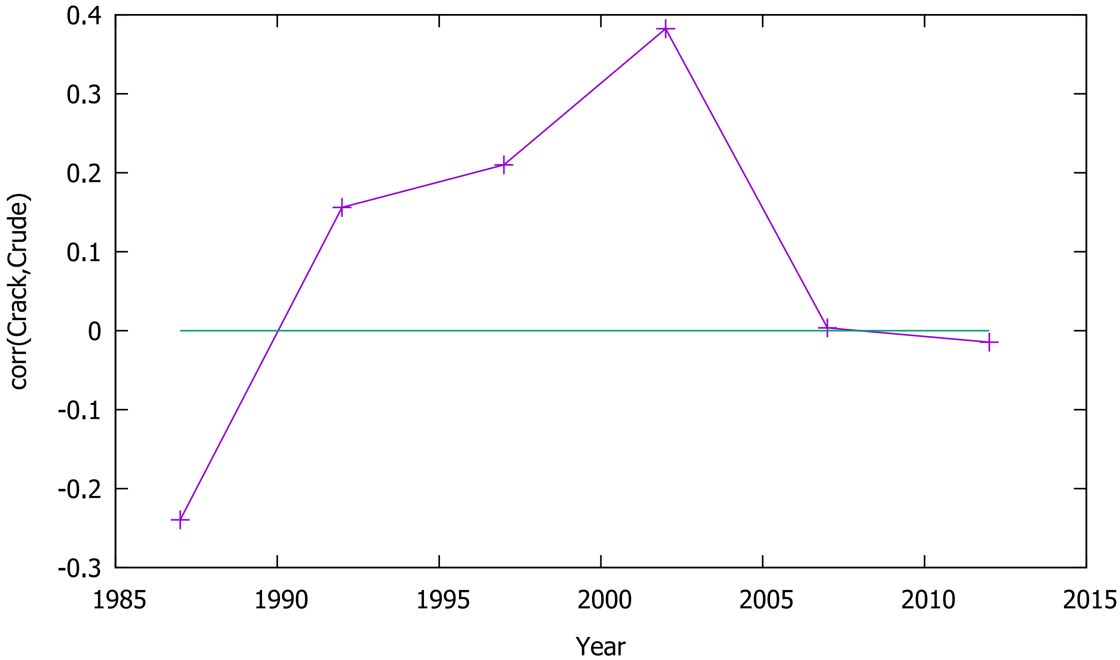


Figure 1: Correlation between Crack Spread and Crude Oil. The correlations are calculated using monthly data. Each correlation is calculated over a 60-month, non-overlapping, window. Correlations calculated between January 1987 and December 1991 are shown under the label for 1987; correlations calculated between January 1992 and December 1996 under the label for 1992; etc.

As our model illustrates, the conventional wisdom may have failed in this example because uncertainty comes from demand shocks as well as supply shocks. When the input cost increases because of a demand shock, e.g., because of a booming economy, the output price can move more than the input price, which means that spreads and input prices move in the same direction. As a result, input prices and the processing margin move together, which means that hedging demand shocks entails selling rather than buying the input forward. The optimal hedge ratio again depends on the convexity of the production cost function, but it is now the elasticity of the supply curve rather than the elasticity of the demand curve that determines the hedge ratio.

In the general case with both demand and supply shocks the correlation between profits, proxied by the spread, and input prices may be low, as illustrated in Figure 1, and it is generally not possible to perfectly hedge with just one financial instrument. However, the intuition developed in the case with a single source of uncertainty still holds. We show that increased supply uncertainty increases the amount of the input that a variance-minimizing hedger will purchase forward, while increased demand uncertainty reduces the magnitude of the forward purchase. The correlation between the spreads and the input prices, and thus the variance-minimizing hedge ratio, is still determined by the interaction between the convexity of the production function and the elasticities of the supply and demand functions.

We derive these results in a simple model where the relevant functions; i.e., the marginal production cost and the supply and demand functions, are all linear. This model can be solved in closed form and conveys our basic intuition. However, to a large extent, the questions that we ask are quantitative: what are the variance-minimizing hedge ratios, and do we expect them to change over time? How effective is hedging in plausible scenarios, i.e., to what extent can hedging reduce the variance of the producer's cash flows? Getting even rough answers to these questions requires a more realistic model that must be solved numerically.

Our more realistic model assumes demand and supply functions that are tailored for the crude oil to refined products supply chain. The main difference between our simple linear model and the model that we take to the data is that the estimated model allows for a more flexible production function that becomes more convex as utilization rates increase and we account for the possibility that production is subject to random production costs, caused for example by hurricanes or fluctuations in the price of natural gas. Using the simulated method of moments, we estimate the parameters of the model, i.e., the convexity of the production function, the volatility of the exogenous demand and supply factors, the demand elasticity of refined products, and the supply elasticity of crude oil.

Based on the estimates of these parameters, we run simulations that allow us to address issues that relate to hedging and its effectiveness. First, we consider the average percentage of the input quantity that is purchased forward in the variance minimizing hedge, and how this quantity changes as capacity utilization changes. We then compare the effectiveness of a dynamic hedging strategy that accounts for the changing levels of capacity with a strategy that simply holds the hedge ratio constant. Finally, we provide a more general analysis of hedging effectiveness, i.e., to what extent can firms reduce the variation of their cash flows by hedging input costs.

We find that in most of the scenarios we consider, hedging reduces cash flow variation by only a modest amount. For our base case, the optimal hedge ratio involves selling – rather than buying – oil forward, and hedging reduces the variance of profits by only 10%. There are two reasons why hedging looks to be ineffective. The first reason is that the correlation between oil prices and the crack spread is quite week because of the confounding effects of supply and demand shocks. Indeed, the correlation between oil prices and the crack spread is weakly positive for most of our sample, indicating that demand shocks are more important than supply shocks. The second reason is that matching the observed volatility of crack spreads requires relatively volatile noise in the portion of processing costs that cannot be hedged.

In the estimated base case, most of the systematic uncertainty comes from demand shocks. Our simulations show that in this case, if the exogenous sources of uncertainty can be eliminated, hedging can reduce cash flow variability by over 80%. However, in our comparative statics we show that when the magnitudes of supply and demand shocks have similar magnitudes, hedging effectiveness is quite low, even when the exogenous sources of uncertainty are eliminated.

The model also highlights the potential advantage of dynamic versus static hedging. A major difference between conventional statistical approaches to optimal hedging versus our structural approach is that the former is backward-looking: hedge ratios are estimated based on past values of realized prices. In contrast, our structural model illustrates how one can incorporate forward-looking information to dynamically adjust the optimal hedge. Example of forward-looking information that can be used include the implied volatility of options on input and output commodity prices, changes in the production function (due to new investments, technology improvements, and expected physical shocks), and the level of future supply and demand.

Using our calibrated model with exogenous shocks to the production function, we show in a simulation that improvement is modest in the base case. Both static and dynamic annual hedging deliver the same level of effectiveness – approximately 10%. The improvement is significantly better when the magnitude of exogenous shocks to the production function is small. The hedging effectiveness of a dynamically adjusted hedging strategy improves by up to 84%, while that of a static strategy by up to 76%. If structural parameters were to change; e.g., the variance of supply or demand shocks, then the benefit of dynamically adjusting the hedging strategy can be potentially even larger.

In summary, we offer the following novel contributions to the literature. First, we characterize the behavior of optimal hedging policies when there are shocks to both the supply of inputs and the demand for the downstream product. As we show, the conventional intuition that is based on shocks that only effect input prices often fail to hold in this equilibrium setting. Second, we offer a framework to highlight the drivers of time variation in dynamic hedge ratios and to incorporate forward-looking information in optimal hedging decisions. Finally, we use a novel calibration approach in the context of hedging and apply it to the case of the refinery industry to provide quantitative insights regarding the value of static versus dynamic hedging.

The rest of the paper is organized as the following. Section II provides a brief review of the relevant literature. We introduce the basic model in Section III. Section IV describes the theoretical results of the model. The quantitative exercise is introduced in Section VI.

II. Literature Review

Our model is most closely related to models developed by Hirshleifer (1988a), Hirshleifer (1988b), and Hirshleifer (1989), which explore the determinants of equilibrium spot and futures prices of a processed commodity. Hirshleifer (1988a) considers supply shocks and shows that the optimal hedge is long the input, while Hirshleifer (1988b) considers the case of demand shocks and shows that the optimal hedge is long the output. The papers focus on the difference between spot and futures prices, and the role played by transaction costs. Hirshleifer (1989) considers how the risk premium between the spot and the futures price varies with stock market variability, limited participation in the futures market, and the elasticity of supply and demand. Beyond the difference in focus, our framework differs in that we consider supply and demand shocks jointly, explore properties of the production function, and consider both competitive and monopolistic markets. We also characterize how the relative size of demand and supply shocks determine the optimal hedge ratio and hedging effectiveness, and derive results regarding the bounds on the ratio of the volatility of the input to the volatility of the output prices.

Carter et al. (2017) provide a recent review of the literature on commodity risk management. Kamara (1993) studies the optimal production and hedging decisions of the owner of a capital asset that can adjust its production plans. The paper shows that production decisions can be independent of risk preferences in the presence of futures contracts. Bessembinder and Lemmon (2002) consider the case of the electricity market and build an equilibrium model for electricity spot and forward prices. Unlike our paper, which focuses on the relationship between input and output prices, the focus of Bessembinder and Lemmon (2002) is on the spot and forward output price and how they are influenced by expected demand and uncertainty about shocks to demand. Another difference is that Bessembinder and Lemmon (2002) do not consider the feedback between a shock to the electricity demand and the price of the input; i.e., they assume that an increased demand for electricity does not affect the price of natural gas. Bessembinder and Lemmon (2002) show that the skewness in output prices is driven from the convexity of the production function but do not study the correlation between input and output prices and their spread, or hedging and its effectiveness.

Casassus et al. (2012) consider the possibility that the inputs and outputs can be stored and show that the ability to store commodities can influence correlations of prices and spreads. While storage is very important for understanding daily fluctuations of input and output prices, especially when demand or supply are seasonal, we expect it to be less important for understanding long term dynamics. For the case of the crude oil, storage accounts for approximately 3 months of consumption. We abstract away from storage and estimate our model using annual data.

Several papers study the behavior of the refinery industry. Wu and Chen (2010) apply a dynamic model of inventory and production to the refined products market. They also consider a market facing shocks to input and output. Dong et al. (2014) model the value of flexibility in the production function of the refinery industry. The focus of this literature is on the optimal level of endogenous production and inventory and its effect on price dynamics.

In contrast, our focus is on hedging and how it depends on supply and demand shocks and on the characteristics of the cost of production.

Our paper is also related to the literature on the dynamics of spreads and the valuation of spread options on a pair of commodities. Carmona and Durrleman (2003) introduce reduced-form pricing methods for spread options. Secomandi (2010) considers the effect of capacity constraints in natural gas pipelines. Our paper is related to this literature but, rather than a reduced-form approach, we build a micro-founded model to understand the dynamics of spreads.

III. Model

In the baseline model we consider a competitive industry which uses a capital asset to convert one unit of input to a unit of output. We consider the optimal production and profit dynamics of a *representative* owner of the capital asset; i.e., an owner of a continuum of small production units with different levels of efficiency; i.e., processing costs. We assume that the production function of the representative owner is proportional to the production function of the entire industry.

In addition to the production function, the primitives of our model are supply and demand functions of input and output. In the baseline model, we assume that the supply and demand functions are linear with respect to quantities, and that the production cost is quadratic with respect to quantity; i.e., the marginal cost of production is linear. We assume that there are no adjustment costs to production; that production is chosen once the levels of supply and demand have been observed and it is instantaneous; and that storage is not possible. These assumptions allow us to derive closed-form solutions – given supply and demand shocks the prices of the input and the output are determined endogenously from the elasticities of supply and demand and the convexity of the production function.

Table 1 summarizes the variables of our baseline model.

Symbol	Definition	Remark
P_C	Price of input	Endogenous
P_g	Price of output	Endogenous
Q^*	Optimal production quantity	Endogenous
X_S	Supply factor	Exogenous
X_D	Demand factor	Exogenous
γ_d	Demand elasticity	Constant parameter
γ_s	Supply elasticity	Constant parameter
$\mathrm{TC}(Q)$	Total cost of producing Q units	Convex in the production quantity
$\phi(Q)$	Capacity-related costs	Convex in capacity utilization
λ	Intensity of capacity-related costs	Capturing heterogeneity in production cost
$P_{I,g}$	Unit cost of other inputs	Constant parameter
F_C	Fixed cost of processing	Constant parameter

Table 1: Notations Used in the Model

A. Input Market

The price of a unit of the input; e.g. crude oil, is determined from a linear inverse supply function.

$$P_c(Q) = X_s + \gamma_s^{-1} Q \tag{1}$$

where $X_s(t)$ is the value of an exogenous, stochastic, supply factor at time t, Q is the supply of the input, and γ_s the elasticity of supply; i.e., the change in quantity supplied for a unit change in price. We assume that one unit of input is processed into one unit of output. Based on this assumption, the variable Q represents both the supply of the input and the quantity of the output. A positive supply shock, which corresponds to a decrease in the supply factor, X_s, decreases the price of the input and increases the quantity supplied.

B. Output Market

The price of a unit of output is determined by a demand factor X_d and the quantity of the output commodity Q. The inverse demand function is linear in quantity:

$$P_g(Q) = X_d - \gamma_d^{-1} Q \tag{2}$$

The demand factor X_d is exogenous and depends on long-term variables like the efficiency of the stock of existing appliances in use, and short-term shocks to income, taste, and seasonal factors. An increase in the demand factor X_d, increases the price as well as the quantity produced. We assume that demand and supply shocks are independent of each other.[3]

C. Output Production

We assume that the competitive industry is populated with many small representative firms which produce according to an increasing marginal cost schedule. This increasing marginal cost can be understood as the outcome of the deployment of production units on a *merit-order*, where the most efficient units of production are utilized first – as the production increases, higher marginal cost units are activated.

We assume that the marginal cost of production increases linearly with the amount produced; i.e., the total cost to produce Q units, $\mathrm{TC}(Q)$, is given by

$$\mathrm{TC}(Q) = F_c + QP_c + P_{I,g}Q + \phi(Q) = F_c + QP_c + P_{I,g}Q + \frac{\lambda}{2}Q^2 \tag{3}$$

where the total cost includes a fixed cost component, F_C, and a variable component that includes the cost of purchasing Q units of the input, QP_c, the cost of inputs other than the main input commodity, such as energy (electricity, natural gas, etc.), labor, materials, maintenance, etc., $P_{I,g}Q$,[4] and an increasing, quadratic, component $\phi(Q) = \lambda Q^2/2$, which

[3]In the case of the crude oil to refined products supply chain, supply shocks correspond to the discovery of additional supply, or to disruptions due to wars. Demand shocks correspond to unexpected economic growth, increased efficiency through new technologies, or the use of refined products in ways that were previously unanticipated. Given this intuition, it is reasonable to expect that supply and demand shocks are uncorrelated, or that, at least, their covariance is small enough that it can be safely ignored.

[4]We assume that the price of these inputs is not stochastic; i.e., it does not depend on the price of the input or the output. Assuming that their price is given by a random factor, creates an additional source of

represents the activation of higher marginal cost units.

D. Equilibrium: The competitive case

We assume that the output market is competitive, which implies that in equilibrium the market price of output equals the marginal cost. The equilibrium price and production satisfy

$$P_g = X_d - \gamma_d Q^* = \left. \frac{\partial \mathrm{TC}(.)}{\partial Q} \right|_{Q=Q^*} \tag{4}$$

The equilibrium quantity produced is given by:

$$Q^* = \frac{X_d - P_{I,g} - X_s}{\gamma_d + \gamma_s + \lambda} \tag{5}$$

The equilibrium price of the input and output, and their spread, are given by

$$\begin{aligned} P_c &= X_s + \gamma_s^{-1} Q^* = \frac{(\gamma_d^{-1} + \lambda) X_s + \gamma_s^{-1} X_d - \gamma_s^{-1} P_{I,g}}{\gamma_d^{-1} + \gamma_s^{-1} + \lambda} \\ P_g &= X_d - \gamma_d^{-1} Q^* = \frac{(\gamma_s^{-1} + \lambda) X_d + \gamma_d^{-1} X_s + \gamma_d^{-1} P_{I,g}}{\gamma_d^{-1} + \gamma_s^{-1} + \lambda} \\ \text{Spread} = P_g - P_c &= X_d - X_s - \left(\gamma_s^{-1} + \gamma_d^{-1}\right) Q^* = \frac{\lambda (X_d - X_s) + (\gamma_d^{-1} + \gamma_s^{-1}) P_{I,G}}{\gamma_d^{-1} + \gamma_s^{-1} + \lambda} \end{aligned} \tag{6}$$

The variances and covariances of input prices, output prices, the spread between output price and input price, and the variance of the quantity produced, are given by

volatility in spreads.

$$
\begin{aligned}
\mathrm{var}(P_c) &= \left(\frac{\gamma_d^{-1}+\lambda}{\gamma_d^{-1}+\gamma_s^{-1}+\lambda}\right)^2 \mathrm{var}(X_s) + \left(\frac{\gamma_s^{-1}}{\gamma_d^{-1}+\gamma_s^{-1}+\lambda}\right)^2 \mathrm{var}(X_d) \\
\mathrm{var}(P_g) &= \left(\frac{\gamma_d^{-1}}{\gamma_d^{-1}+\gamma_s^{-1}+\lambda}\right)^2 \mathrm{var}(X_s) + \left(\frac{\gamma_s^{-1}+\lambda}{\gamma_d^{-1}+\gamma_s^{-1}+\lambda}\right)^2 \mathrm{var}(X_d) \\
\mathrm{var}(P_g - P_c) &= \frac{\lambda^2}{(\gamma_d^{-1}+\gamma_s^{-1}+\lambda)^2}\left(\mathrm{var}(X_d)+\mathrm{var}(X_s)\right) \\
\mathrm{var}(Q^*) &= \frac{1}{(\gamma_d^{-1}+\gamma_s^{-1}+\lambda)^2}\left(\mathrm{var}(X_d)+\mathrm{var}(X_s)\right) \\
\mathrm{covar}(P_g - P_c, P_c) &= \frac{-\lambda(\lambda+\gamma_d^{-1})\mathrm{var}(X_s) + \lambda\gamma_s^{-1}\mathrm{var}(X_d)}{(\gamma_d^{-1}+\gamma_s^{-1}+\lambda)^2} \\
\mathrm{covar}(P_g - P_c, P_g) &= \frac{\lambda(\lambda+\gamma_s^{-1})\mathrm{var}(X_d) - \lambda\gamma_d^{-1}\mathrm{var}(X_s)}{(\gamma_d^{-1}+\gamma_s^{-1}+\lambda)^2}
\end{aligned}
\tag{7}
$$

IV. Properties of Prices and Spreads

Our model allows the study of the transmission of shocks across the supply chain. It helps us determine how properties of prices, such as the variance and the correlations of the price of input, the price of output, and their spread, depend on the convexity of the production function and the elasticities of supply and demand.

Proposition 1. *In a market described by Equations (1)-(5), the variance of the spread between the output price and the input price increases when the supply elasticity, γ_s, the demand elasticity, γ_d, and the coefficient of convexity, λ, increase; i.e.,*

$$
\begin{aligned}
\frac{\partial Var(P_g - P_c)}{\partial \gamma_s} &> 0 \\
\frac{\partial Var(P_g - P_c)}{\partial \gamma_d} &> 0 \\
\frac{\partial Var(P_g - P_c)}{\partial \lambda} &> 0.
\end{aligned}
$$

The proof of Proposition 1 follows from Equation (7).

Our next proposition illustrates the mechanism for risk transmission along the supply chain.

Proposition 2. *In a market described by Equations (1)-(5), the ratio of the variance of the price of the input to the variance of the price of the output is greater than 1 when only supply shocks are present, and smaller than 1 when only demand shocks are present. Moreover, with both supply and demand shocks, the value of the ratio is bounded by*

$$\frac{\gamma_s^{-2}}{(\gamma_s^{-1}+\lambda)^2} \leq \frac{var(P_c)}{var(P_g)} \leq \frac{(\gamma_d^{-1}+\lambda)^2}{\gamma_d^{-2}}$$

Proposition 2 follows from Equation (7), which shows that

$$\frac{\mathrm{var}(P_c)}{\mathrm{var}(P_g)} = \frac{(\gamma_d^{-1}+\lambda)^2\mathrm{var}(X_s) + \gamma_s^{-2}\mathrm{var}(X_d)}{\gamma_d^{-2}\mathrm{var}(X_s) + (\gamma_s^{-1}+\lambda)^2\mathrm{var}(X_d)},$$

which implies that the ratio increases (decreases) when the variance of the supply (demand) shocks increases.

Proposition 2 shows that the value of the ratio of the variance of the price of the input to the variance of the price of the output depends on the relative variance of supply shocks and demand shocks. If implied values for the variance of input and output prices are available, the proposition provides a mechanism to identify whether expected shocks are likely to be in the supply of the input or the demand of the output. Assuming that the elasticities of demand and supply, as well as the convexity of the production function remain constant, then an increase (decrease) in the value of the ratio can be attributed to an increase in the variance of supply (demand) shocks.

Proposition 3. *In a market described by Equations (1)-(5), an increase in the variance of the supply: a) decreases the covariance between the input price and the spread between the output price and the input price; b) decreases the covariance between the output price and the spread between the output price and the input price; and, c) increases the ratio of the* *variance of the input price to the variance of the output price. Similarly, an increase in the variance of the demand: d) increases the covariance between the input price and the spread between the output price and the input price; e) increases the covariance between the output price and the spread between the output price and the input price; and, f) decreases the ratio of the variance of the input price to the variance of the output price.*

The proof of the proposition follows from Equation (7).

Proposition 3 implies that when uncertainty is concentrated on the supply side, the input and output prices covary negatively with the spread, and when uncertainty is concentrated on the demand side, input and output prices covary positively. For the producer of the output, this behavior has an impact on whether to hedge by buying or selling forward contracts. The proposition also shows the relative importance of supply and demand shocks on prices. Parts c) and f) of the proposition show that supply shocks have a bigger impact on the input price, while demand shocks have a bigger impact in the output price.

A. The monopoly case

Up to this point we have assumed that producers are perfectly competitive price takers. In this section, we consider the case of a single producer that accounts for the effect of its production on output prices.

The monopolist maximizes the total profit $\pi = P_g(Q)Q - TC(Q)$ by considering the effect of its production decisions on the output market:

$$\max_{Q} Q(X_d - \gamma_d Q) - Q(P_{I,g} + X_s + \gamma_s^{-1} Q + \lambda Q) \tag{8}$$

The optimal amount produced by a monopolist is:

$$Q^*_{\mathrm{M}} = \frac{X_d - P_{I,g} - X_s}{2\gamma_d^{-1} + \gamma_s^{-1} + \lambda} \tag{9}$$

The equilibrium price of the input and output, and their spread, are given by

$$
\begin{aligned}
P_c = X_s + \gamma_s^{-1} Q_M^* &= \frac{(2\gamma_d^{-1} + \lambda)X_s + \gamma_s^{-1}X_d - \gamma_s^{-1}P_{I,g}}{2\gamma_d^{-1} + \gamma_s^{-1} + \lambda} \\
P_g = X_d - \gamma_d^{-1} Q_M^* &= \frac{(\gamma_s^{-1} + \gamma_d^{-1} + \lambda)X_d + \gamma_d^{-1}X_s + \gamma_d^{-1}P_{I,g}}{2\gamma_d^{-1} + \gamma_s^{-1} + \lambda} \\
\text{Spread} = P_g - P_c = X_d - X_s - \left(\gamma_s^{-1} + \gamma_d^{-1}\right) Q_M^* &= \frac{(\lambda + \gamma_d^{-1})(X_d - X_s) + (\gamma_d^{-1} + \gamma_s^{-1})P_{I,G}}{2\gamma_d^{-1} + \gamma_s^{-1} + \lambda}
\end{aligned}
\tag{10}
$$

The following propositions summarize the difference between a competitive market and a monopoly.[5]

Proposition 4. *In a market described by Equations (1)-(3), the variance of the quantity produced by a monopolist is smaller than the variance of the quantity produced in a competitive market.*

Proof. Comparing Q^*, and Q^*_{M} from Equations (5) and (9), we observe that the numerators of the three equations are the same but the denominator is greater in the monopoly case. Thus, the level and the variance of quantity are both smaller in the case of a monopoly.

□

Proposition 5. *In a market described by Equations (1)-(3), the variance of the spread between the output price and the input price is smaller in a competitive market compared to the variance of the spread in a monopoly.*

Proof. We note that the variance of the spread between the output and the input price in the case of a monopoly is given by

$$
\mathrm{Var}(P_g - P_c)^{\mathrm{M}} = \left(\frac{\lambda + \gamma_d^{-1}}{2\gamma_d^{-1} + \gamma_s^{-1} + \lambda}\right)^2 \left(\mathrm{Var}(X_d) + \mathrm{Var}(X_s)\right) \tag{11}
$$

[5] Similar results can be derived for a firm that has monopsony power in the input market.

The result follows from the fact that

$$\frac{d}{dx}\frac{(\lambda+x)^2}{(\gamma_s^{-1}+\gamma_d^{-1}+\lambda+x)^2}=2(\gamma_s^{-1}+\gamma_d^{-1})\frac{\lambda+x}{(\gamma_s^{-1}+\gamma_d^{-1}+\lambda+x)^3}$$

which is positive for $x \geq 0$.

$\square$

The intuition underlying Propositions 4 and 5 is that the production choices of the monopolist, compared to those of a producer in a competitive market, are less sensitive to both demand and supply shocks.

V. Minimizing variance of profits through hedging

Our results have implications for the optimal hedging program of the owner of the capital asset. We do not model the incentives of the firm to hedge. Instead, we assume that the objective of the commodity processing firm is to minimize the variance of its profit.

A. Profitability and Hedging

To determine the optimal hedge, we assume the firm has access to a single financial contract, a forward contract on the price of the input.

The firm's profit depends on two variables: 1) the profit margin $(P_g - P_c)$; and 2) the production quantity(Q).

$$\pi(Q) \approx Q(P_g - P_c) - QP_{I,g}$$

where $\pi(Q)$ is the profit for producing Q units of output – we have ignored costs that do not depend on the production quantity.

For a representative producer of the competitive industry, with a marginal cost that mirrors the marginal cost of the industry, the amount produced is proportional to the aggregate optimal amount produced, Q^*. As long as the price elasticity of demand is sufficiently small

– the gasoline market is one such case – the changes in quantity are negligible compared to changes in spreads. In such markets, the production quantity of a larger producer changes very little over time; however, its profit margin can vary significantly.[6]

When the changes in quantity are negligible compared to changes in the profit margin, the producer minimizes the variance of its profit by selling β units of the input to minimize the residuals in the relation

$$(P_g - P_c)_t = \alpha + \beta(P_{c,t} - F_{c,0}) + \epsilon \tag{12}$$

where $F_{c,0}$ is the futures price at time 0 for a unit of input delivered at time t. The optimal β, the hedge ratio, is given by the standard formula:

$$\text{hedge ratio} = -\frac{\text{covar}(P_g - P_c, P_c)}{\text{var}(P_c)} \tag{13}$$

B. Properties of the Variance-Minimizing Hedge Ratio

Equation (13) suggests that the variance-minimizing hedge ratio depends on the variance of the spread, the variance of the price of the input, and the correlation between the spread and the price of the input. Proposition 3 shows that, if uncertainty is driven by supply shocks, then the optimal hedge for the owner of the capital asset is to buy the input commodity in the forward market. On the other hand, if uncertainty is driven from demand shocks, the owner of the capital asset should sell the input commodity in the forward market.

In addition to determining the direction of the hedge; i.e., whether to buy or sell forward contracts on the input, our model offers guidance regarding the hedging amount, and the effectiveness of hedging.

The following proposition describes the optimal hedge ratio:

[6] For oil, the capacity utilization rate of the refinery industry only changes by a few percent. On the other hand, the profit margins can vary from \$6 per barrel to \$25 per barrel.

Proposition 6. *In a competitive market described by Equations (1)-(5), when demand is certain and supply is uncertain, an increase in the convexity coefficient, λ, increases the optimal amount bought; when only demand is uncertain, an increase in the convexity coefficient, λ, increases the optimal amount sold. The hedge ratio that minimizes the variance of profits is positive; i.e., long futures contracts, when demand is certain, but supply is uncertain, and negative when supply is certain and demand is uncertain. When both supply and demand are uncertain, the hedge ratio is bounded between*

$$-\frac{\lambda}{\gamma_s^{-1}} \leq \textit{hedge ratio} \leq \frac{\lambda}{(\lambda + \gamma_d^{-1})}$$

Proof. The results in Proposition 6 follow from the optimal hedge ratio

$$-\frac{\text{covar}(P_g - P_c, P_c)}{\text{var}(P_c)} = -\frac{-\lambda(\lambda + \gamma_d^{-1})\text{var}(X_s) + \lambda\gamma_s^{-1}\text{var}(X_d)}{(\lambda + \gamma_d^{-1})^2\text{var}(X_s) + \gamma_s^{-2}\text{var}(X_d)} \tag{14}$$

When demand is deterministic; i.e., $\text{var}(X_d) = 0$, and supply is uncertain, the optimal hedge is to buy $\lambda/(\lambda + \gamma_d^{-1})$ forward contracts, which implies that the hedge ratio is less than one, but tends to one when the convexity coefficient approaches infinity.

It is easy to see, from Equation (14), that the hedge ratio decreases as the variance of demand increases relative to the variance of supply. In the limit when supply is deterministic; i.e., $\text{var}(X_s) = 0$, and demand is uncertain, the optimal hedge is to sell λ/γ_s forward contracts. The bounds on the hedge ratio follow.

□

We note that, with only supply uncertain, the hedge ratio is between 0 and 1, while, when only demand is uncertain the hedge ratio can be very large in magnitude, especially when the convexity of the production function is large.

Two corollaries follow from the hedge ratio in Equation (14).

Corollary 1. *In a competitive market described by Equations (1)-(5), when demand is certain and supply is uncertain, the hedge ratio that minimizes the variance of profits decreases as the variance of demand shocks increases relative to the variance of supply shocks, and increases as the variance of supply shocks increases relative to the variance of demand shocks.*

Corollary 2. *In a competitive market described by Equations (1)-(5), the hedge ratio is positive when*

$$\frac{var(X_d)}{var(X_s)} > \frac{\lambda + \gamma_d^{-1}}{\gamma_s^{-1}}$$

and is negative otherwise.

In addition to changes in the optimal hedge ratio as uncertainty shifts from the input to the output, we can quantify the effectiveness of hedging using futures contracts for the input. We define hedging effectiveness as the reduction in the variance of the profits by hedging; i.e., the coefficient of determination, R^2, in the regression of the spread between profits and the value of futures contracts on the input.

Proposition 7. *In a competitive market described by Equations (1)-(3), the variance of hedged profits is zero, and hedging effectiveness is 100%, when there is a single source of uncertainty; either supply or demand shocks. If both supply and demand are uncertain, hedging effectiveness is zero when the hedge ratio that minimizes the variance of profits is zero.*

Proof. The coefficient R^2 is given by the square of the correlation between the profits of the firm and the value of the futures contract in the input.

$$R^2 = \frac{\left(-\lambda(\lambda + \gamma_d^{-1})\text{var}(X_s) + \lambda\gamma_s^{-1}\text{var}(X_d)\right)^2}{\left((\lambda + \gamma_d^{-1})^2\text{var}(X_s) + \gamma_s^{-2}\text{var}(X_d)\right)\lambda^2(\text{var}(X_s) + \text{var}(X_d))} \tag{15}$$

From Equation (15), it is obvious that for fixed variance of demand (supply), the numerator initially decreases (increases) and subsequently increases (decreases) as the variance of supply (demand) increases. When the variance of supply and demand are such that the hedge ratio is zero, the effectiveness of the hedge is also zero; i.e., there is no benefit in hedging the profits of the refiner with a futures contract on the input. Hedging effectiveness reaches 100% when either the variance of the supply shocks, $\text{var}(X_s)$, or the variance of the demand shocks, $\text{var}(X_d)$, is equal to zero. $\square$

Proposition 7 shows that hedging effectiveness is highest when there is a single source of uncertainty, either supply or demand, and that, when both demand and supply are uncertain, it is a non-monotonic function of the variance of supply or demand.

Another implication of Proposition 7 is that higher input or output volatility does not necessarily result in larger hedging positions; firms may indeed optimally reduce their hedging in response to higher volatility.

Our next proposition compares the hedge ratios of a monopolist to that of an operator in a competitive market.

Proposition 8. *In a market described by Equations (1)-(3), when there are only supply shocks, the hedge ratio that minimizes the variance of profits for a monopolist is greater than for an operator in a competitive market; i.e., the monopolist buys more of the input forward. On the other hand, when there are only demand shocks, the monopolist sells more of the input forward than an operator in a competitive market.*

Proof. The hedge ratio for a monopolist is given by

$$\frac{\lambda\gamma_s^{-1}\text{var}(X_d) - \lambda(\lambda+\gamma_d^{-1})\text{var}(X_s)}{\gamma_s^{-2}\text{var}(X_d) + (\lambda+\gamma_d^{-1})^2\text{var}(X_s)}$$

while the hedge ratio for an operator in a competitive market is given by

$$\frac{(\lambda+\gamma_d^{-1})\gamma_s^{-1}\text{var}(X_d) - (\lambda+\gamma_d^{-1})(\lambda+2\gamma_d^{-1})\text{var}(X_s)}{\gamma_s^{-2}\text{var}(X_d) + (\lambda+2\gamma_d^{-1})^2\text{var}(X_s)}$$

The proof follows from comparing the hedge ratios when the variance of the demand shocks, $\text{var}(X_d)$, and the variance of the supply shocks, $\text{var}(X_s)$ are equal to zero. With both supply and demand shocks, the hedge ratio for the monopolist is equal to zero when

$$\frac{\text{var}(X_d)}{\text{var}(X_s)} = \frac{\lambda + 2\gamma_d^{-1}}{\gamma_s^{-1}}$$

while, for an operator in a competitive market, it is equal to zero when

$$\frac{\text{var}(X_d)}{\text{var}(X_s)} = \frac{\lambda + \gamma_d^{-1}}{\gamma_s^{-1}}$$

$\square$

VI. Quantitative Analysis of Hedging in the Refinery Industry

We apply our model to the specific case of the refinery industry: the supply chain that transforms crude oil to refined products through a refinery. This is arguably one of the most important supply chains in the world, influencing almost every industry. There are major refining hubs located in North America, Western Europe, the Middle East, South/East Asia, and significant recent growth in refining capacity in South America. The refined products market in the United States is the largest in the world, consuming almost 20% of the global crude oil production.

In Section III we examined a model with linear demand and supply functions and linear marginal production cost and were able to derive closed-form solutions. The setting that we explore in this section is more realistic, but must be solved numerically. There are multiple objectives for such a numerical analysis. First, we want to test whether the theoretical results, derived in a simple model, hold in a real-world application. Second, we are interested in determining the magnitude of hedging effectiveness in a particular context. For example, one of the theoretical lessons of the model is that in cases with both supply and demand shocks one may not be able to reduce variance very much by hedging. However, this is a quantitative question, and we need a calibrated model to illustrate this.

A. Data

Crude oil can be refined to an array of products, including gasoline, distillates – i.e. diesel fuel, jet fuel, and heating oil – and heavy, or residual, products, such as gas oil, lubricants, and asphalt. Gasoline, represents close to 60% of refinery revenue and is the most important product for the majority of refineries. Gasoline yield; i.e., the ratio of gasoline output to total refinery output ranges between 42% and 48%.

We use a weighted basket of gasoline and heating oil to proxy for the output of a refinery. The difference in the price of this basket, made of two parts gasoline and one part heating oil, and the price of the crude oil input, is called the crack spread, and is a proxy for the gross margins of the refining industry.[7] The same 3-2-1 ratio is used in the crack spread contracts traded in NYMEX and other commodity exchanges.

We use annual data on prices, quantities of crude oil refined, and refining capacity.[8,9,10] Our data spans the period from 1987 to 2017. We obtain prices for crude oil, gasoline, and heating oil from the Energy Information Administration (EIA).[11] We obtain prices for the total amount of crude oil refined and the global refining capacity from the statistical review report issued by British Petroleum.[12]

[7]The profit of individual refineries can vary. For example, refineries that can process heavy or sour crude benefit significantly from the price difference between these types and light/sweet crude oil. We do not model the heterogeneity in crude oil prices.

[8]Gasoline prices are typically higher during the summer in the northern hemisphere, while heating oil prices are typically higher during the winter – using annual prices avoids the potential seasonality.

[9]Annual prices are the average of monthly prices within the year.

[10]Our choice to use annual data somewhat mitigates the need to model storage. Storage is particularly relevant for explaining monthly data – for example it is common to transfer excess production of gasoline from the low demand months of winter to the high demand months of summer. However, storage from year to year is less common, potentially due to the limited amount of storage available – total, worldwide, available storage is approximately equal to 3 months of consumption.

[11]`https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm`

[12]`https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.`

Statistic	Crude Oil	Gasoline	Heating Oil	Spread
Mean	54.12	64.27	64.35	10.12
Median	38.93	49.05	48.35	9.78
Maximum	117.56	128.17	136.13	15.51
Minimum	18.62	26.18	24.04	5.99
Standard Deviation	31.30	31.76	34.87	2.57
Skewness	0.90	0.79	0.90	0.27
Kurtosis	2.38	2.21	2.42	2.09

Table 2: Descriptive Statistics for prices of crude oil, gasoline, heating oil, and crack spread. Prices are real, deflated to 2012 dollars.

Figure 2 presents the time-series of the price of crude oil and the crack spread for the period 1987-2017. The prices are deflated using the consumer price index – the base year is 2012.

Table 2 provides descriptive statistics for annual prices and crack spreads. We use New York harbor gasoline, NYMEX heating oil, and Brent crude oil prices to calculate the crack spread. The values for the skewness and kurtosis suggest that the distributions, especially for the crack spread, are not symmetric, and that there is a large right tail.

Figure 3 presents a time-series of the ratio of the volatilities of the prices of crude oil and refined product. Similar to Figure 1, that illustrates that correlations between the price of the input and the price of crack spread vary over time, the relative volatility also varies over time. Both Figure 1 and Figure 3 suggest that the relative magnitude of supply and demand shocks varies over time – supply shocks are relatively larger in the late '80s and early '90s, while demand shocks dominate in the mid '90s and mid '00s.

B. Supply and demand functions

While the model in Section III allows for closed form solutions, to estimate a structural model for the crude oil to refined products supply chain we need a more general model.

Based on the crude oil production of low-cost members of the Organization of Petroleum

html

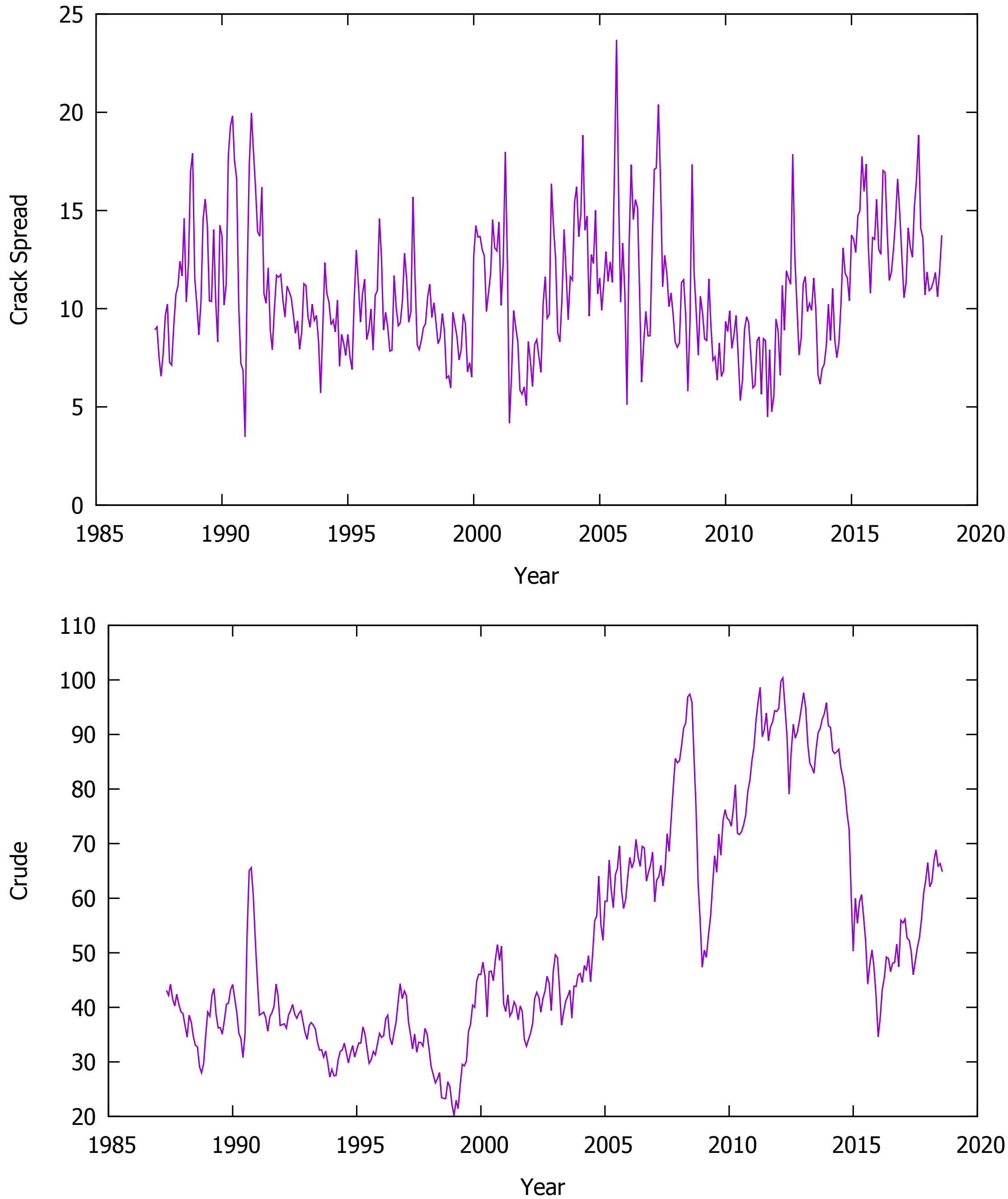


Figure 2: The figure presents the crack spread – top panel, measured in real terms, based in 2012 values and deflated using the consumer price index – and the price of crude oil – specifically the price of Brent, bottom panel. Prices and spreads are observed monthly.

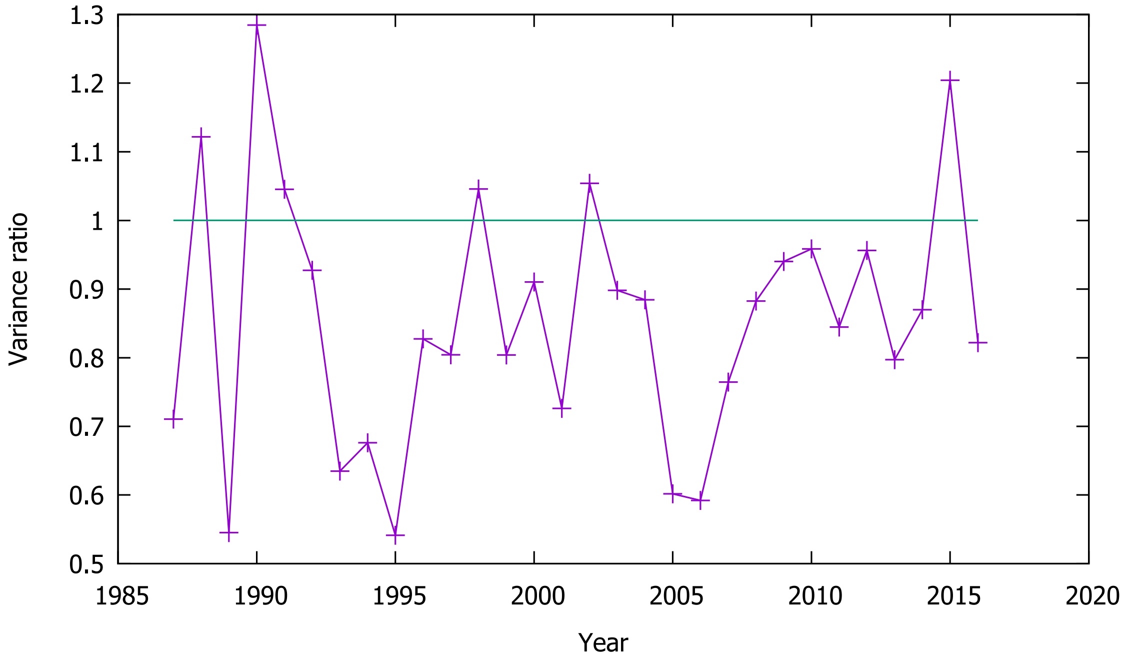


Figure 3: The figure presents the ratio of the volatility of the price of crude oil to the volatility of the price of refined products. The ratio is reported annually, with the volatilities calculated from the standard deviation of corresponding monthly returns within each year.

Exporting Countries (OPEC), we assume that the supply function is flat for low levels of production. The inverse supply function becomes steeper when the marginal supply moves to non-OPEC producers and unconventional sources. According to the literature; e.g. Dale (2016), the slope of the supply curve starts to increase at a quantity equal to 80% of global refining capacity. The inverse supply function is given by:

$$P_C = e^{X_s} + \Big((Q - \underline{Q}) \mathbb{1}_{Q > \underline{Q}} \Big)^{\gamma_s} \tag{16}$$

where the factor, X_s, captures exogenous shocks to the global supply of oil, and shifts the inverse supply function; $\underline{Q}$ is the threshold where the steep part of the inverse supply function begins; and γ_s is the sensitivity of the marginal cost of crude oil to the level of production of crude oil at levels of production above the threshold.

We model the dynamics of the stochastic factor X_s as a mean-reverting process.

$$dX_s = \mu_s(\overline{X}_s - X_s)dt + \sigma_s dW_s \tag{17}$$

where the mean reversion rate μ_s, the long term level $\overline{X}_s$, and the volatility σ_s are assumed constant.

Consistent with the demand elasticity literature; e.g., Liu (2014) we use a constant-elasticity function for the demand of refined products[13]

$$P_G = e^{X_d} Q^{\gamma_d} \tag{18}$$

We assume that the stochastic demand factor, X_d, follows a mean-reverting process with shocks that are correlated with the shocks to the net supply of crude oil

$$dX_d = \mu_d(\overline{X}_d - X_d)dt + \sigma_d dW_d \tag{19}$$

where μ_d is the speed of mean reversion, X_d is the long term level of the demand factor, and σ_d is the volatility of the shocks to the demand of refined products.

Variable	Functional Form
Crude oil supply	$P_C = e^{X_s} + \left((Q - \underline{Q})\mathbb{1}_{Q>\underline{Q}}\right)^{\gamma_s}$
Supply shift factor	$dX_s = \mu_s(\overline{X}_s - X_s)dt + \sigma_s dW_s$
Refined products demand	$P_G = e^{X_d} Q^{\gamma_d}$
Demand shift factor	$dX_d = \mu_d(\overline{X}_d - X_d)dt + \sigma_d dW_d$
Marginal cost of production	$\mathrm{MC}_t = P_{I,g} + P_C + \phi_t(Q)$
Capacity-related costs	$\phi_t(Q) = \lambda_t Q^{\eta}$
Shocks to capacity	$\lambda_t = \overline{\lambda} e^{\sigma_\lambda \epsilon_t}$

Table 3: Supply, demand, and production functions for the crude oil to refined products supply chain.

[13]This functional form is equivalent to the log-log specification typically used to estimate gasoline demand elasticity in the literature.

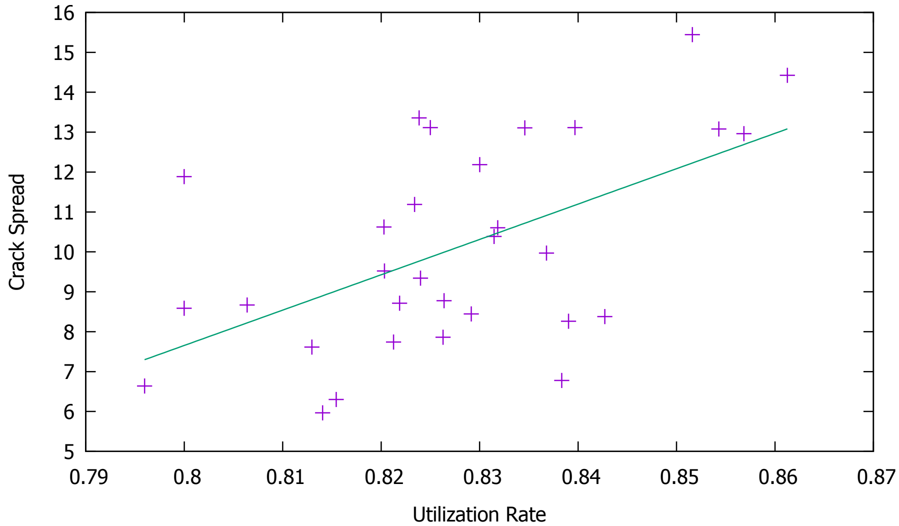


Figure 4: Crack Spread vs. Global Capacity Utilization. The crack spread values are expressed in real terms – deflated to 2012 prices using the consumer price index. The data are reported annually between 1987 and 2017.

C. Production Function

Figure 4 illustrates the relationship between global capacity utilization and the level of crack spreads. In a competitive market the crack spread equals the marginal production cost – the figure suggests that the marginal production cost is increasing as capacity utilization increases. The figure includes the linear least squares fit – a linear function is a close match to the data. The coefficient of determination, R^2; i.e., the percentage of the variance in the crack spread explained by capacity utilization is 32%.

To accommodate the residuals in Figure 4, potentially due to changes in production technologies, the delay between planning and building a refinery which results in shortages and oversupply, and exogenous shocks; e.g., hurricanes, that affect production, we assume that the production function is influenced by random fluctuations.

We note that, during the period we study, the amount of crude oil refined has increased significantly. However, Figure 4, suggests that the cost to refine a certain percentage of the total refining capacity remains relatively stable, possibly due to improvements in refining technology. To account for this feature, rather than expressing the quantity of crude oil that is refined each year in absolute terms; e.g., number of barrels of oil per day, we instead express it as a percentage of refining capacity.

Given the information in Figure 4, we assume that the marginal cost of refining Q units of crude oil in period t is given by

$$\mathrm{MC}_t = P_{I,g} + P_C + \phi_t(Q) \tag{20}$$

where $P_{I,g}$ is the cost of other inputs, P_C is the price of crude oil, and $\phi_t(Q)$ are costs that are capacity-related. The capacity-related costs, $\phi_t(Q)$, are assumed to be given by a power function, with two unknown coefficients λ_t and η.

$$\phi_t(Q) = \lambda_t Q^{\eta} \tag{21}$$

We assume that the convexity of the production function, described by the power η, is constant. However, to account for the residuals in Figure 4, we assume that the coefficient λ_t is a random variable drawn from the following distribution

$$\lambda_t = \overline{\lambda} e^{\sigma_\lambda \epsilon_t} \tag{22}$$

where $\overline{\lambda}$ is the baseline value for the parameter, σ_λ is the standard deviation of shocks to the parameter, and ϵ_t are standard, normally distributed, i.i.d. random variables.

The fluctuations in the production function have significant empirical consequences: for example, they influence the volatility of the crack spread. Given the i.i.d. nature of the random variables, ϵ_t, we expect that the marginal production costs will be autocorrelated.

D. Caveats

Our data consists of a time-series of annual prices and quantities between 1987-2017 – a total of 31 observations. It is clear that the small number of observations limits our choices.

There are several, potential, improvements that we could attempt with a longer time-series.

Supply Function of Crude Oil. We use a two-piece power function to model the increasing marginal cost of crude oil. In reality, the crude oil cost curve is much more complicated. In addition, several papers in the literature argue that the oil sector contains dynamic and delayed responses, see Wirl (2008).[14]

Time-Varying Variables. Given the small number of observations, we assume that the parameters of our model are constant. In particular, we assume that the parameters for the production function for the refinery sector, the demand for refined products, and the production function of crude oil are constant. However, refining technology has significantly improved, and our adjustment of using quantity produced as a percentage of refining capacity may not be accurate. In addition, there is evidence that consumer demand elasticity for gasoline has changed over time, see Hughes et al. (2006). Finally, the supply curve for crude oil has shifted significantly due to several factors, including the introduction of unconventional energy sources.

Storage. Our model is an instantaneous production model, with no inter-temporal storage of refined products or crude oil. In reality, both crude oil and refined products are stored. In the presence of storage, the optimal production will be less sensitive to demand shocks and more sensitive to input shocks because the refiners can produce in low-demand periods and store the refined products in order to sell them in periods with higher demand, and possibly with more expensive input. Since we do not model storage, our estimate of the elasticity of demand may be biased.

Shut-Downs. Refiners need to stop operation for necessary maintenance. The planned shut-downs generate patterns in the production rate that we do not capture. A maintenance

[14]For example, the presence of price adjusting mechanisms, such as the Organization of Petroleum Exporting Countries, prevent the oil market from following a static cost curve.

schedule chosen strategically; e.g., plan shut-downs when demand is expected to be low, would generate endogenous capacity adjustments. For example, in years in which oil prices are high, refiners may decide to postpone maintenance activities and keep the average operable capacity at a higher level. Since most of the planned shut-downs happen within the year, using annual data helps mitigate this potential bias.

E. Estimation

We use the Simulated Method of Moments (SMM) to estimate the structural parameters of the model. Starting with an initial guess of the parameters, we simulate several scenarios for the model. We calculate the model-generated moments for each scenario, and then average over all the scenarios. We compare the model-generated moments to the empirical ones and modify the initial guess of the model parameters until the percentage difference between the model-generated moments and the empirical moments is smaller than a cut-off. Additional details on SMM can be found in Strebulaev and Whited (2012).

The empirical moments we use in the SMM procedure, are the mean and standard deviation of the price of refined products; the mean and standard deviation of crude oil prices; the annual autocorrelation of the price of refined products; the correlations between the price of refined products and the crack spread, the price of refined products and crude oil, and the price of crude oil and the crack spread; the mean and standard deviation of capacity utilization; the mean of the crack spread; and the explained variance of the crack spread. Overall, we use 12 empirical moments.

The model parameters are: the demand and supply elasticities; the mean-reversion rates for supply and demand shocks; the standard deviation of supply and demand shocks; the long-run level of the supply and demand factors; the degree of convexity and the coefficient of the convexity term in the production function; and the standard deviation of fluctuations to the output from the production function. The total number of parameters is 11, implying that our model is over-identified.

F. Starting Values

A subset of parameters can potentially be estimated directly. This is possible by either using estimates provided in the literature, or through direct identification. In order to improve the performance of the SMM estimation, we use starting values for every parameter.

Elasticity of Demand. The literature on demand for refined products; especially gasoline, reports a wide range of values for short-term price elasticity of demand between 0.00 to -0.15 for different countries – see Cooper (2003), and Hughes et al. (2008).[15] We choose a starting value of γ equal to -20.

Supply Elasticity. In anticipation of, and during the disruption in 2011 to the supply of Libyan crude oil to the global markets, prices increased by close to \$10 per barrel. Given that Libyan crude production is close to one million barrels per day, we use a starting value for supply elasticity, that corresponds to the assumption that an additional production of a million barrels of crude oil per day increases the price of crude oil by \$10 per barrel. This assumption corresponds to a starting value for supply elasticity of $\gamma_s = 1.45$.

Stochastic Processes for Supply and Demand Factors. We use statistical properties of the historical prices of crude oil and refined products to determine the starting values for the long-term levels, volatility parameters, and mean-reversion rates of the supply and demand processes.

[15] Note that our elasticity parameter γ is the inverse of typically reported elasticity parameters. The estimates in the literature are based on price elasticity for a particular country or region. Since our model reflects global demand, it is possible that the value of the estimated demand elasticity may be outside the range given in the literature.

Cost of Processing. Based on industry estimates, we assume that the processing cost, net of the cost of crude oil and the nonlinear, convex, term, is \$3 per barrel.[16] Going forward, we report the crack spread net of this processing cost.

Convexity of the Production Function. Figure 4 suggests that the relationship between the deflated crack spread and the global capacity utilization is close to linear. Since, in a competitive market, the crack spread corresponds to the marginal cost of the production function, we choose a starting value for the power of the convex part of the marginal cost to be $\eta = 1$, corresponding to quadratic production costs with respect to the quantity of oil refined. To estimate a starting value for the coefficient λ we use the intercept of the linear relationship between crack spreads and capacity utilization, which provides a starting value equal to 3.2. To approximately match the explained variance of the crack spread, we set the initial value of the standard deviation of the random variables in the production function to 10%.

Table 4 lists the 11 parameters we estimate, their starting values, and the values that result in model-generated moments that best match the empirically observed moments.

Simulation Parameters. For the SMM procedure, we use 100 iterations. For each iteration, we set the number of time periods of the simulated vector T to 200 – each period corresponds to one year. We discard the first 100 periods in each iteration and use the next 100 periods to estimate the various moments.

We report the performance of the model in matching empirically observed moments in Table 5.

We observe that the model provides a reasonable match for the level and standard deviation of the price of crude oil, the price of refined products, and the crack spread. The autocorrelation between the prices of crude oil, refined products, and crack spread, are also

[16] See `https://www.iea.org/media/omrreports/Refining_Margin_Supplement_OMRAUG_12SEP2012.pdf`.

Notation	Parameter	Starting Value	Source of Starting Value	Estimated Value
γ_d	Demand elasticity	-20	Literature	-22.4
γ_s	Quantity sensitivity of crude oil price	1.45	Change of $10/barrel per extra 1M barrels	2.07
μ_d	Mean-reversion rate for demand (year^{-1})	0.3	Statistics of prices	0.11
μ_s	Mean-reversion rate for supply (year^{-1})	0.3	Statistics of prices	0.05
σ_d	Standard deviation of demand	0.3	Statistics of prices	0.53
σ_s	Standard deviation of supply	0.3	Statistics of oil prices	0.14
$\overline{X}_s$	Long-run supply factor	3	Statistics of crude oil prices	2.90
$\overline{X}_d$	Long-run demand factor	100	Statistics of refined products prices	101.24
η	Convexity of marginal cost function	1.00	Regression of crack spreads on capacity utilization	1.85
λ	Coefficient of marginal cost function	3.42	Regression of crack spreads on capacity utilization	0.0022
σ_λ	Shocks to Capacity	0.1	Regression of crack spreads on capacity utilization	0.07

Table 4: List of Parameters

Moment	Empirical Value	Model-Generated Value	Weight
Mean Refined Prices	64.27	64.85	20
S.D of Refined Prices	31.76	31.23	10
Mean Crude Oil Prices	54.12	54.62	20
S.D of Crude Oil Prices	31.30	30.93	10
Mean of Crack Spreads	7.12	7.23	20
Explained Volatility of Crack Spreads	0.10	0.10	20
Average Capacity Utilization	82%	80%	10
S.D of Capacity Utilization	1.7%	2.4%	5
Autocorrelation of Annual Refined Prices	0.86	0.86	5
Correlation of Crack Spread and Refined Prices	0.39	0.27	5
Correlation of Crack Spread and Crude Oil Price	0.26	0.23	5
Correlation of Crude Oil and Refined Prices	0.99	0.99	5

Table 5: Moments Used in the Simulated Method of Moments Estimation

accurate in terms of the sign and reasonably close in terms of size.

The worst match occurs for the standard deviation of capacity utilization. The model-generated standard deviation of capacity utilization is 50% larger than the empirical value.[17]

Our results suggest that demand shocks are the major source of shocks during the period we study. This finding is in line with recent empirical results; e.g., Kilian (2009), which decompose the volatility of the crude oil prices to supply and demand driven shocks and conclude that the share of demand shocks in explaining the volatility of crude oil prices has increased over time.

[17] This mismatch is largely due to the method trying to match the model volatility of crack spreads with the empirical value. Since we do not account for certain shocks; e.g., shocks to the prices of energy and materials used by in the refining process, the method increases the volatility of capacity-utilization to match the volatility of crack spreads with the empirical value.

VII. Hedging Effectiveness

Our calibrated model can provide quantitative guidance regarding the optimal hedging policies for a refiner that can only access the forward market on input (i.e. crude oil).[18]

The optimal number of forward contracts required to hedge the refiners profits is given by

$$\frac{Q_F}{Q} = -\frac{\text{covar}(P_g - P_c, P_c)}{\text{var}(P_c)}$$

The correlation of profits with input or output prices can change because of changes in the level of supply and demand, as well as changes in their variances.[19]

To evaluate the effectiveness of hedging, we generate 1000 random scenarios for demand and supply shocks for the next year. The variance of realized refinery profits, defined as the crack spread plus the pay-off from hedging, is then calculated under different hedging policies.

A. Baseline

We first consider the effectiveness of hedging using our model for the calibrated set of parameters. We find that, for the basecase, since demand shocks are relatively larger than supply shocks, to hedge profits, oil refiners with a production cost that reflects the production function of the entire industry, should sell forward a small number of contracts of crude oil. Without hedging, we find that the variance of next-period profits, in dollars per barrel per unit of production is 26.0%, while the variance of hedged profits is 23.5%, an improvement

[18] The effectiveness of optimal hedging strategies for refiners using both input and output forward contracts has been discussed by several recent papers; e.g., Ji and Fan (2011), Alexander et al. (2013), and Liu et al. (2017). Rather than use a structural approach, these papers use statistical estimation to compare various strategies using both forward contracts for both crude oil and refined products to hedge the crack spread.

[19] We note that our results assume that the decision-maker knows the true parameters of the model and that the model is correctly specified. Otherwise, the calculation is subject to both sampling and specification errors. The literature on optimal decision making under parameter uncertainty; e.g., Smith and Winkler (2006), shows that a rational decision-maker, in the face of sampling and specification errors, may refrain from taking the suggested optimal decision.

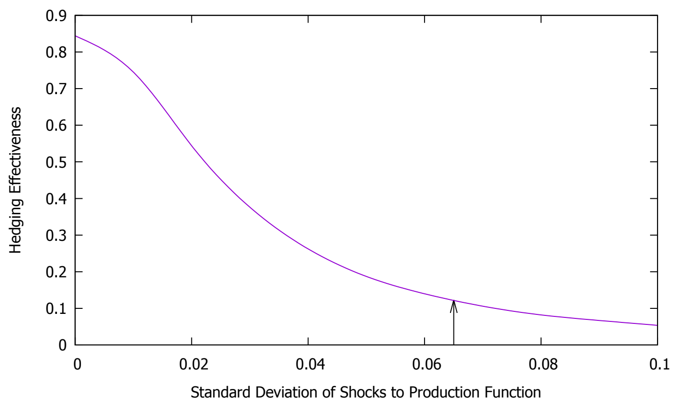


Figure 5: Impact of Exogenous Shocks on Production Function on Hedging Effectiveness. The vertical line with the arrow corresponds to the base case of the calibrated model.

of approximately 10%.

Hedging effectiveness is low for two reasons: the hedge ratio is small, due to both supply and demand shocks; and, there are fluctuations to the production cost function that further diminish hedging effectiveness.

We plot the impact of the volatility of shocks to the production function on the effectiveness of the hedging strategy in Figure 5. We note that hedging effectiveness can be above 80% if the fluctuations in the production function are small. Hedging effectiveness quickly deteriorates as the size of the fluctuations increases.

B. Dynamic versus Static Hedging

The optimal hedge ratio in our structural model varies over time for two reasons: endogenously – following the realization of supply and demand shocks, the optimal quantity produced, the crack spread, the correlation between the crack spread and the input price, and the hedge ratio change; and exogenously – if, for example, the variance of supply or demand shocks changes, so does the hedge ratio.

We consider the variation due to endogenous changes and quantify the impact on hedging effectiveness in two ways. First, we consider the cross-sectional variation of the hedge ratio and hedging effectiveness as the variance of supply and demand change. Figure 6 shows

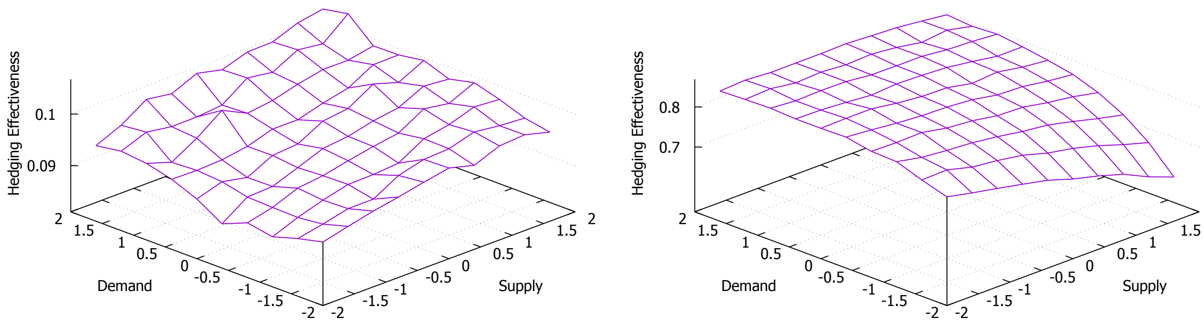


Figure 6: Hedging effectiveness vs. supply and demand shocks over a single year. The left panel shows hedging effectiveness for the baseline case. The right panel shows hedging effectiveness in the case without exogenous shocks to the production function. The units for the supply and demand shocks are in terms of standard deviations away from their long term mean.

the optimal hedge ratio and hedging effectiveness, over a single year, for different levels of the initial realization of demand and supply shocks, keeping all structural parameters of the model constant. The value of the supply and demand shocks varies within two standard deviations of their long-term level. The figure illustrates that the presence of fluctuations to the production function reduces hedging effectiveness significantly across all values of supply and demand shocks. We note that the fluctuations, evident in the left panel, are due to the large variance of the results – the values of hedging effectiveness are not statistically different across the values of the supply and demand shocks displayed. On the other hand, when random fluctuations to the production function are eliminated, the right panel illustrates that hedging effectiveness is high

Without fluctuations to the production function, hedging effectiveness depends on capacity utilization. Since, in the calibrated model, the production function is convex, when capacity utilization is low convexity is low, and the correlation of spreads and input prices is close to zero. On the other hand, when the capacity utilization is high, the correlation and hedging effectiveness can be much higher.

The second way we quantify hedging effectiveness is across time. Starting with the calibrated values for the parameters and the levels of supply and demand, we simulate 500 paths, each for 100 years, and evaluate hedging effectiveness by comparing the cumulative variance of profits for a dynamic strategy, where the hedge ratio is updated each year, and a static strategy, where the hedge ratio is set once, based on the long-term level of the supply and demand shocks. The results suggest that, relative to static hedging, dynamic hedging decreases the variance of profits only when fluctuations to the production function are small. In particular, without fluctuations to the production function, the cumulative variance of profit is reduced by 84% using dynamic hedging, and 76% using static hedging. With the calibrated level of fluctuations to the production function on the other hand, both strategies reduce the cumulative variance of profit by 10%.[20]

C. Comparative Statics

Our results suggest that hedging effectiveness is high when there is a single source of uncertainty; either supply or demand shocks; when fluctuations to the production function are small; and when the convexity of the production function is large.

Figure 7 explores hedging effectivess as the volatilities of supply and demand change. The figure shows that hedging effectiveness is non-monotonic. For example, as the volatility of supply increases, the hedging effectiveness initially drops and then starts increasing. The intuition behind this result is that the correlation between crude oil prices and crack spreads, when the volatility of supply is very low, is large and positive, and hedging effectiveness high. As the volatility of supply increases, the correlation drops. At some point – which is a function of demand and supply volatility as well as other, structural, parameters – the correlation of the crack spread and the price of crude oil, the hedge ratio, and the hedging effectiveness

[20]Our simulation assumes that the model parameters are constant. The benefit of dynamic hedging can hedging can potentially be significantly higher when volatitilies of supply and demand change, and can be determined from the price of forward looking financial instruments; e.g., the implied volatility of option prices.

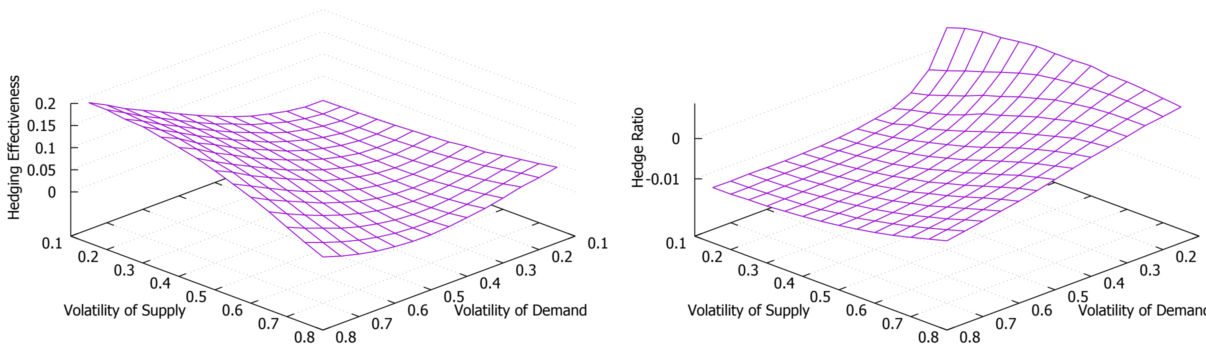


Figure 7: Hedging effectiveness (left panel) and hedge ratio (right panel) against the volatility of supply and demand shocks.

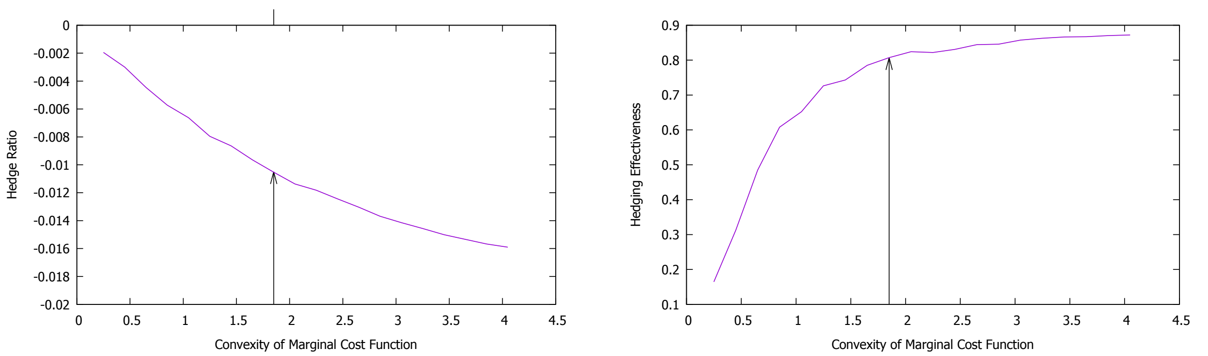


Figure 8: The left panel shows the hedge ratio that minimizes the variance of profits against the convexity of the marginal cost function. The right panel shows the corresponding hedging effectiveness. The vertical lines with the arrow correspond to the base case of the calibrated model

become zero. As the volatility of supply increases further, the correlation between crude oil price and crack spreads becomes negative, and hedging effectiveness increases.

Figure 8 shows the optimal hedge ratio (left panel) and hedging effectiveness (right panel) as the convexity of the production function varies. When the convexity of the production function is low, the crack spread is less volatile, and its correlation with input and output prices, as well as the hedge ratio drop. When the production function becomes more convex, the hedge ratio increases in absolute terms. The right panel of Figure 8 shows that hedging effectiveness increases as the convexity of the production function increases.

VIII. Conclusions

There is a substantial literature that examines the motivations for firms to reduce the variance of their cash flows by hedging.[21] This literature implicitly assumes that firms can in fact effectively hedge. In this paper we have taken a careful look at the economic factors that affect the effectiveness of hedging in an equilibrium model with endogenous input and output prices that are determined by stochastic supply and demand shocks as well as by the characteristics of the production process that transforms the input into the output. We estimated our model with data on the oil refinery industry to get a calibrated base case and provided numerical comparative statics on the determinants of hedging effectiveness. As we show, hedging is not expected to be particularly effective in most plausible cases.

To summarize, our analysis generates the following insights: 1) the source of shock (demand versus supply) is a key determinant of the direction of the co-movement between profits and input/output costs and consequently optimal hedging position of firms. In industries where most shocks come from the supply side, the firm should take a short position on the futures contracts for the input; when most come from the demand side, the optimal hedging requires a long position in input commodity; 2) there are natural bounds on the effectiveness of hedging, which is determined by the convexity of the cost function and the elasticity of demand; 3) market competition reduces the volatility of spreads in commodity processing industries; 4) one-sided financial hedging, where a producer can hedge the costs of his input, but not the price of his output, may not be efficient in plausible scenarios, when supply shocks and demand shocks are of similar magnitudes, or in the presence of significant exogenous noise in the production function.

Although our message is relatively negative about the efficacy of one-sided hedging, it

[21] This theoretical literature started in the 1980s and 1990s with papers by Shapiro and Titman (1986), Smith and Stulz (1985) and Froot et al. (1993). Empirical studies include Tufano (1996) and Geczy et al. (1997).

does point to a number of ways that an informed dynamic hedging policy can increase its effectiveness. For example, to account for the observed variation in the crack spread, our model includes a substantial amount of exogenous noise in the cost of refining that we assume cannot be observed in advance by the hedger. In reality, some of these costs can be observed in advance, and the producer can hedge more effectively by accounting for them. In addition, we have assumed that producers cannot anticipate changes in the variance of future demand and supply shocks. In reality, changes in these variances can be directly observed, e.g., an increase in tensions in the Middle East implies an increase in the variance of future supply shocks, or they can be indirectly observed from information in the financial markets. Indeed, an important implication of our model is that the variance of the input price increases more than the variance of the output price when the supply shock variance increases, while the variance of the output price increases more when the variance of the demand shock increases. In theory, one can glean useful information about these variances from the commodity options markets.

We focused our attention on the petroleum refining industry because of the availability of data. However, our analysis applies to any industry that converts an input commodity into an output. In some sense, the refining industry is not ideal for our analysis of effectiveness of hedging inputs, since this is an industry where it is possible to at least partially hedge the outputs as well as the inputs. For this reason, in future research, it makes sense to study markets with outputs like broiler chickens and airline travel, which cannot be directly hedged.

The tight relation between its inputs and outputs also makes the petroleum refining industry somewhat unique. It might also be interesting to consider how our model applies to industries where these linkages are not as tight. For example, electricity is an important input for aluminum smelters, however, the link between electricity prices and aluminum prices do not co-move to the same extent as the prices of oil and gasoline for a couple of reasons. First, since aluminum production accounts for a very small part of the overall use of electricity, shocks to aluminum demand is likely to have a very small effect on electricity prices. Second, the price of aluminum is determined on world markets, since it is inexpensive to ship, while electricity is a local commodity, whose price is determined by regional supply and demand conditions that do not directly influence the price of aluminum. For these reasons, shocks to the supply of electricity have very little influence on aluminum prices, which is why aluminum producers tend to buy electricity with long term forward contracts.

A final extension of our model that we will leave to future research has to do with the production function. In our current model producers convert one unit of the input good into one unit of the output good at a cost that increases as capacity utilization increases. Alternatively, we can assume that the process converts one unit of the input to X units of the output, where X can vary from plant to plant (some are more or less efficient) and it can vary as capacity utilization increases and decreases. For example, older gas fired power plants have higher heat rates than newer plants, which means that they use more natural gas to produce the same amount of electricity. Similarly, big airplanes use less fuel per passenger than smaller airplanes.

For these industries, the optimal hedge ratio for the more or less efficient producer can be very different. We leave these extensions for future research.

Alexander, C., Prokopczuk, M., Sumawong, A., 2013. The (de) merits of minimum-variance hedging: Application to the crack spread. Energy Economics 36, 698–707.

Bessembinder, H., Lemmon, M.L., 2002. Equilibrium pricing and optimal hedging in electricity forward markets. Journal of Finance 57, 1347–1382.

Carmona, R., Durrleman, V., 2003. Pricing and hedging spread options. SIAM Review 45, 627–685.

Carter, D.A., Rogers, D.A., Simkins, B.J., Treanor, S.D., 2017. A review of the literature on commodity risk management. Journal of Commodity Markets 8, 1–17.

Casassus, J., Liu, P., Tang, K., 2012. Economic linkages, relative scarcity, and commodity futures returns. Review of Financial Studies 26, 1324–1362.

Cooper, J.C., 2003. Price elasticity of demand for crude oil: estimates for 23 countries. OPEC Energy Review 27, 1–8.

Dale, S., 2016. New economics of oil. Oil & Gas, Nat. Resources & Energy J. 1, 365.

Dong, L., Kouvelis, P., Wu, X., 2014. The value of operational flexibility in the presence of input and output price uncertainties with oil refining applications. Management Science 60, 2908–2926.

Froot, K.A., Scharfstein, D.S., Stein, J.C., 1993. Risk management: Coordinating corporate investment and financing policies. Journal of Finance 48, 1629–1658.

Geczy, C., Minton, B.A., Schrand, C., 1997. Why firms use currency derivatives. Journal of Finance 52, 1323–1354.

Hirshleifer, D., 1988a. Residual risk, trading costs, and commodity futures risk premia. Review of Financial Studies 1, 173–193.

Hirshleifer, D., 1988b. Risk, futures pricing, and the organization of production in commodity markets. Journal of Political Economy 96, 1206–1220.

Hirshleifer, D., 1989. Determinants of hedging and risk premia in commodity futures markets. Journal of Financial and Quantitative Analysis 24, 313–331.

Hughes, J., Knittel, C., Sperling, D., et al., 2008. Evidence of a shift in the short-run price elasticity of gasoline demand. Energy Journal 29, 113–134.

Hughes, J.E., Knittel, C.R., Sperling, D., 2006. Evidence of a shift in the short-run price elasticity of gasoline demand. Technical Report. National Bureau of Economic Research.

Ji, Q., Fan, Y., 2011. A dynamic hedging approach for refineries in multiproduct oil markets. Energy 36, 881–887.

Kamara, A., 1993. Production flexibility, stochastic separation, hedging, and futures prices. Review of Financial Studies 6, 935–957.

Kilian, L., 2009. Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. American Economic Review 99, 1053–69.

Liu, P., Vedenov, D., Power, G.J., 2017. Is hedging the crack spread no longer all it's cracked up to be? Energy Economics 63, 31–40.

Liu, W., 2014. Modeling gasoline demand in the united states: A flexible semiparametric approach. Energy Economics 45, 244–253.

Secomandi, N., 2010. On the pricing of natural gas pipeline capacity. Manufacturing & Service Operations Management 12, 393–408.

Shapiro, A.C., Titman, S., 1986. An integrated approach to corporate risk management. The revolution in corporate finance 3, 251–265.

Smith, C.W., Stulz, R.M., 1985. The determinants of firms' hedging policies. Journal of Financial and Quantitative Analysis 20, 391–405.

Smith, J.E., Winkler, R.L., 2006. The optimizers curse: Skepticism and postdecision surprise in decision analysis. Management Science 52, 311–322.

Strebulaev, I.A., Whited, T.M., 2012. Dynamic models and structural estimation in corporate finance. Foundations and Trends in Finance 6, 1–163.

Tufano, P., 1996. Who manages risk? an empirical examination of risk management practices in the gold mining industry. Journal of Finance 51, 1097–1137.

Wirl, F., 2008. Why do oil prices jump (or fall)? Energy Policy 36, 1029–1043.

Wu, O.Q., Chen, H., 2010. Optimal control and equilibrium behavior of production-inventory systems. Management Science 56, 1362–1379.

Financialization and Commodity Market Serial Dependence*

Zhi Da[†], Ke Tang[‡], and Yubo Tao[§]

This Draft: January 2019

Abstract

Recent financialization in the commodity market makes it easier for institutional investors to trade a portfolio of commodities via various commodity index products. Using news-based sentiment measures, we find that such trading can propagate non-fundamental shocks from some commodities to others in the same index, giving rise to price overshoots and subsequent reversals, or "excessive co-movement" at daily frequency. Excessive co-comovement results in negative daily commodity return autocorrelations even at the index level (but not for non-indexed commodities) and such autocorrelations move with our commodity index exposure measures. Taking advantage of the fact that index weights of the same commodity can vary across different indices in a relatively ad-hoc and pre-determined fashion, we provide causal evidence that index trading drives excessive co-movement. Overall, our paper adds value to the understanding of price discovery and market efficiency of financialized commodity futures markets.

JEL Classification: G12, G40, Q02.
Keywords: Return autocorrelation; Index trading; News sentiment; Non-fundamental shocks.

*We have benefited from comments and suggestions from Li Guo, Jing Wu and seminar participants at Notre Dame University and the 3rd Australian Commodity Markets Conference.

†University of Notre Dame, Mendoza College of Business, Notre Dame, IN, 46556, USA. Email: zda@nd.edu.
‡Tsinghua University, Institute of Economics, Beijing, 100084, China. Email: ketang@mail.tsinghua.edu.cn.
§Singapore Management University, School of Economics, 90 Stamford Road, 178903, Singapore. Email: yubo.tao.2014@phdecons.smu.edu.sg

1 Introduction

The last two decades witnessed the financialization of the commodity markets. According to the estimates from the Commodity Futures Trading Commission (CFTC), investment flows to various commodity indices exceeded $600 billion during the period from 2000 to 2017. Coinciding with the large investment inflow to commodity indices, different commodities started to display synchronized boom and bust cycles. In addition, Tang and Xiong (2012) find such co-movement to be more severe for commodities in popular indices (indexed commodities) than for those excluded from indices (non-indexed commodities), as shown in Figure 1. [1]

[Figure 1 is about here.]

Co-movement among indexed commodities in itself, however, does not necessarily imply that financialization is the cause, since indexed commodities could have been endogenously selected into an index, precisely because they are exposed to the same fundamental shocks. In a review article, Cheng and Xiong (2014) write "direct tests of price impacts and impacts on correlations should incorporate clear identification strategies in the spirit of Angrist and Pischke (2010)."

Our paper fills this gap. Our main variable of interest is the daily return autocorrelation instead of return correlation of different commodities. When we do that, we observe a clear divergence between the indexed commodity portfolio and the non-indexed commodity portfolio, as evident in Figure 2. With a backward rolling window of ten years, we do not observe a clear trend in the past 38 years in the daily autocorrelation in returns of the non-indexed commodity portfolio (NIDX). In sharp contrast, the daily autocorrelations in popular commodity indices (S&P Goldman Sachs Commodity Index (GSCI) and the Bloomberg Commodity Index (BCOM)) have steadily declined since 2004 when financialization began.[2] They entered the negative territory around

[1] We first calculate an equal weighted index for each sector of indexed and non-indexed commodities, then calculated the average correlation among five sector indices for an annual rolling window. Since there are no non-indexed commodities in energy and live cattle sectors, we take heating oil and RBOB and lean hogs as non-indexed commodities due to their small weights in the index.

[2] GSCI was originally developed in 1991, by Goldman Sachs. In 2007, ownership was transferred to Standard &

2005 and became significantly negative since 2006. While a declining index return autocorrelation can be consistent with improved information efficiency when common fundamental shocks are simultaneously and efficiently incorporated into the prices of multiple indexed commodities, *a negative return autocorrelation unambiguously signals inefficiency of price discovery.* It suggests that prices across multiple indexed commodities can overshoot and subsequently revert at the same time, resulting in “excessive co-movement,” even at the index level. Our paper hence provides empirical evidence to the theoretical hypothesis by Goldstein and Yang (2017), which writes ”The trading of index traders can inject both fundamental information and unrelated noise into the futures prices. Therefore, market efficiency can either increase or decrease with financialization.” The negative autocorrelation certainly demonstrates a negative side effect of the financialization, likely due to unrelated noise injected into the futures markets.

[Figure 2 is about here.]

Negative return autocorrelation at daily frequency is hard to explain using fundamental factors. For example, common discount rate or risk premium variations which can also cause negative return autocorrelations tend to operate at business cycle frequency. Instead, we attribute it to financialization and the resulting commodity index trading that propagate “non-fundamental shocks” from some commodities in the index to the rest. Figure 3 provides some supporting evidence. We plot rolling average daily return autocorrelations with a shorter backward window of three years against a measure of exposure to index trading, in a shorter sample starting from 2006 (due to the availability of the indexing exposure measure). We see a clear negative relation between the index autocorrelation and the index exposure measure. In other words, when institutional investors are trading commodity index more actively, the commodity index return (on both GSCI and BCOM) becomes more negatively autocorrelated. No such relation is observed for the portfolio

Poor’s. BCOM was originally launched in 1998 as the Dow Jones-AIG Commodity Index (DJ-AIGCI) and renamed to Dow Jones-UBS Commodity Index (DJ-UBSCI) in 2009, when UBS acquired the index from AIG. On July 1, 2014, the index was rebranded under its current name.

of non-indexed commodities (NIDX).

[Figure 3 is about here.]

The rest of the paper provides additional evidence that pins down the link between financialization and excessive return co-movement among indexed commodities.

We ran three sets of tests. In the first, we directly measure daily sentiment on a commodity derived from its news articles. Specifically, the sentiment measure is constructed as the residual from orthogonalizing news tone of articles about a commodity on fundamental factors of the same commodity. We then study the spillover of such sentiment across indexed commodities. Take an indexed commodity, corn, as an example. We compute the "connected" index sentiment by averaging the sentiment measures on other non-Grains indexed commodities (such as energy, metal, etc.) using institutional investors' total exposure to that commodity as the weight. We find that the "connected" index sentiment is related to contemporaneous return on corn positively and significantly, but to negatively and significantly predict corn's return tomorrow. The positive contemporaneous correlation could suggest that sentiment is propagated from some commodities to others in the same index. It could also suggest that our sentiment measure may still contain common fundamental factors. Nevertheless, the fact that such a positive correlation reverts on the next day confirms the existence of "non-fundamental" shocks. As index trading propagates such "non-fundamental" shocks across commodities in the same index, it results in synchronized price overshoots and reversals and therefore "excessive" co-movement. We confirm that the results are not driven by the 2008-2009 great financial crisis. In fact, the results are stronger after excluding the financial crisis period. As a placebo test, we repeat the same tests among non-indexed commodities but do not find evidence for such "non-fundamental" shocks.

Our second set of tests directly link "excessive" co-movement to index trading. We first confirm that the index sentiment propagation results are much stronger during periods when the index is more exposed to institutional trading. More formally, we also regress daily autocorrelation of indexed commodities on the abnormal index exposure measure and find a significantly negative coefficient. Abnormal index exposure does not affect the autocorrelation on the return of non-indexed commodities at all.

Our third test aims at establishing causality from commodity index trading to excessive comovement in the commodity index. We take advantage of the fact that the same commodity can receive different weights across two popular commodity indices (GSCI and BCOM). The relative weight difference arises in a rather ad-hoc fashion and is determined in the beginning of each year. We find that the negative daily return autocorrelation on commodities overweighted in GSCI (relative to BCOM) correlates more with the excessive exposure to ETFs based on GSCI (relative to that based on BCOM).

Our paper links to three strands of literatures. First, it adds value to the debates of price impact of the index investments. Using a theoretical model, Basak and Pavlova (2016) show that the excess correlation among commodities can arise if institutional investors care about outperforming a commodity index. Sockin and Xiong (2015) theoretically show that financial inflows and outflows (through index investing) to commodity markets can be misread as a signal about global economic growth if informational frictions exist in the commodity future markets. Singleton (2013) and Gilbert (2010) show that index investments do predict movements of commodity prices. Mou (2011) and Yan *et al.* (2019) document that index rebalancing causes futures prices shift significantly. Henderson *et al.* (2014) document that the hedging activities of issuers of commodity-linked notes can significantly influence commodity futures prices. Brogaard *et al.* (2018) documents that firms using commodity indices have relatively worse performance, and hence index investing distorts the price signal in commodity markets. However, Büyükşahin and Harris (2011), Irwin and Sanders (2012) and Sanders and Irwin (2011) find little evidence that the index position changes link to price movements in futures markets. Hamilton and Wu (2015) presents a mixed result. Our paper adds value to this debate by presenting a supporting evidence on the price pressure from the index investment. Particularly, prices of indexed commodities over-shoot and reverse subsequently when reacting to non-fundamental sentiment shocks, while non-indexed commodities do not show such a pattern.

Second, our paper relates to the study of co-movements among different commodities. For example, Pindyck and Rotemberg (1990) document a co-movement of unrelated commodities and attribute it to common effects of macroeconomic variables. However, Ai *et al.* (2006) argues that the co-movement among commodities is not excessive, and can be explained by common tendencies in demand and supply factors. Different with these studies, our results suggest that commodity index trading can propagate price pressure across commodities in the same index. To clarify, while such a price pressures results in "excessive" co-movement at the index level, we do not claim it to drive the boom and bust commodity cycles entirely. Indeed, our results seem to be stronger after excluding the greater financial crisis from our analysis.

Third, our paper also speaks to existing literature that links indexing to side effects, mostly in equity markets. Such side effects include the amplification of fundamental shocks (Hong *et al.*, 2012), non-fundamental price changes (Chen *et al.*, 2004), excessive comovement (Barberis *et al.*, 2005; Greenwood, 2005, 2008; Da and Shive, 2018), a deterioration of the firms information environment (Israeli *et al.*, 2017), increased non-fundamental volatility in individual stocks (Ben-David *et al.*, 2017), and reduced welfare of retail investors (Bond and García, 2017). Our results indicate that similar side effects may exist in the commodity market as well.

The remainder of the paper goes as follows. Section 2 describes the data and constructs variables used in this research. Section 3 presents the empirical results, and section 4 concludes.

2 Data and Variable Construction

In this section, we describe the commodities used in our analyses and introduce two most popular commodity indices and their construction. We then describe how we measure the exposure of a commodity to indexing. Finally, we discuss our news database and how we construct a news-based sentiment measure for each commodity.

2.1 Commodities and commodity indices

Commodity price data are obtained from Pinnacle Corporation. Following Kang *et al.* (2019), we compute the daily excess return for each commodity using the nearest-to-maturity (front-month) contract and we roll positions on the 7th calendar day of the maturity month into the next-to-maturity contract.[3] The excess return r_{it} on commodity i on date t is calculated as:

$$r_{it} = \frac{F_i(t,T) - F_i(t-1,T)}{F_i(t-1,T)}. \tag{1}$$

where $F_i(t,T)$ is the futures price on day t for a futures contract maturing on date T.

Table 1 lists the 27 commodities we examined. They are categorized into five sectors: Energy, Grains, Livestock, Metals, and Softs. Futures listing exchanges and coverage periods are also provided for each commodity.

[Table 1 is about here.]

The recent financialization makes it easy for institutional investors to trade various commodity indices. A commodity index functions like an equity index, such as the S&P 500, in which its value is derived from the total value of a specified basket of commodities. Currently, the largest two indices by market share are the S&P Goldman Sachs Commodity Index (GSCI) and the Bloomberg Commodity Index (BCOM). These two indices use different selection criteria and weighting schemes: the GSCI is weighted by the world production of each commodity, whereas the BCOM focuses on the relative amount of trading activity of a particular commodity. Importantly, for both indices, the weights are set in the beginning of the year and do not vary during the year. Table 1 provides index membership information for each of the 27 commodities in our sample.

Investors can use three types of financial instruments to gain exposure to a commodity index: commodity index swaps, exchange-traded funds (ETF), and exchange-traded notes (ETN). We

[3] If the 7th is not a business day, we use the next business day as our roll date.

collect the daily price data of GSCI and BCOM from Yahoo finance and calculate their daily returns as $(P_t - P_{t-1})/P_{t-1}$. We also construct an equal-weighted non-indexed commodities index (NIDX) and calculate its daily returns by simply equally averaging the daily returns across non-indexed commodities. Table 2 provides summary statistics regarding daily returns on individual commodities and commodity indices.

[Table 2 is about here.]

Table 2 highlights the benefit of investing in the commodity market. Commodities offer attractive annual Sharpe ratios that are comparable to that in the equity market. More importantly, their return correlations with the equity market (proxied using the S&P 500 index) are fairly low with an average correlation of 0.16, bringing in additional diversification benefit. Take Gold for example, its annual Sharpe ratio is 0.47 and its return correlation with the equity market is almost zero in our sample period. Not surprisingly, given these attractive features, institutional investors became more willing to invest in commodities, especially since the start of financialization that makes it easy for them to trade commodity indices.[4]

Energy sector, especially crude oil (CL) and natural gas (NG), did not perform well in our sample period from 2003 to 2015. Since both GSCI and BCOM indices place heavy weights in the energy sector, both indices suffered losses in the same period.

2.2 Commodity index exposure

The exposure of a commodity to index trading by institutional investors (or index exposure in short), is defined as the total market cap of index trading on that commodity as the percentage of total market cap of all trading on that commodity. Each Tuesday, the CFTC releases a weekly

[4]Tang and Xiong (2012) argue that the capital inflow into commodity futures markets integrates the segmented commodity futures markets with mainstream financial markets, for example the equity markets; particularly they show an increasing correlation between commodity and equity indexes especially during the financial crisis. However, the correlation declines in recent years (Bhardwaj *et al.*, 2015) likely caused by the capital outflow from the commodity markets. The overall correlation between GSCI and SP500 is around 0.3 in our sample.

Commitments of Traders (CoT) report, which includes the total open interest of each commodity and the long/short positions of each type of traders.[5] It also includes a supplemental Commodity Index Trader (CIT) report that shows the positions of a set of index traders identified by the CFTC since January 3, 2006. According to the manual of CIT, the total open interest in the supplementary CIT report can be recovered from the 9 components that are detailed in the report:

$$2(\mathit{Open\ Interest}^{All}) = \underbrace{(\mathit{Long} + \mathit{Short} + 2\mathit{Spread})}_{\mathit{Non\text{-}commercial}} + \underbrace{(\mathit{Long} + \mathit{Short})}_{\mathit{Commercial}} + \underbrace{(\mathit{Long} + \mathit{Short})}_{\mathit{Index\ Trading}} + \underbrace{(\mathit{Long} + \mathit{Short})}_{\mathit{Non\text{-}reportable}}. \quad (2)$$

In this paper, we define the index open interest as the average of the long and short positions of index traders: $\mathit{Open\ Interest}^{Idx} = (\mathit{Long}^{Idx} + \mathit{Short}^{Idx})/2$. Therefore, we can derive the dollar value open interests (Market Cap) on index/total trading as:

$$\mathit{Market\ Cap}_t^{Idx} = \sum_{i=1} \mathit{Open\ Interest}_{it}^{Idx} \times \mathit{Contract\ Size}_i \times \mathit{Price}_{it}, \quad (3)$$

$$\mathit{Market\ Cap}_t^{All} = \sum_{i=1} \mathit{Open\ Interest}_{it}^{All} \times \mathit{Contract\ Size}_i \times \mathit{Price}_{it}. \quad (4)$$

Unfortunately, the CIT only reports 13 agricultural commodities (listed in Table 1) and it covers no commodities in the energy and metals sectors. Masters (2008) and Hamilton and Wu (2015) proposed to estimate the unreported index trading positions by making use of the reported data and their weights in each commodity index. Taking crude oil (CL) as an example, the general idea of Masters (2008) is to use the fact that both GSCI and BCOM have their own uniquely included commodities, i.e. soybean oil (BO) and soybean meal (SM) in BCOM[6] and cocoa (CC), feeder cattle (FC) and Kansas wheat (KW) for GSCI. Then, note that index traders replicate the index by allocating across the commodities according to the known weights[7] $\delta_{jy(t)}^{(i)}, i \in \{G, B\}$, we can separately estimate CL's dollar value long/short positions on index trading, $X_{CL,t}$, on GSCI/BCOM

[5] The traders are classified into three types: commercial (C), noncommercial (NC), and non-reportables (NR). In CIT report, CFTC separates the index trading positions (Idx) from the positions of the commercial traders.

[6] Note that soybean meal (SM) was added to BCOM in 2013.

[7] Both weights reported in the GSCI and BCOM manuals are dollar value weights.

trading as below:

$$\widehat{X}^B_{CL,t} = \begin{cases} \dfrac{\delta^B_{CL,y(t)}}{\delta^B_{BO,y(t)}} X_{BO,t}, & \text{if } y(t) < 2013, \\ \dfrac{1}{2}\left(\dfrac{\delta^B_{CL,y(t)}}{\delta^B_{BO,y(t)}} X_{BO,t} + \dfrac{\delta^B_{CL,y(t)}}{\delta^B_{SM,y(t)}} X_{SM,t}\right), & \text{if } y(t) \geq 2013. \end{cases} \tag{5}$$

$$\widehat{X}^G_{CL,t} = \frac{1}{3}\left(\frac{\delta^G_{CL,y(t)}}{\delta^G_{CC,y(t)}} X_{CC,t} + \frac{\delta^G_{CL,y(t)}}{\delta^G_{FC,y(t)}} X_{FC,t} + \frac{\delta^G_{CL,y(t)}}{\delta^G_{KW,y(t)}} X_{KW,t}\right). \tag{6}$$

where $y(t)$ denotes the year of t. Note that the weights of commodities in an index are determined at the beginning of a year and stay the same during the year. Thus, the dollar value of index trading for commodity i at time t is estimated as

$$X_{it} = Position_{it} \times ContractSize_i \times Price_{it}. \tag{7}$$

Combining the estimates above, Masters (2008) propose to estimate the total market cap of CL on index trading as:

$$\widehat{Market\ Cap}^{Idx}_{CL,t} = \widehat{Market\ Cap}^{B}_{CL,t} + \widehat{Market\ Cap}^{G}_{CL,t}. \tag{8}$$

However, as pointed out by Irwin and Sanders (2011), Masters' estimator is severely biased when there is a huge difference between $\frac{\delta^G_{CL,y(t)}}{\delta^G_{CC,y(t)}} X_{CC,t}$, $\frac{\delta^G_{CL,y(t)}}{\delta^G_{FC,y(t)}} X_{FC,t}$ and $\frac{\delta^G_{CL,y(t)}}{\delta^G_{KW,y(t)}} X_{KW,t}$. To deal with this issue, Hamilton and Wu (2015) propose to generalize Masters' method by using all the reported commodities' positions for estimation. Specifically, they choose $\widehat{X}^G_{it}$ and $\widehat{X}^B_{it}$ to minimize the sum of squared discrepancies in predicting the CIT reported value for X_{it} across 12 commodities. Thus, the estimated dollar value positions on index trading for commodity i in day t is given by

$$\widehat{X}^{Idx}_{it} = \begin{bmatrix} \delta^G_{iy(t)} & \delta^B_{iy(t)} \end{bmatrix} \begin{bmatrix} \sum\limits_{j\in CIT} \left(\delta^G_{jy(t)}\right)^2 & \sum\limits_{j\in CIT} \delta^G_{jy(t)}\delta^B_{jy(t)} \\ \sum\limits_{j\in CIT} \delta^B_{jy(t)}\delta^G_{jy(t)} & \sum\limits_{j\in CIT} \left(\delta^B_{jy(t)}\right)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum\limits_{j\in CIT} \delta^G_{jy(t)} X^{Idx}_{jt} \\ \sum\limits_{j\in CIT} \delta^B_{jy(t)} X^{Idx}_{jt} \end{bmatrix}, \tag{9}$$

where $\delta_{jy(t)}$ is the weight of a commodity j in a certain index in year $y(t)$, and the superscripts G and B denote the index GSCI and BCOM, respectively. From Equation (9) we obtain both the long and short dollar value positions for unreported commodities, and thus the total market cap. Combining the CIT-reported open interest on index trading data, we can estimate the daily market cap of the total (index) trading by

$$\widehat{Market\ Cap}_t^{Idx} = \sum_{j \in Idx} \widehat{Market\ Cap}_{jt}. \tag{10}$$

Figure 4 plots the market capital of index traders. The figure shows that before the financial crisis and around 2011, the market capital of index trading reached its highest level. It trended down afterwards.

[Figure 4 is about here.]

Then, the index exposure is defined as

$$Indexing_t = Market\ Cap_{it}^{Idx} \Big/ Market\ Cap_{it}^{All}, \tag{11}$$

To study the impact of index trading on commodities returns, we finally propose a flow measure, abnormal index exposure, by evaluating how much incremental market cap on index trading contributes to the total market cap traded, i.e.,

$$Abn.\ Indexing_t = \frac{Market\ Cap_t^{Idx} - Market\ Cap_{t-1}^{Idx}}{Market\ Cap_t^{All}}. \tag{12}$$

2.3 Commodity sentiment measure

The news data we use come from the Thomson Reuters News Analytics - Commodities data (TRNA-C). TRNA-C data provides 3 news tones (positive, negative and neutral) for each piece of commodity news and the sample coverage starts from January 2003.[8] By averaging all the news tones on each piece of news in a trading day for each commodity, we obtain a daily panel of 3 news tones for each commodity.

For each commodity, we first compute the net tone as the difference between the positive tone and the negative tone. We then calculate the abnormal net tone as the residual of regressing the net tones on its first lag and the month dummies. Finally, to extract news sentiment, we then orthogonalize the abnormal net tones on commodity fundamentals such as returns, basis and liquidity. The reasons to include those controls are as follows: as shown in Brennan (1958) and Gorton *et al.* (2012) basis mainly represents the level of inventory, which can be considered as the mismatch between demand and supply of a certain commodity; Szymanowska *et al.* (2014) have shown that basis is a determinant of commodity risk premium. Since hedging activity from production firms may cause extra trading in futures markets, we include Amihud illiquidity as a control in our regression, which is considered as the best liquidity measure in commodity markets (Marshall *et al.* (2011)). Specifically, for each commodity, we run the following regression:

$$Abn.\ Net\ Tone_t = \alpha + \beta' \begin{bmatrix} r_t \\ r_{t-1} \end{bmatrix} + \theta' \begin{bmatrix} Basis_t \\ Basis_{t-1} \end{bmatrix} + \phi' \begin{bmatrix} Illiquidity_t \\ Illiquidity_{t-1} \end{bmatrix} + \epsilon_t, \tag{13}$$

where *Basis* is the log basis[9] and *Illiquidity* is the Amihud (2002) illiquidity measure.[10] We then treat the residual of the regression as the sentiment measure for each commodity. The descriptive statistics of our sentiment measure for each commodity are shown in Table 3.

[Table 3 is about here.]

[8] According to the TRNA-C manual, the news tones are calculated base on neural network algorithm and the reported accuracy is around 75%.

[9] The log basis is defined as $\frac{\ln(F_i(t,T_1))-\ln(F_i(t,T_2))}{T_2-T_1}$, where $F_i(t,T_1)$ and $F_i(t,T_2)$ are the prices of the closest and next closest to maturity contracts for commodity i.

[10] For each commodity, we compute its Amihud's (2002) illiquidity measure as the ratio of the absolute value of its daily return divided by its dollar trading volume in the same day.

As evident in Table 3, crude oil receives more news coverage than other commodities.[11] The sentiment measures have zero means by construction. Their average standard deviations is 0.1069 ranging from 0.0522 for oat (O-) to 0.1953 for soybean (S-).

3 Empirical Analysis

We conduct three sets of empirical analyses. We first study the propagation of "non-fundamental" shocks across commodities using our news-based sentiment measure. We then examine daily return autocorrelations for commodity indices and relate them to measures of their index exposure. Finally, we provide causal evidence that index trading drives negative index return autocorrelations.

3.1 Sentiment Spillover

To study the sentiment spillover across the indexed commodities, we construct a "connected" sentiment measure for each commodity. Take corn (C-) for example. To construct its "connected" sentiment on day t, we take a weighted average of sentiment measures on all other indexed commodities from other sectors on that day, i.e.,

$$Cnn.\ Sentiment_{it} = \sum_{S(j)\neq S(i)} W_{jy(t)} Sentiment_{jt}, \tag{14}$$

where $S(i)$ is the sector that commodity i belongs to, and the weight $W_{jy(t)}$ is defined as

$$W_{jy(t)} = \frac{E_{y(t)}(\$Open\ Interest^{Idx}_{jt})}{\sum_j E_{y(t)}(\$Open\ Interest^{Idx}_{jt})}, \tag{15}$$

with $E_{y(t)}(\$Open\ Interest^{Idx}_{jw(t)})$ being the average of the weekly dollar-valued open interest on index trading in year $y(t)$. In other words, the weight on "connected" indexed commodity j is determined

[11] Since oat and rough rice are close substitutes, the news tone in our dataset treats them as one commodity; we hence use identical news tone for both oat and rough rice.

by its average dollar-valued open interest relative to total dollar-valued open interests across both indices.

In the above definition, the set of indexed commodities "connected" to corn only includes indexed commodities from other sectors such as energy and metals, but not other indexed commodities from the same grains sector such as soybean (S-) and wheat (W-). To the extent that sentiment measure includes commodities from the same sector may still contain fundamental factors, they are more likely to co-move within sector than across sectors. In this sense, our measure alleviates the concerns for fundamental-driven co-movement. As a placebo test, we construct the "connected" sentiment measure for non-indexed commodities in the same fashion, except that we use an equal weighting scheme as in the construction of NIDX.

Based on the "connected" sentiment measure, we run the following day / commodity panel regressions to examine both contemporaneous and predictive relations between the "connected" sentiment measure and the commodity returns, for indexed and non-indexed commodities separately,

$$r_{it} = \beta_0 + \beta_1 \cdot \mathit{Cnn.\ Sentiment}_{it} + \theta' \boldsymbol{X}_{it-1} + \epsilon_{it}, \tag{16}$$

$$r_{it} = \beta_0 + \beta_1 \cdot \mathit{Cnn.\ Sentiment}_{it-1} + \theta' \boldsymbol{X}_{it-1} + \epsilon_{it}, \tag{17}$$

where $\boldsymbol{X}$ is a vector of control variables including lagged returns, lagged log basis, and lagged Amihud's (2002) illiquidity measure for each commodity. We also include the lagged change in implied volatility of crude oil options with nearest maturity in the regressions to control for systematic volatility shock in commodity markets in the spirit of Christoffersen and Pan (2018).[12] Both individual/sector fixed effect and year fixed effect are controlled in the regression. The results are reported in Table 4.

[12] Christoffersen and Pan (2018) shows that shocks to oil volatility are strongly related to various measures of funding constraints of financial intermediaries, which is arguably a key driver of pricing kernel dynamics.

[Table 4 is about here.]

Focusing on Panel A, we find a positive and significant contemporaneous relation between the indexed commodity return and its "connected" sentiment measure. The relation is consistent with the notion that index trading could propagate sentiment across commodities within the same index. It may also reflect fundamental-driven co-movement that are not fully controlled for using return, basis and liquidity measures in our sentiment measure construction. Indeed, we also observe positive and significant coefficients on "connected" sentiment measures for non-indexed commodities even though index trading is not possible here.

While both sentiment spillover and common fundamental factor can explain the positive contemporaneous relation we observed in Panel A, only sentiment predicts future return negatively. This is because sentiment-induced trading induces a "non-fundamental" shock or price pressure in contemporaneous return and such a shock will be reverted in the future. For example, positive sentiment on energy may induce institutional investors to buy the commodity index. Such trading propagates the sentiment from energy sector to other indexed commodities and results in a positive price pressure in corn today, explaining the positive contemporaneous relation between the return on corn and its "connected" sentiment. As the positive price pressure on corn reverts tomorrow, the "connected" sentiment today should negatively predict corn's return tomorrow.

Such a negative return predictability by "connected" sentiment is exactly what we find in Panel B for indexed commodities. The coefficient on "connected" sentiments is likely to capture the impact of sentiment spillover. For instance, a coefficient of -0.0052 (t-value of -1.86) on the "connected" sentiment implies that a one-standard-deviation increase in the sentiment of "connected" indexed commodities propagate a price pressure of 1.9 basis points. The last column in Panel B does not find significant return predictability among non-indexed commodities, again consistent with the notion that sentiment is mainly propagated via index trading.

Figure 1 highlights the well-known commodity boom and bust cycles during the 2008-2009 financial crisis. A natural concern is whether our results so far are completely driven by such market wide fluctuations. To that end, we repeat in Table 5 our previous analyses, but after excluding the great financial crisis (2008/09/15 to 2009/06/30 according to the bankruptcy of Lehman Brothers and NBER recession dating as in Tang and Xiong (2012)). We find our results to be similar, if not stronger, after removing the great financial crisis. For example, the "connected" sentiment measure is now associated with a bigger coefficient (in absolute term) of -0.0072 in predicting the return reversal on the next day. In other words, our results so far are more consistent with high-frequency (daily) price overshoots and reversals than low-frequency (business cycle) variation in fundamental risk factors.

[Table 5 is about here.]

Turning to the control variables, consistent with Szymanowska *et al.* (2014), lagged basis makes a positive prediction (although insignificant) on commodity returns listed in Table 4. The coefficient of illiquidity also agrees with Kang *et al.* (2019), i.e. in an illiquid market hedgers have to pay a higher liquidity premium to speculators in order to execute their hedging. Past returns positively predict futures returns for non-indexed commodities, likely caused by under-reaction, since non-indexed commodities are relatively illiquid, and hence their futures markets are not efficient in transmitting fundamental information. This can be seen from Table 2, where non-indexed commodities normally have positive autocorrelations. However, indexed commodities such as energy and metals often have negative autocorrelation, which explains the negative coefficient of lagged returns in Table 4 and 5. Implied volatility shocks of crude oil significantly predict commodity returns of indexed commodities, but with less strength for non-indexed commodities.

If index trading propagates sentiment and creates price pressure at the index level, we should observe stronger effect during times when index trading exposure is high. To test this conjecture, we split the sample into two subsamples based on our abnormal index exposure measure defined in the previous section. Specifically, we first calculate each week's average abnormal index exposure.

"High" of H ("Low" or L) index exposure period includes weeks whose average abnormal index exposure measure is above (below) the median of the full sample. We then re-run the previous regression analyses in the "H" and "L" subperiods separately:

$$r_{it} = \beta_0 + \beta_1 \cdot Cnn.\ Sentiment_{it} + \theta' \boldsymbol{X}_{it-1} + \epsilon_{it}, \quad t \in \{H, L\} \tag{18}$$

$$r_{it} = \beta_0 + \beta_1 \cdot Cnn.\ Sentiment_{it-1} + \theta' \boldsymbol{X}_{it-1} + \epsilon_{it},, \quad t \in \{H, L\} \tag{19}$$

where $w(t)$ denotes the week of date t. Both individual/sector fixed effect and year fixed effect are controlled in the regression. The results are reported in Table 6.

[Table 6 is about here.]

Focusing on the sentiment return predictability results in Panel B, we find that the return reversal is only significant during the "High" period for the indexed commodities. The coefficient on the sentiment measure is -0.0146 (t-value of -3.55) during months with abnormally high amount of index trading. The economic magnitude is large. A coefficient of -0.0146 implies that a one-standard-deviation increase in the sentiment of "connected" indexed commodities propagate a price pressure of 5.5 basis points. Consistent with the notion that index trading results in price overshoot and reversal, when we focus our attention on non-indexed commodities, we observe no return reversals in either"High-" or "Low-" index exposure period.

So far, our results using news-based sentiment measures provide supporting evidence that as index trading propagates "non-fundamental" shocks across commodities in the same index, it creates correlated price overshoots and reversals at daily frequency. Such excessive co-movements will result in negative daily return autocorrelations even at the index level. In the next two subsections, we therefore focus our attention on index daily return autocorrelation measures.

3.2 Return autocorrelation and index exposure

Focusing on the daily return autocorrelation, Figures 2 and 3 confirm the link between index trading and excessive co-movement in the commodity market. With a backward rolling window of ten years, we observe a continuing decline in the average index daily return autocorrelations for both GSCI ad BCOM indices in Figure 2. They became significantly negative since 2006 when financialization made index trading easy. A negative return autocorrelation unambiguously signals excessive co-movement and price inefficiency at the index level. In sharp contrast, no such decline is observed for the average daily return autocorrelation for a portfolio of non-indexed commodities (NIDX). It is almost always positive throughout the sample period.

Using our index exposure measure, Figure 3 shows that the gradual decline in the index return autocorrelation is accompanied with a rising exposure to index trading during the same sample period since 2006. Of course, "missing" factors can drive the common trend in the low-frequency variations of both daily index return autocorrelation and index exposure.

In this subsection, we examine and confirm the linkage between index trading and index return autocorrelation at daily frequency, taking advantage of the high-frequency nature of our measures. In the first test, we check if the autocorrelation coefficient of the returns will be decreasing when abnormal index exposure increases. In particular, we include an interaction term between lagged return and abnormal index exposure in the autoregressive model of commodity returns:

$$r_{it} = \beta_0 + \beta_1 \cdot r_{i,t-1} + \beta_2 \cdot Abn.\ Indexing_{t-1} + \gamma \cdot (r_{i,t-1} \times Abn.\ Indexing_{t-1}) + \theta' \boldsymbol{X}_{i,t-1} + \epsilon_{it}, \quad (20)$$

where the vector $\boldsymbol{X}$ contains each commodity's lagged basis, lagged Amihud illiquidity, and lagged change in implied volatility of crude oil options with nearest maturity as control variables in the spirit of Nagel (2012) and Bianchi *et al.* (2016). The key parameter of interest is γ and the results are presented in Table 7.

[Table 7 is about here.]

Table 7 reveals three sets of interesting results. First, we observe negative and significant coefficients on the interaction term only for the indexed commodities. In other words, abnormally high index trading today implies a negative correlation between the indexed commodity return today and that tomorrow, consistent with the notion that index trading results in price pressure at the index level today and such a price pressure is reverted tomorrow. Second, we find index exposure measure to have nothing to do with the return autocorrelation of non-indexed commodities, consistent with the pattern in Figure 3. This result serves as a nice placebo. Finally, the result is still significant, when excluding the financial crisis period.

Complementary to interaction results, we use abnormal index exposure to directly predict the return serial dependence following Baltussen *et al.* (2018). Specifically, we run the following daily panel regression of each commodity's serial dependence measure on the lagged abnormal index exposure measure and controls:

$$(r_{it} \cdot r_{i,t-1})/2\sigma_i^2 = \beta_0 + \beta_1 \cdot \textit{Abn. Indexing}_{t-1} + \theta' \boldsymbol{X}_{i,t-1} + \epsilon_{it}, \tag{21}$$

where σ_i^2 is the sample variance of commodity i's returns, and vector $\boldsymbol{X}$ contains each commodity's lagged serial dependence measure, lagged basis, lagged Amihud illiquidity, and lagged change in implied volatility of crude oil options with nearest maturity as control variables in the spirit of Baltussen *et al.* (2018), Nagel (2012) and Bianchi *et al.* (2016). We run the panel regression for indexed and non-indexed commodities separately. The results are presented in Table 8.

[Table 8 is about here.]

Consistent with Table 7, we observe negative and significant coefficients on the index exposure measure only for the indexed commodities. The economic magnitude of such effect is large. For example, coefficient of -6.2068 means that a one-standard-deviation increase in the abnormal index exposure makes its daily return autocorrelation -2.55% more negative. Lastly, the result remains significant after excluding the great financial crisis from our analysis.

3.3 Causal evidence

Can some missing factors drive the link between index trading and negative daily return autocorrelation we documented so far? Maybe in the last 15 years, institutional investors simply became more willing to invest in a basket of certain commodities as an asset class. Such an investment demand will result in correlated order flow across these commodities and will result in negative commodity portfolio return autocorrelations, regardless whether commodities index products have been introduced or not. It is simply a coincidence that part of that correlated order flow is also satisfied through index products (rather than through trading the underlying commodity futures directly). One could even argue that the commodity indexed products was introduced precisely to cater for correlated demand from institutional investors in trading these commodities (that are chosen to be included in GSCI and BCOM indices).

While such a correlated demand story could explain the low-frequency trends displayed in Figure 3, it is harder to explain the high-frequency relation (between index trading and negative daily return autocorrelation) we documented in Table 7 and 8. This is no reason to believe that a broadly increasing trend to invest in a general commodity basket should be correlated with abnormal trading activities in two specific commodity indices on a day-to-day basis. Nevertheless, in this subsection, we conduct additional tests, aiming at pinning down the causality from index trading on index return autocorrelation.

The additional causality tests are similar in spirit to those in Greenwood (2008) and Baltussen *et al.* (2018) that take advantage of different weighting schemes across two Japanese equity indices. Similar to the case of equity indices, the same commodity can receive different weights across GSCI and BCOM indices. The relative weighting is determined in a fairly ad-hoc fashion, and importantly for our purpose, is determined at the beginning of the year and then held constant throughout the year. A testable implication of index trading therefore goes as follows: for a portfolio of commodities that are overweighted in GSCI index (relative to BCOM index), its daily return autocorrelation should be more negatively correlated with trading measure on GSCI (relative to that on BCOM).

We implement the test by constructing zero-investment portfolio ("GSCI/BCOM OW portfolio"). Take GSCI OW portfolio as an example, we first compare commodity i's weight in GSCI, $w_{jy(t)}^{GSCI}$, to its weight in BCOM, $w_{jy(t)}^{BCOM}$:

$$OW_{jy(t)}^{GSCI} = \begin{cases} w_{jy(t)}^{GSCI} - w_{jy(t)}^{BCOM}, & \text{if } w_{jy(t)}^{GSCI} > 0, \\ 0, & \text{if } w_{jy(t)}^{GSCI} = 0, \end{cases} \tag{22}$$

where $y(t)$ is the year of date t. Then, we assign weight ϖ_{jt}^{GSCI} on each commodity j as

$$\varpi_{jt}^{GSCI} = \left(OW_{jy(t)}^{GSCI} - \frac{1}{N} \sum_{j=1}^{N} OW_{jy(t)}^{GSCI} \right) r_{t-1}^{GSCI}, \tag{23}$$

and calculate the portfolio return $R_t^{GSCI} = \sum_{j=1}^{N} \varpi_{jt}^{GSCI} r_{jt}$, where r_{t-1}^{GSCI} is the return on GSCI index and r_{jt} is the return on commodity j. A portfolio that bases on the commodities that are overweighted in BCOM can be constructed in the same way.

Next, we compute the ETF indexing measure for each commodity index. Specifically, we first compute the total market cap ($Shares\ Outstanding_t \times NAV_t$) of each commodity index's ETF/ETNs (i.e., GSG and GSP for GSCI and DJCI and DJP for BCOM). In the second step, we calculate the total market cap of commodities on GSCI/BCOM trading by estimating each commodity's dollar value trading positions on each index with modified Hamilton and Wu (2015) method as below:

$$\widehat{X}_{it}^{G} = \begin{bmatrix} \delta_{iy(t)}^{G} & 0 \end{bmatrix} \begin{bmatrix} \sum\limits_{j\in CIT} \left(\delta_{jy(t)}^{G}\right)^2 & \sum\limits_{j\in CIT} \delta_{jy(t)}^{G}\delta_{jy(t)}^{B} \\ \sum\limits_{j\in CIT} \delta_{jy(t)}^{B}\delta_{jy(t)}^{G} & \sum\limits_{j\in CIT} \left(\delta_{jy(t)}^{B}\right)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum\limits_{j\in CIT} \delta_{jy(t)}^{G} X_{jt}^{Idx} \\ \sum\limits_{j\in CIT} \delta_{jy(t)}^{B} X_{jt}^{Idx} \end{bmatrix}, \tag{24}$$

$$\widehat{X}_{it}^{B} = \begin{bmatrix} 0 & \delta_{iy(t)}^{B} \end{bmatrix} \begin{bmatrix} \sum\limits_{j\in CIT} \left(\delta_{jy(t)}^{G}\right)^2 & \sum\limits_{j\in CIT} \delta_{jy(t)}^{G}\delta_{jy(t)}^{B} \\ \sum\limits_{j\in CIT} \delta_{jy(t)}^{B}\delta_{jy(t)}^{G} & \sum\limits_{j\in CIT} \left(\delta_{jy(t)}^{B}\right)^2 \end{bmatrix}^{-1} \begin{bmatrix} \sum\limits_{j\in CIT} \delta_{jy(t)}^{G} X_{jt}^{Idx} \\ \sum\limits_{j\in CIT} \delta_{jy(t)}^{B} X_{jt}^{Idx} \end{bmatrix}. \tag{25}$$

The abnormal ETF indexing is then defined as a ratio of the index-related ETF/ETNs' total market cap changes to the total market cap of the index, i.e.,

$$Abn.\ ETF\ Indexing_t^{(i)} = \frac{ETF\ Market\ Cap_t^{(i)} - ETF\ Market\ Cap_{t-1}^{(i)}}{Market\ Cap_t^{(i)}}, \quad i \in \{GSCI, BCOM\}. \tag{26}$$

Then, we define the GSCI/BCOM relative ETF indexing measure as

$$Relative\ ETF\ Indexing_t^{(i)} = Abn.\ ETF\ Indexing_t^{(i)} - Abn.\ ETF\ Indexing_t^{(j)}, \tag{27}$$

where $i = GSCI, j = BCOM$ or $i = BCOM, j = GSCI$.

Finally, we regress the GSCI/BCOM OW portfolio's returns on the lagged GSCI/BCOM relative ETF indexing measure with controls, i.e.,

$$R_t^{(i)} = \beta_0 + \beta_1 \cdot Relative\ ETF\ Indexing_{t-1}^{(i)} + \theta' \boldsymbol{X}_{t-1} + \epsilon_t, \quad i \in \{GSCI, BCOM\}. \tag{28}$$

where $\boldsymbol{X}$ is a vector of control variables motivated by Bianchi *et al.* (2016), Nagel (2012) and Baltussen *et al.* (2018), which contains the lagged GSCI (BCOM) average return over the past 21 trading days, lagged realized GSCI (BCOM) volatility over the past 250 trading days, the lagged log GSCI (BCOM) related ETF/ETNs' trading volume detrended with one-year average log trading volume, and lagged implied volatility of crude oil options with nearest maturity. The results using the full sample and the sample excluding financial crisis for both portfolios are shown in Table 9.

[Table 9 is about here.]

The results in Table 9 strongly support a causal interpretation that index trading drives excessive co-movement and negative index return autocorrelation. For a portfolio of commodities that are relatively overweighted in index i, its daily return autocorrelation is indeed more negatively correlated with relative trading exposure to index i. The results hold for both GSCI and BCOM indices and after excluding the great financial crisis.

4 Conclusion

We examine the impact of recent financialization in the commodity market on excessive co-movement among indexed commodities. Using news-based sentiment measures, we find that index trading enabled by financialization can propagate non-fundamental shocks from some commodities to others in the same index, giving rise to price overshoots and subsequent reversals, or "excessive co-movement" at daily frequency. Excessive co-comovement results in negative daily commodity return autocorrelations even at the index level (but not for non-indexed commodities) and such autocorrelations move with our commodity index exposure measures. Taking advantage of the fact that index weights of the same commodity can vary across different indices in a relatively ad-hoc and pre-determined fashion, we provide causal evidence that index trading drives excessive co-movement. Such excessive co-movement could contribute to the boom-and-bust cycles observed during the recent financial crisis, even though it does not drive such cycles.

Given the attractive risk-return tradeoff and the diversification benefits associated with commodity index investments, the commodity financialization process can be expected to continue. We do not dispute such benefits. We simply highlight an unexpected side effect to these benefits as negative serial dependence in commodity index return signals excessive price co-movements even at the index level. Excessive price movement could impose costs on institutional investors who trade often and individual investors who invest in commodities through those institutions. In addition, as the recent paper by Brogaard *et al.* (2018) argues, inefficient commodity prices could even distort real decisions of a firm.

References

Ai, C., Chatrath, A. and Song, F. (2006). On the comovement of commodity prices. *American Journal of Agricultural Economics*, **88** (3), 574–588.

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets*, **5** (1), 31–56.

Angrist, J. D. and Pischke, J.-S. (2010). The credibility revolution in empirical economics: How better research design is taking the con out of econometrics. *Journal of Economic Perspectives*, **24** (2), 3–30.

Asness, C. S., Moskowitz, T. J. and Pedersen, L. H. (2013). Value and momentum everywhere. *The Journal of Finance*, **68** (3), 929–985.

Baltussen, G., van Bekkum, S. and Da, Z. (2018). Indexing and stock market serial dependence around the world. *Journal of Financial Economics, Forthcoming.*

Barberis, N., Shleifer, A. and Wurgler, J. (2005). Comovement. *Journal of Financial Economics*, **75** (2), 283–317.

Basak, S. and Pavlova, A. (2016). A model of financialization of commodities. *The Journal of Finance*, **71** (4), 1511–1556.

Ben-David, I., Franzoni, F. and Moussawi, R. (2017). Do etfs increase volatility? *Journal of Finance, Forthcoming.*

Bhardwaj, G., Gorton, G. and Rouwenhorst, G. (2015). Facts and fantasies about commodity futures ten years later. *NBER Working Paper.*

Bianchi, R. J., Drew, M. E. and Fan, J. H. (2016). Commodities momentum: A behavioral perspective. *Journal of Banking & Finance*, **72**, 133–150.

Bond, P. and García, D. (2017). Informed trading, indexing, and welfare. *Working Paper (University of Washington).*

Brennan, M. J. (1958). The supply of storage. *The American Economic Review*, **48** (1), 50–72.

Brogaard, J., Ringgenberg, M. and Sovich, D. (2018). The economic impact of index investing. *The Review of Financial Studies, Forthcoming.*

Büyükşahin, B. and Harris, J. H. (2011). Do speculators drive crude oil futures prices? *The Energy Journal*, pp. 167–202.

Chen, H., Noronha, G. and Singal, V. (2004). The price response to s&p 500 index additions and deletions: Evidence of asymmetry and a new explanation. *The Journal of Finance*, **59** (4), 1901–1930.

Cheng, I.-H. and Xiong, W. (2014). Financialization of commodity markets. *Annual Review of Finance and Economics*, **6** (1), 419–441.

Christoffersen, P. and Pan, X. N. (2018). Oil volatility risk and expected stock returns. *Journal of Banking & Finance*, **95**, 5–26.

Da, Z. and Shive, S. (2018). Exchange traded funds and asset return correlations. *European Financial Management*, **24** (1), 136–168.

Gilbert, C. L. (2010). How to understand high food prices. *Journal of Agricultural Economics*, **61** (2), 398–425.

Goldstein, I. and Yang, L. (2017). Commodity financialization and information transmission. *Rotman School of Management Working Paper No. 2555996*.

Gorton, G. B., Hayashi, F. and Rouwenhorst, K. G. (2012). The fundamentals of commodity futures returns. *Review of Finance*, **17** (1), 35–105.

Greenwood, R. (2008). Excess comovement of stock returns: Evidence from cross-sectional variation in nikkei 225 weights. *The Review of Financial Studies*, **21** (3), 1153–1186.

Greenwood, R. M. (2005). A cross sectional analysis of the excess comovement of stock returns. *HBS Finance Research Paper No. 05-069*.

Hamilton, J. D. and Wu, J. C. (2015). Effects of index-fund investing on commodity futures prices. *International Economic Review*, **56** (1), 187–205.

Henderson, B. J., Pearson, N. D. and Wang, L. (2014). New evidence on the financialization of commodity markets. *The Review of Financial Studies*, **28** (5), 1285–1311.

Hong, H., Kubik, J. D. and Fishman, T. (2012). Do arbitrageurs amplify economic shocks? *Journal of Financial Economics*, **103** (3), 454–470.

Irwin, S. H. and Sanders, D. R. (2011). Index funds, financialization, and commodity futures markets. *Applied Economic Perspectives and Policy*, **33** (1), 1–31.

— and — (2012). Testing the masters hypothesis in commodity futures markets. *Energy economics*, **34** (1), 256–269.

Israeli, D., Lee, C. M. C. and Sridharan, S. A. (2017). Is there a dark side to exchange traded funds? an information perspective. *Review of Accounting Studies*, **22** (3), 1048–1083.

Kang, W., Rouwenhorst, K. G. and Tang, K. (2019). A tale of two premiums: The role of hedgers and speculators in commodity futures markets. *Journal of Finance, Forthcoming.*

Marshall, B. R., Nguyen, N. H. and Visaltanachoti, N. (2011). Commodity liquidity measurement and transaction costs. *The Review of Financial Studies*, **25** (2), 599–638.

Masters, M. W. (2008). *Testimony before Committee on Homeland Security and Governmental Affairs of the United States Senate*. Tech. rep., Commodity Futures Trading Commission.

Miffre, J. and Rallis, G. (2007). Momentum strategies in commodity futures markets. *Journal of Banking & Finance*, **31** (6), 1863–1886.

Mou, Y. (2011). Limits to arbitrage and commodity index investment: front-running the goldman roll. *Working Paper, Columbia University.*

Nagel, S. (2012). Evaporating liquidity. *The Review of Financial Studies*, **25** (7), 2005–2039.

Pindyck, R. S. and Rotemberg, J. J. (1990). The excess co-movement of commodity prices. *The Economic Journal*, **100** (403), 1173–1189.

Sanders, D. R. and Irwin, S. H. (2011). The impact of index funds in commodity futures markets: A systems approach. *Journal of Alternative Investments*, **14** (1), 40–49.

Singleton, K. J. (2013). Investor flows and the 2008 boom/bust in oil prices. *Management Science*, **60** (2), 300–318.

Sockin, M. and Xiong, W. (2015). Informational frictions and commodity markets. *The Journal of Finance*, **70** (5), 2063–2098.

Szymanowska, M., De Roon, F., Nijman, T. and Van Den Goorbergh, R. (2014). An anatomy of commodity futures risk premia. *The Journal of Finance*, **69** (1), 453–482.

Tang, K. and Xiong, W. (2012). Index investment and the financialization of commodities. *Financial Analysts Journal*, **68** (5), 54–74.

Yan, L., Irwin, S. H. and Sanders, D. R. (2019). Is the supply curve for commodity futures contracts upward sloping? *Working Paper, UIUC.*

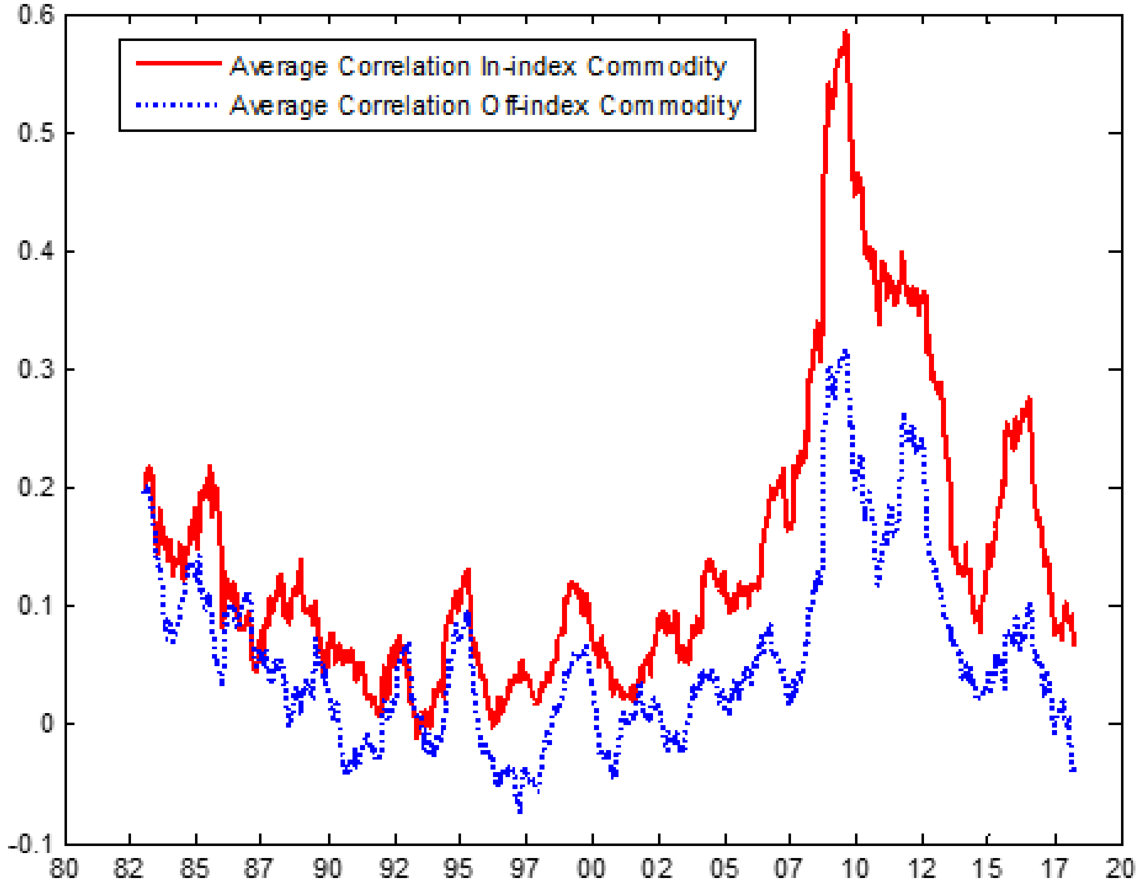


Figure 1: This figure plots the average return correlations of commodities in the GSCI and BCOM indices (indexed commodities) and those not included in these indices (non-indexed commodities). We follow the spirit of Tang and Xiong (2012) in computing these correlations. Specifically, we first calculate an equal weighted index for each sector of indexed and non-indexed commodities, then calculated the average correlation among five sector indices for an annual rolling window. Since there are no non-indexed commodities in energy and live cattle sectors, we take heating oil and RBOB and lean hogs as non-indexed commodities due to their small weights in the index. The sample period is from 1980 to 2018.

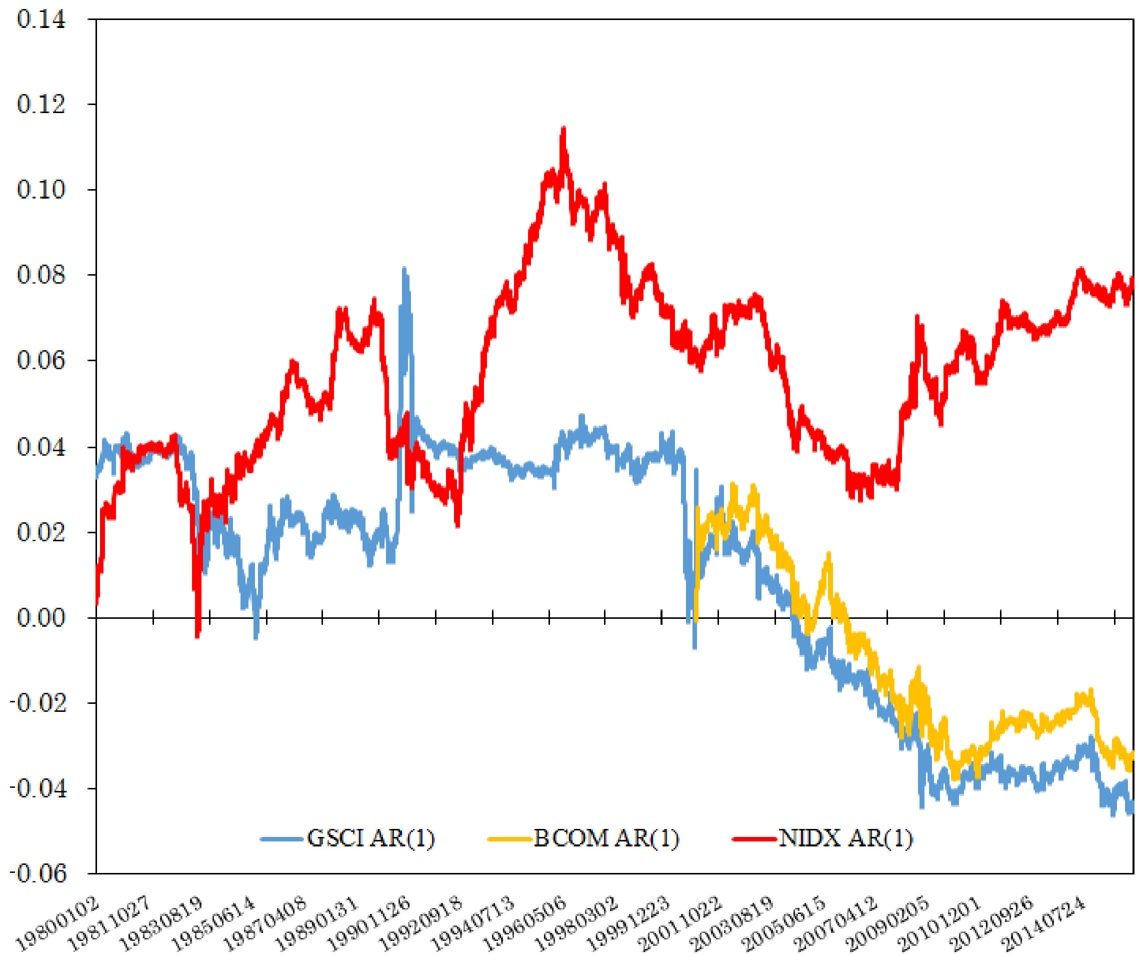


Figure 2: This figure plots the evolution of serial dependence in index returns from 1980 to 2018. Serial dependence is measured by first-order autocorrelation from index returns at the daily frequency. The indices are GSCI, BCOM and an equal-weighted portfolio non-indexed commodities (NIDX). Plotted series are the moving averages of these measures using a 10-year backward rolling window.

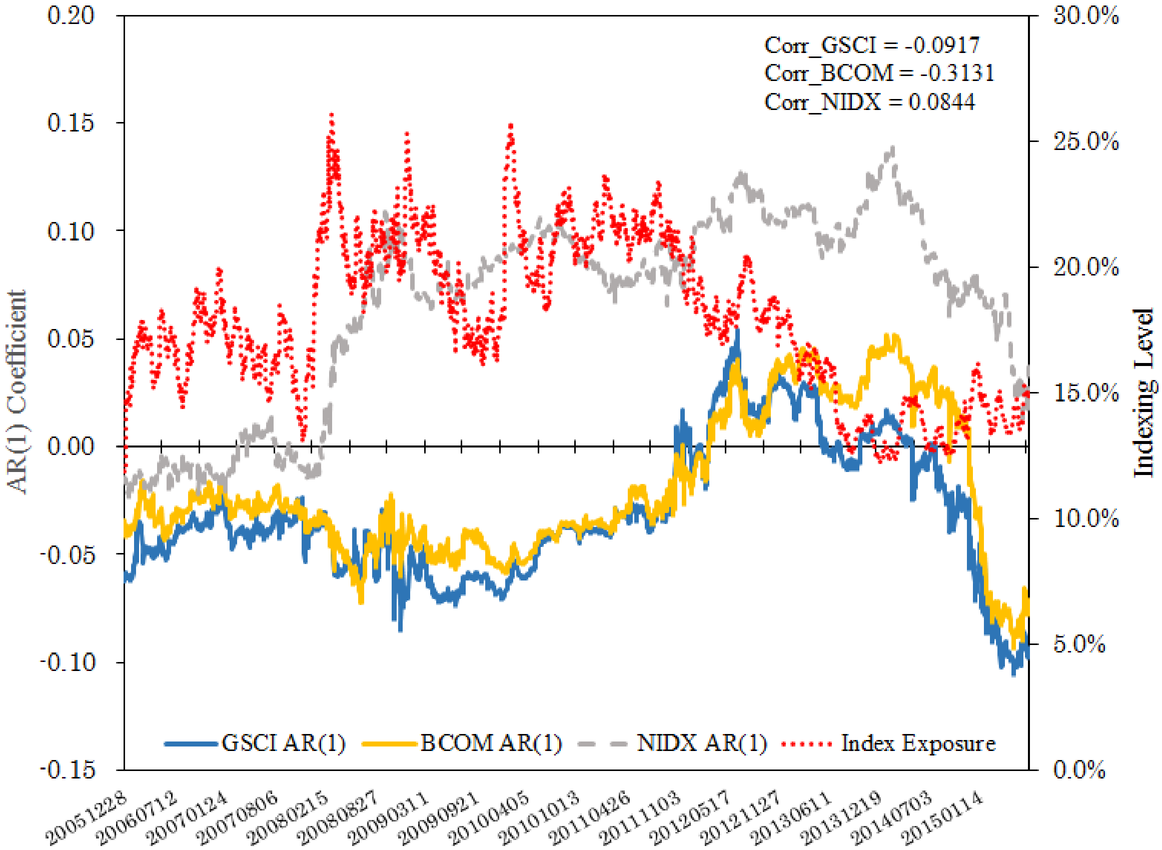


Figure 3: This figure plots the evolution of serial dependence in index returns and index exposure measure from 2006 to 2015. Serial dependence is measured by first-order autocorrelation from index returns at the daily frequency. The indices are GSCI, BCOM and an equal-weighted portfolio non-indexed commodities (NIDX). The index exposure is defined in equation (11). Plotted series are moving averages of these measures using a 3-year backward rolling window.

Figure 4: This figure plots the estimated market capital of index trading from 2006 to 2015 using Hamilton and Wu (2015) method.

Table 1: Detailed List of Commodities for Analysis

This table provide a detailed list of the commodities studied in this paper and their basic information. The futures contracts of these commodities are all traded in the United States. The GSCI and BCOM also include commodities traded in London, which are not included in our analysis. The commodities that are included in both indices are classified as "Indexed" commodities while these are not included in both indices are classified as "Non-indexed" commodities.

Ticker	Name	Full Name	Exchange	Inception	GSCI	BCOM	CIT	Indexed	Non-indexed
Panel A: Energy									
CL	Crude Oil	Crude Oil, WTI / Global Spot	NYMEX	1983/03/30	✓	✓		✓	
HO	Heating Oil	ULSD NY Harbor	NYMEX	1978/11/14	✓	✓		✓	
NG	Natural Gas	Natural Gas, Henry Hub	NYMEX	1990/04/04	✓	✓		✓	
RB	Gasoline	Gasoline, Blendstock	NYMEX	2005/10/03	✓	✓		✓	
Panel B: Grains									
BO	Soybean Oil	Soybean Oil / Crude	CBOT	1959/07/01		✓	✓		
C-	Corn	Corn / No. 2 Yellow	CBOT	1959/07/01	✓	✓	✓	✓	
KW*	KC Wheat	Wheat / No. 2 Hard Winter	CBOT	1970/01/05	✓		✓	*	
MW	Minn Wheat	Wheat / Spring 14% Protein	MGEX	1979/01/02					✓
O-	Oat	Oats / No. 2 White Heavy	CBOT	1959/07/01					✓
RR	Rough Rice	Rough Rice #2	CBOT	1986/08/20					✓
S-	Soybean	Soybeans / No. 1 Yellow	CBOT	1959/07/01	✓	✓	✓	✓	
SM*	Soybean Meal	Soybean Meal / 48% Protein	CBOT	1959/01/07		*	*		✓
W-	Wheat	Wheat / No. 2 Soft Red	CBOT	1959/07/01	✓	✓	✓	✓	
Panel C: Livestock									
FC	Feeder Cattle	Cattle, Feeder / Average	CME	1971/11/30	✓		✓		
LC	Live Cattle	Cattle, Live / Choice Average	CME	1964/11/30	✓	✓	✓	✓	
LH	Lean Hogs	Hogs, Lean / Average Iowa/S Minn	CME	1966/02/28	✓	✓	✓	✓	
Panel D: Metals									
GC	Gold	Gold	NYMEX	1974/12/31	✓	✓		✓	
HG	Copper	Copper High Grade / Scrap No. 2 Wire	NYMEX	1959/01/07	✓	✓		✓	
PA	Palladium	Palladium	NYMEX	1977/01/03					✓
PL	Platinum	Platinum	NYMEX	1968/03/04					✓
SI	Silver	Silver 5,000 Troy Oz.	NYMEX	1963/06/12	✓	✓		✓	
Panel E: Softs									
CC	Cocoa	Cocoa / Ivory Coast	ICE	1959/07/01	✓		✓		
CT	Cotton	Cotton / 1-1/16"	ICE	1959/07/01	✓	✓	✓	✓	
JO	Orange Juice	Orange Juice, Frozen Concentrate	ICE	1967/02/01					✓
KC	Coffee	Coffee 'C' /Colombian	ICE	1972/08/16	✓	✓	✓	✓	
LB	Lumber	Lumber / Spruce-Pine Fir 2x4	CME	1969/10/01					✓
SB	Sugar	Sugar #11/World Raw	ICE	1961/01/04	✓	✓	✓	✓	

*Note: KW and SM are both included in BCOM from 2013. Since SM is included in BCOM from 2013, its position on index trading is reported in CIT report since 2013.

Table 2: Descriptive Statistics of Commodities' Returns

This table provides some descriptive statistics of each commodity/Index' returns in columns 2-7. In column 8, we calculate the annualized Sharpe ratio (scaled by $\sqrt{252}$) of each commodity. In the last column, we show the correlation coefficient between the returns on each commodity/index and the return on S&P 500 composite index. NIDX denotes the equal weighted portfolio of non-indexed commodities. The sample is of daily frequency ranging from January 2, 2003 to December 29, 2015.

Commodity	Obs.	Mean	StDev.	Min	Max	AR(1)	Sharpe Ratio	ρ_{SP500}
Panel A: Energy								
CL	3266	-0.006%	0.0221	-0.1225	0.1427	-0.0650	-0.0455	0.2710
HO	3266	0.010%	0.0205	-0.0923	0.1041	-0.0401	0.0808	0.2267
NG	3266	-0.111%	0.0295	-0.1362	0.2064	-0.0583	-0.5974	0.0521
RB	3266	0.041%	0.0227	-0.1066	0.1385	-0.0395	0.2842	0.2291
Panel B: Grains								
BO	3272	0.009%	0.0157	-0.0698	0.0837	0.0094	0.0939	0.2122
C-	3272	0.008%	0.0186	-0.0989	0.0905	0.0316	0.0644	0.1491
KW	3272	0.009%	0.0188	-0.0860	0.0843	0.0082	0.0786	0.1419
MW	3272	0.037%	0.0180	-0.1071	0.2468	0.0379	0.3273	0.1187
O-	3272	0.045%	0.0209	-0.1125	0.1109	0.0744	0.3441	0.1279
RR	3272	0.017%	0.0154	-0.0595	0.0971	0.0777	0.1725	0.1272
S-	3272	0.055%	0.0162	-0.0782	0.0670	0.0081	0.5389	0.1588
SM	3272	0.084%	0.0182	-0.0805	0.0971	0.0374	0.7329	0.0971
W-	3272	-0.015%	0.0209	-0.0949	0.0919	-0.0049	-0.1145	0.1370
Panel C: Livestock								
FC	3264	0.024%	0.0090	-0.0583	0.0454	0.1277	0.4232	0.1298
LC	3264	0.014%	0.0096	-0.0616	0.0378	0.0797	0.2368	0.1347
LH	3272	-0.003%	0.0138	-0.0547	0.0543	0.0313	-0.0362	0.0407
Panel D: Metals								
GC	3266	0.036%	0.0122	-0.0934	0.0901	-0.0066	0.4667	0.0047
HG	3266	0.060%	0.0191	-0.1105	0.1235	-0.0826	0.4952	0.2942
PA	3266	0.040%	0.0209	-0.1237	0.1054	0.0721	0.3003	0.2210
PL	3266	0.022%	0.0144	-0.0916	0.0774	0.0509	0.2482	0.1769
SI	3266	0.049%	0.0217	-0.1771	0.1328	-0.0208	0.3593	0.1196
Panel E: Softs								
CC	3262	0.034%	0.0187	-0.0931	0.0974	0.0172	0.2900	0.1382
CT	3254	-0.006%	0.0188	-0.1236	0.0906	0.0592	-0.0477	0.1849
JO	3262	0.040%	0.0206	-0.1277	0.1669	0.0787	0.3064	0.0864
KC	3262	0.005%	0.0205	-0.1064	0.1385	-0.0237	0.0423	0.1433
LB	3272	-0.040%	0.0188	-0.0669	0.1064	0.0831	-0.3416	0.1233
SB	3262	0.004%	0.0201	-0.1163	0.0853	-0.0117	0.0291	0.1292
Panel F: Commodity Indices								
GSCI	3272	-0.010%	0.0152	-0.0829	0.0748	-0.0442	-0.1007	0.2869
BCOM	3266	-0.004%	0.0111	-0.0620	0.0581	-0.0308	-0.0584	0.2859
NIDX	3272	0.027%	0.0096	-0.0472	0.0441	0.0643	0.4531	0.2607

Table 3: Descriptive Statistics of Commodities' News Sentiment

This table provides some descriptive statistics of each commodity's news sentiment. The news sentiment of each commodity is calculated from the news tones data provided in Thomson Reuters News Analytics. The news are calculated in two steps. In the first step, we obtain the abnormal net tones by regressing the net tone (positive - negative) on its first lag and the month dummies. Then, in the second step, we orthogonalize the abnormal net tones (residuals in the first step regression) of each commodity on the its fundamentals (log basis and Amihud illiquidity). The whole sample is of daily frequency ranging from January 2, 2003 to December 29, 2015.

Commodity	Total # of News	Data Range	Obs.	StDev.	Min	Max
Panel A: Energy						
CL	994888	2003/01/02 - 2015/12/29	3257	0.0636	-0.2468	0.2184
HO	168937	2003/01/08 - 2015/12/29	3019	0.1343	-0.7825	0.8574
NG	492454	2003/01/02 - 2015/12/29	3257	0.0670	-0.2682	0.2752
RB	165861	2005/12/14 - 2015/12/29	2523	0.0959	-0.3365	0.3422
Panel B: Grains						
BO	559769	2003/01/02 - 2015/12/29	3257	0.0556	-0.6330	0.1941
C-	61972	2003/01/07 - 2015/12/29	2808	0.1666	-0.7501	0.6574
KW	54229	2008/12/05 - 2015/12/29	1308	0.1450	-0.7189	0.6274
MW	54229	2008/12/05 - 2015/12/29	1308	0.1460	-0.7012	0.6223
O-	806511	2003/01/02 - 2015/12/29	3257	0.0522	-0.2221	0.1697
RR	806511	2003/01/02 - 2015/12/29	3257	0.0544	-0.2426	0.1754
S-	60180	2003/02/06 - 2015/12/29	2040	0.1953	-0.8654	0.7206
SM	551181	2003/01/02 - 2015/12/29	3255	0.0617	-0.3280	0.2322
W-	54229	2008/12/05 - 2015/12/29	1308	0.1436	-0.7200	0.6036
Panel C: Livestocks						
FC	465815	2003/01/02 - 2015/12/29	3257	0.0691	-0.3273	0.2344
LC	465815	2003/01/02 - 2015/12/29	3257	0.0678	-0.3210	0.3229
LH	465815	2003/01/02 - 2015/12/29	3257	0.0695	-0.3290	0.2539
Panel D: Metals						
GC	282645	2003/01/02 - 2015/12/29	3257	0.1002	-0.4176	0.3684
HG	35100	2009/12/18 - 2015/12/29	1304	0.1747	-0.7621	0.7534
PA	282645	2003/01/02 - 2015/12/29	3257	0.1050	-0.4134	0.3545
PL	282645	2003/01/02 - 2015/12/29	3257	0.1009	-0.3438	0.3178
SI	282645	2003/01/02 - 2015/12/29	3257	0.1016	-0.4014	0.3273
Panel E: Softs						
CC	88814	2003/01/02 - 2015/12/29	3253	0.1346	-0.5673	0.4840
CT	105907	2003/01/02 - 2015/12/29	3255	0.1043	-0.6056	0.3229
JO	36098	2003/01/02 - 2015/12/29	2943	0.1638	-0.8048	0.7349
KC	107244	2003/01/02 - 2015/12/29	3255	0.1239	-0.5875	0.4784
LB	251274	2003/01/02 - 2015/12/29	3257	0.0841	-0.4554	0.4819
SB	149012	2003/01/02 - 2015/12/29	3255	0.1058	-0.4753	0.3903

*Note: As Thomson Reuters only provides some news tones up to sector level, we have to use sector news tones for some commodities. Specifically, (1) GC, SI, PA, and PL use scores for "Gold and Precious Metals"; (2) W-, MW and KW use scores for "Wheat"; (3) FC, LC, and LH use scores for "Livestocks"; (4). O- and RR use scores for "Grains".

Table 4: Spillover Effect of Sentiment on Returns across Indexed/Non-indexed Commodities

This table presents the results of regressing commodities returns on the "connected" sentiment meausure and controls. The "connected" sentiment measure is constructed in two steps. In the first step, we obtain each commodity's new sentiment by orthogonalizing the abnormal news tones on its fundamentals (i.e., basis and Amihud illiquidity). Then, we aggregate the sentiment of indexed commodities that belong to other sectors. For "connected" sentiment of non-indexed commodities, we take a simple average on the sentiment of non-indexed commodities from other sectors. The control variables include lagged returns, lagged basis, lagged Amihud Illiquidity, and lagged change in implied volatility of crude oil options with nearest maturity. The news tones and controls are of daily frequency ranging from January 2, 2003 to December 29, 2015. The CIT data is of weekly frequency ranging from January 3, 2006 to January 5, 2016. The *t*-statistics reported in the parenthesis are based on Newey-West standard errors with 4 lags. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	Panel A: Contemporaneous		Panel B: Predictive	
	Indexed	Non-indexed	Indexed	Non-indexed
Cnn. Sentiment	0.0605***	0.0507***		
	(21.47)	(13.21)		
L.Cnn. Sentiment			-0.0052*	-0.0015
			(-1.86)	(-0.41)
L.Return	-0.0121*	0.0722***	-0.0116	0.0721***
	(-1.69)	(7.61)	(-1.60)	(7.53)
L.Basis	0.0039	0.0055	0.0037	0.0048
	(0.61)	(0.40)	(0.58)	(0.35)
L.Illiquidity	1.58e-05***	1.08e-07	1.56e-05***	1.11e-07
	(2.66)	(1.23)	(2.59)	(1.30)
L.ΔOil ImVol	0.0001***	1.06e-05	0.0001***	2.32e-05
	(4.06)	(0.23)	(4.12)	(0.49)
Intercept	-0.0006*	0.0004	0.0010**	0.0004
	(-1.64)	(0.88)	(2.41)	(0.87)
Sector Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
# of Obs.	38,165	19,312	38,149	19,305
# of Individuals	16	8	16	8
Overall R-squared	1.50%	1.69%	0.29%	0.71%

Table 5: Spillover Effect of Sentiment on Returns across Indexed/Non-indexed Commodities excluding Financial Crisis Period

This table presents the subperiod results of regressing commodities returns on individual/sectoral "connected" sentiment meausures and controls. According to NBER's record of expansions and troughs, the sample excludes the period December 2007 to June 2009. The "connected" sentiment measure is constructed in two steps. In the first step, we obtain each commodity's new sentiment by orthogonalizing the abnormal news tones on its fundamentals (i.e., basis and Amihud illiquidity) and take the residuals as the sentiment. Then, we aggregate the sentiment of indexed commodities that belong to other sectors according to the weights based on estimated index trader's open interests (Hamilton and Wu, 2015). For "connected" sentiment of non-indexed commodities, we take a simple average on the sentiment of non-indexed commodities from other sectors. The control variables include lagged returns, lagged basis, lagged Amihud Illiquidity, and lagged change in implied volatility of crude oil options with nearest maturity. The news tones and controls are of daily frequency ranging from January 2, 2003 to December 29, 2015. The CIT data is of weekly frequency ranging from January 3, 2006 to January 5, 2016. The *t*-statistics reported in the parenthesis are based on Newey-West standard errors with 4 lags. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	Excluding Financial Crisis (2008/09/15 – 2009/06/30)			
	Panel A: Contemporaneous		Panel B: Predictive	
	Indexed	Non-indexed	Indexed	Non-indexed
Cnn. Sentiment	0.0600***	0.0375***		
	(21.92)	(9.98)		
L.Cnn. Sentiment			-0.0072***	-0.0028
			(-2.62)	(-0.76)
L.Return	-0.0047	0.0813***	-0.0032	0.0819***
	(-0.64)	(8.09)	(-0.43)	(8.11)
L.Basis	0.0049	-0.0055	0.0038	-0.0073
	(0.73)	(-0.39)	(0.46)	(-0.51)
L.Illiquidity	1.00e-05**	9.49e-08	9.07e-06*	9.77e-08
	(1.96)	(1.05)	(1.76)	(1.09)
LD.Oil ImVol	6.06e-05*	-2.95e-05	5.18e-05	-2.76e-05
	(1.69)	(-0.56)	(1.45)	(-0.53)
Intercept	-0.0003	-5.31e-06	5.55e-05	0.0015**
	(-0.53)	(-0.01)	(0.10)	(2.05)
Sector Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
# of Obs.	35,465	17,872	35,449	17,865
# of Individuals	16	8	16	8
Overall R-squared	1.47%	1.38%	0.17%	0.82%

Table 6: Spillover Effect of Sentiment on Returns across Indexed/Non-indexed Commodities under High/Low Abnormal Index Exposure

This table presents the results of regressing commodities returns on individual/sectoral "connected" sentiment measures and controls under different levels of abnormal index exposure. The abnormal index exposure is measured as the ratio of index investment flow to the total market cap. Based on this measure, we calculate the median of the weekly average abnormal index exposure. Then, we classify the daily sample within the week whose average exposure is above/below the median into "High/Low" exposure groups. The "connected" sentiment measure is constructed in two steps. In the first step, we obtain each commodity's new sentiment by orthogonalizing the abnormal news tones on its fundamentals (i.e., basis and Amihud illiquidity) and take the residuals as sentiment. Then, we aggregate the sentiment of indexed commodities that belong to other sectors according to estimated index trader's open interests (Hamilton and Wu, 2015). For "connected" sentiment of non-indexed commodities, we take a simple average on the sentiment of non-indexed commodities from other sectors. The control variables include lagged returns, lagged basis, lagged Amihud Illiquidity, and lagged change in implied volatility of crude oil options with nearest maturity. The news tones and controls are of daily frequency ranging from January 2, 2003 to December 29, 2015. The CIT data is of weekly frequency ranging from January 3, 2006 to January 5, 2016. The t-statistics reported in the parenthesis are based on Newey-West standard errors with 4 lags. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

	Panel A: Contemporaneous				Panel B: Predictive			
Variables	Indexed		Non-indexed		Indexed		Non-indexed	
	High	Low	High	Low	High	Low	High	Low
Cnn. Sentiment	0.0493***	0.0697***	0.0406***	0.0582***				
	(12.08)	(16.90)	(7.60)	(10.20)				
L. Cnn. Sentiment					-0.0146***	0.0022	-0.0046	0.0008
					(-3.55)	(0.54)	(-0.86)	(0.15)
L.Return	-0.0343***	-0.0220**	0.0431***	0.0838***	-0.0462***	-0.0091	0.0434***	0.0805***
	(-3.28)	(-2.06)	(2.83)	(5.90)	(-4.19)	(-0.84)	(2.96)	(5.69)
L.Basis	0.0017	0.0093	0.0162	0.0104	0.0009	0.0084	0.0170	0.0085
	(0.17)	(1.17)	(0.87)	(0.53)	(0.09)	(1.02)	(0.88)	(0.44)
L.Illiquidity	9.91e-06	2.09e-05***	-1.02e-07	6.23e-07**	7.55e-06	2.20e-05***	-8.41e-08	5.29e-07***
	(1.57)	(3.56)	(-1.52)	(2.27)	(1.15)	(4.44)	(-1.37)	(2.67)
L.ΔOil ImVol	0.0002***	0.0001**	-6.22e-05	0.0001*	6.50e-05	0.0003***	-8.41e-05	0.0002***
	(3.04)	(2.28)	(-1.01)	(1.68)	(1.36)	(4.76)	(-1.23)	(2.62)
Intercept	0.0067***	-0.0045***	0.0021	-0.0007	0.0066**	-0.0033***	0.0034	-0.0006
	(9.07)	(-7.37)	(0.82)	(-1.11)	(2.02)	(-5.53)	(1.46)	(-0.95)
Sector Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
# of Obs.	17,433	17,387	8,787	8,741	17,460	17,344	8,806	8,715
# of Individuals	16	16	8	8	16	16	8	8
Overall R-squared	2.11%	2.64%	1.69%	2.77%	1.14%	0.87%	0.79%	1.33%

Table 7: Return Serial Dependence and Commodity Indexing: Interaction Effect

This table presents the results of regressing commodities returns on the interaction between commodities' lagged returns and abnormal index exposure. The abnormal index exposure is measured as the ratio of change in market cap on index trading to the total market cap. In the regression, we control for each commodity's lagged change in returns, lagged change in basis, lagged change in Amihud illiquidity, and lagged change in implied volatility of crude oil options with nearest maturity. The financial crisis period is 2008/09/15 – 2009/06/30. The CIT data is of weekly frequency ranging from January 3, 2006 to January 5, 2016. The *t*-statistics reported in the parenthesis are based on Newey-West standard errors with 4 lags. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	Full Sample		Exclude Financial Crisis	
	Indexed	Non-indexed	Indexed	Non-indexed
L.(Abn. Index Exposure $\times$ Return)	-4.3122***	-0.8201	-5.7340***	0.2486
	(-2.83)	(-0.35)	(-3.42)	(0.10)
L.Abn. Index Exposure	0.1186***	0.0747*	0.1162***	0.0547
	(3.87)	(1.74)	(3.75)	(1.21)
L.Return	-0.0223***	0.0703***	-0.0129	0.0825***
	(-2.83)	(6.80)	(-1.62)	(7.55)
L.Basis	0.0004	0.0001	0.0013	-0.0129
	(0.05)	(0.01)	(0.18)	(-0.79)
L.Illiquidity	1.78e-05***	9.43e-08	1.12e-05*	7.57e-08
	(3.01)	(1.26)	(1.87)	(0.99)
LD.Oil ImVol	0.0002***	4.78e-05	6.92e-05*	-6.89e-06
	(4.78)	(0.97)	(1.86)	(-0.12)
Intercept	0.0081***	0.0079***	0.0080**	0.0078***
	(15.25)	(2.84)	(2.17)	(2.78)
Individual Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
# of Obs.	34,789	17,513	32,089	16,073
# of Individuals	16	8	16	8
Overall R-squared	0.45%	0.85%	0.32%	0.98%

Table 8: Return Serial Dependence and Commodity Indexing: Predict Correlation

This table presents the results of regressing commodities serial dependence measure on commodities' abnormal index exposure. The serial dependence measure is defined as $(r_{it}r_{it-1})/2\sigma_i^2$. The abnormal index exposure is measured as the ratio of change in market cap on index trading to the total market cap. In the regression, we control for each commodity's lagged change in returns, lagged change in basis, lagged change in Amihud illiquidity, and lagged change in implied volatility of crude oil options with maturity less than one month. The CIT data is of weekly frequency ranging from January 3, 2006 to January 5, 2016. The financial crisis period is 2008/09/15 – 2009/06/30. The t-statistics reported in the parenthesis are based on Newey-West standard errors with 4 lags. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	Full Sample		Exclude Financial Crisis	
	Indexed	Non-indexed	Indexed	Non-indexed
L.Abn. Index Exposure	-6.2068***	-1.3423	-4.8122**	1.3344
	(-2.80)	(-0.44)	(-2.44)	(0.52)
L.Serial Dependence	-0.0175	-0.1116**	-0.0072	-0.1200*
	(-0.74)	(-2.13)	(-0.30)	(-1.94)
L.Basis	-0.1601	2.3330**	-0.2128	2.3660**
	(-0.33)	(2.14)	(-0.43)	(2.02)
L.Illiquidity	-0.0016**	-5.96e-06*	-0.0003	-6.67e-06*
	(-2.13)	(-1.71)	(-0.51)	(-1.74)
LD.Oil ImVol	-0.0072**	0.0007	0.0031	0.0057**
	(-2.42)	(0.25)	(1.50)	(2.11)
Intercept	0.2251*	0.0063	0.2798***	0.0067
	(1.87)	(0.09)	(8.13)	(0.15)
Individual Fixed Effect	Yes	Yes	Yes	Yes
Year Fixed Effect	Yes	Yes	Yes	Yes
# of Obs.	34,789	17,513	32,089	16,073
# of Individuals	16	8	16	8
Overall R-squared	0.34%	1.51%	0.22%	1.87%

Table 9: Serial Dependence and Indexing in Overweighted GSCI/BCOM Commodities

This table presents the results based on relative index overweighting in GSCI/BCOM commodities after regressing a zero-investment trading strategies against GSCI (BCOM) indexing relative to BCOM (GSCI) indexing ("Relative indexing"). Relative indexing is the change in GSCI (BCOM) related ETF/ETN exposure minus BCOM (GSCI) related ETF/ETN exposure, where an index's ETF/ETN exposure is defined as the change in market cap of ETF/ETNs following GSCI/BCOM divided by the estimated GSCI/BCOM market cap. The market cap of each ETF is calculated by multiplying its shares outstanding by its NAV. The GSCI/BCOM market cap is obtained by summing up each indexed commodity's dollar value positions with CIT reported data using Hamilton and Wu (2015). The GSCI/BCOM OW portfolio returns are defined as $R_t^p := \sum_{j=1}^N \varpi_{jt}^p r_{jt}$, where $p \in \{GSCI, BCOM\}$ and ϖ_{jt}^p is a weight assigned to return r_{jt} on commodity j based on the overweighting measure

$$OW^p_{jy(t)} = \begin{cases} w^p_{jy(t)} - w^q_{jy(t)}, & \text{if } w^p_{jy(t)} > 0, \\ 0, & \text{if } w^p_{jy(t)} = 0. \end{cases}$$

with $p = GSCI$, $q = BCOM$ or vice versa, and $w^p_{jy(t)}$ being the commodity j's weight in GSCI/BCOM for year $y(t)$ respectively. We then calculate strategy weight $\varpi^p_{jt} = (OW_{jy(t)} - N^{-1}\sum_{j=1}^N OW_{jy(t)})r^p_{t-1}$. In the regression for GSCI (BCOM) OW portfolio, we control for the lagged relative (GSCI-BCOM) return, lagged relative realized volatility over the past 250 trading days, the lagged relative log ETF/ETNs' trading volume detrended with one-year average log trading volume, and lagged implied volatility of crude oil options with maturity less than 1 month. The CIT data is of weekly frequency ranging from January 3, 2006 to January 5, 2016. The ETF trading data is of daily frequency starting from June 7, 2006 to December 29, 2015. The t-statistics reported in the parenthesis are based on Newey-West standard errors with 4 lags. ***, ** and * denote statistical significance at the 1%, 5%, and 10% levels, respectively.

Variables	GSCI OW Portfolio		BCOM OW Portfolio	
	Full Sample	Exclude Financial Crisis	Full Sample	Exclude Financial Crisis
L.Relative ETF Indexing	-8.82e-05*	-0.0001**	-9.80e-05**	-0.0001**
	(-1.79)	(-2.01)	(-2.44)	(-2.50)
L.Relative Return	0.0012	0.0003	-0.0007	-0.0004
	(1.26)	(0.56)	(-1.25)	(-1.08)
L.Relative RVOL	-0.0001	0.0002**	9.21e-06	-0.0002**
	(-1.00)	(2.27)	(0.09)	(-2.35)
L.Relative Volume	-1.15e-06	-2.42e-06	-6.35e-08	1.15e-06
	(-0.38)	(-1.10)	(0.00)	(0.82)
LD.Oil ImVol	-2.50e-06*	-3.42e-07	1.73e-06*	7.73e-08
	(-1.68)	(-0.36)	(1.79)	(0.13)
Intercept	-8.83e-07	-4.46e-06**	1.98e-06	4.10e-06***
	(-0.33)	(-2.36)	(1.02)	(2.74)
# of Obs.	1,930	1,750	1,930	1,750
Adj. R-squared	1.20%	0.17%	1.01%	0.38%

Returns and Risk Premia on Energy Options

Kris Jacobs*

University of Houston

Bingxin Li†

West Virginia University

(Preliminary draft, please do not cite or circulate)

September 18, 2019

Abstract

We study returns on NYMEX futures options on crude oil, natural gas, gasoline, and heating oil. Average returns in the natural gas option market strongly differ from the other three markets. The moneyness and maturity patterns in call and put returns, straddle returns and delta-hedged returns display similarities across all four markets. These patterns are also similar to stylized facts in equity index option markets. Variance risk premiums are mostly negative, again consistent with index option markets, but positive for the one-month maturity.

JEL Classification: G13; G17

Keywords: Futures Options; Energy; Commodities; Option Returns; Variance Risk Premiums.

*Kris Jacobs, C. T. Bauer College of Business, University of Houston, Phone: 713-743-2826, Email: kjacobs@bauer.uh.edu.

†Bingxin Li, John Chambers College of Business and Economics, Center for Innovation in Gas Research and Utilization, West Virginia University, Phone: 304-293-2777, Email: bingxin.li@mail.wvu.edu.

1 Introduction

Commodity markets have grown rapidly over the past decade and have become an increasingly important component of financial markets. Energy commodities account for more than 70% of the S&P GSCI (the Goldman Sachs Commodity Index) and are thus one of the most critical components of commodity markets. Commodity derivatives markets have also experienced very rapid growth. The Bank for International Settlements (BIS) estimates that the gross market value of the global over the counter (OTC) commodity derivatives market increased from $42 billion in June 1998 to $220 billion in December 2018.

The study of options on commodity futures, and energy futures options, in particular, is therefore of great interest. In this paper, we study return patterns in four energy option markets: NYMEX WTI crude oil futures, Henry Hub natural gas futures, NYMEX heating oil futures, and Reformulated Gasoline Blendstock for Oxygen Blending (RBOB) gasoline futures. Liquidity and data availability differ across these markets. Crude oil option markets are very liquid, followed by natural gas option markets. The market for gasoline and heating oil options is considerably less liquid. Crude oil options, heating oil options, and natural gas options are available over a twenty-five year period, similar to most studies in the index option and equity option literature. The sample for RBOB gasoline options is considerably shorter.

The existing literature on energy option returns is very limited. Christoffersen and Pan (2019) investigate investors' belief and state price densities in the crude oil market and document the option returns on crude oil futures. They find that crude oil OTM option returns support U-shaped state price densities. Dew-Becker, Giglio, and Kelly (2019) document that returns on commodity straddles are generally negative at one-month maturities, increase as maturity increases, and become statistically significantly positive at the five-month maturity. Kang and Pan (2019) find negative variance risk premia in the crude oil market. Procopczuk, Symeonidis, and Simen (2017) document 60-day variance risk premiums for a cross-section of commodities.

We first consider the sign of average call and put returns, and return patterns as a function of moneyness and maturity. In index option markets, Coval and Shumway (2001) establish that with a conventional stochastic discount factor that is decreasing in aggregate wealth, returns on index calls are positive, returns on index puts are negative, and both call and put returns increase as a function of moneyness. Bakshi and Kapadia (2003b) study returns in equity option markets based on an index model, and distinguish between the priced market risk and idiosyncratic risk in option returns. We find that both put and call returns are on average negative for crude oil, heating oil, and gasoline, but positive for natural gas. The different returns patterns in the natural gas options market are not likely to be driven by differences in the returns on the underlying futures markets.

Put returns conform to theory more than calls when inspecting average returns as a function of the strike price. For crude oil, heating oil, and gasoline puts, option returns increase as a function of the strike price. In the call option market, the theoretically anticipated pattern is resoundingly rejected in all four markets, and the most stringent rejections occur for OTM calls. Some of these patterns in call option returns are consistent with existing findings in the index option literature (Bakshi, Madan, and Panayotov, 2010; Ni, 2009; Bondarendo, 2014). In that literature, several studies have proposed U-shaped pricing kernels to explain these return patterns (Bakshi, Madan, and Panyotov, 2010; Christoffersen, Heston, and Jacobs, 2013).

For crude oil, heating oil, and gasoline, returns on longer-maturity options exceed returns on shorter-maturity options. In fact, average returns are negative for short maturities and positive for longer maturities. This finding is also consistent with existing findings in the index option literature. Bakshi and Kapadia (2003a) suggest that this can be explained by a negative variance risk premium if volatilities are lower at longer horizons. Volatilities are indeed lower at long horizons in the energy markets we study.

We then compute various other returns on options and option strategies. We start with returns on straddles and zero-beta straddles, which reduce the exposure to underlying price risk. We find that returns on straddles generally increase as a function of maturity. Returns are often negative at short maturities, especially for ATM straddles, and positive at long maturities, except for gasoline options. These results are consistent with Dew-Becker, Giglio, and Kelly (2019), who investigate a broad cross-section of commodity markets and report that straddles generally earn negative returns at short maturities, while average returns are positive at longer maturities. Returns on zero-beta straddles at short maturities are on average negative. This finding is consistent with Coval and Shumway's (2001) finding for index option markets. However, we find zero-beta straddles become positive when maturities are longer than three months. These results suggest that, just as in index option markets, energy option returns are affected by important risk factors other than market returns.

We find that energy delta-hedged gains are negative for most short maturities, but for longer maturities some delta-hedged gains are positive. Bakshi and Kapadia (2003a) study the relation between delta-hedged returns and the variance risk premium. They find negative delta-hedged index returns and show that this is consistent with a negative market volatility risk premium. We find that the average variance risk premium at the one-month maturity is positive and significant in the crude oil, natural gas, and heating oil markets. For maturities longer than one month, variance risk premia are consistently negative. These findings are therefore difficult to reconcile with the delta-hedged returns.

Our results are related to various other strands of literature. On the one hand, there is an important literature on commodity futures returns, starting with Gorton and Rouwenhorst (2006).[1] There is also smaller literature on pricing models for commodity options.[2] The studies by Coval and Shumway (2001), Bakshi and Kapadia (2003a, 2003b), and Bakshi, Madan, and Panayotov (2010) are part of a larger literature that studies patterns in and determinants of index and equity option returns and delta-hedged returns.[3] Finally, there is of course a large literature on option pricing in equity, currency, and fixed income markets, including pricing models for these markets. This literature is too voluminous to cite here.

[1]See, among others, Erb and Harvey (2006), Gorton, Hayashi, and Rouwenhorst (2012), and Bhardwaj, Gorton, and Rouwenhorst (2015).

[2]See, for example, Trolle and Schwartz (2009) and Christoffersen, Jacobs, and Li (2016).

[3]See Jones (2006), Cao and Huang (2007), Broadie, Chernov, and Johannes (2009), Constantinides, Jackwerth, and Perrakis (2009), Constantinides, Jackwerth, and Savov (2013), Bondarenko (2014), Chaudhuri and Schroder (2015), Muravyev (2016), and Hu and Jacobs (2019).

The rest of the paper proceeds as follows. Section 2 discusses the energy futures and futures option data. Section 3 presents returns on energy call and put options. We analyze return patterns as a function of moneyness and maturity, and discuss potential explanations for these patterns. Section 4 presents results for options straddles and delta-hedged returns. Section 5 analyzes variance risk premiums and relates them to returns patterns. Section 6 concludes.

2 Energy Futures and Options Data

We use futures and futures option data from the Chicago Mercantile Exchange (the CME Group, formerly NYMEX). These data are relatively new in the literature and we first discuss the data source and the filters we impose. We then present descriptive statistics, discuss the stylized facts in the data, and compare them to stylized facts in equity option markets that have been extensively studied in the literature.

2.1 Data Source

The Chicago Mercantile Exchange (the CME Group, formerly NYMEX) energy derivatives market is the world's largest and most liquid commodity derivatives market. We use all available derivatives contracts on four energy products traded on the CME: NYMEX WTI crude oil futures options; Henry Hub natural gas futures options; NYMEX heating oil futures options; and the RBOB gasoline futures options. For crude oil options and heating oil options, data are available from January 2nd, 1990 to August 12th, 2016. Data on natural gas options are available beginning October 5th, 1992, and data on RBOB gasoline options are available starting May 15th, 2006. While these are relatively short time series, the data include several recessions and geopolitical events, and for three of the four underlying contracts, the length of the resulting time series is comparable to data typically used in the index and equity option literature.

The CME commodity futures and futures options data include a wide range of energy, agricultural and other commodities. Here we exclusively focus on energy futures and futures options. These options are more liquid than some of the other commodity options, and perhaps as a result of the data quality also seems higher. The energy contracts allow us to study options with a wide range of maturities and a broad range of strike prices. The number of traded energy futures and futures options contracts has significantly increased through the sample period, especially for longer maturities. For example, in our sample, the number of crude oil futures with positive open interest increased from 17 to 45, and the number of crude oil options with positive open interest increased from 77 to 1435. Liquidity has also significantly increased throughout the sample period. Open interest (daily trading volume) for the nearest crude oil futures contract increased from 66,925 (45,177) to 273,746 (86,622). The combined open interest (daily trading volume) for all options on that futures contracts increased from 92,083 (16,427) to 376,694 (32,820) contracts.

Because of liquidity pattern and the space concerns in reporting the results, we focus on the nearest twelve monthly futures contracts. We refer to these contracts as M1-M12. We exclusively retain options on these twelve futures contracts, but we also impose additional filters. We discard options that violate the standard no-arbitrage conditions and options with implied volatilities greater than two. If we have fewer than five remaining usable strikes in a given cross-section, we discard the entire cross-section. For each option maturity, we assign the remaining available options to various bins according to moneyness. We use seven moneyness intervals, 0.86-0.90, 0.90-0.94, 0.94-0.98, 0.98-1.02, 1.02-1.06, 1.06-1.10, and 1.10-1.14, where moneyness is defined as option strike divided by the price of the underlying futures contract.

The crude oil, natural gas, heating oil, and gasoline options we use are American options on the corresponding futures contracts. The CME has also introduced European-style options; however, the trading history on these contracts is much more limited than that of the American options. The literature has proposed different methods to convert American option prices to European prices. The existing literature on commodity markets (Trolle and Schwartz, 2009; Christoffersen, Jacobs, and Li, 2016) applies the Barone-Adesi and Whaley (1987) (BAW henceforth) adjustment and uses ATM and OTM options to minimize the effect of well-known errors in the BAW early exercise approximation. Because we are also interested in returns of the in-the-money (ITM) options, we use the method of Bjerksund and Stensland (2002) to convert the American option prices to European option prices.[4] We invert the option prices using the Black (1976) model to get the implied volatilities.

2.2 Descriptive Statistics

Our focus is on the energy futures options data, which we discuss in detail below. We first provide a brief discussion of the underlying futures data to provide additional perspective on the option data. The descriptive statistics for the futures data are presented in Table A1 in the Appendix. All prices in this paper are settlement prices.[5] Not surprisingly, the WTI crude oil futures contracts are the most liquid contracts, judged by open interest and trading volume. Liquidity generally decreases with time to maturity in all four markets, meaning investors trade more aggressively as the futures contracts approach expiration when the uncertainty of the market increases and hedging is more urgent.

The average futures price is relatively similar across different maturities for crude oil, heating oil, and gasoline. For natural gas futures contracts, the average futures price increases slightly as the time to maturity increases. In Figure 1, the left-side panels plot the prices of the one-month futures contracts. Futures prices increase dramatically between 2003 and 2008 and sharply decline after July 2008. Prices have since recovered in three of the four markets under consideration, but not in the natural gas futures market. The right-side panels plot the difference between the one-month and one-year futures prices. While Table A1 indicates that the futures term structure is relatively flat in our sample period for crude oil, heating oil, and gasoline, Figure 1 indicates that the term structure slope strongly varies over time. This is also the case for the natural gas market, which is on average in contango over our

[4] Alternatively, the approach of Ju and Zhong (1999) can be used.

[5] The CME light sweet crude oil futures contract trades in units of 1000 barrels and prices are quoted in U.S. dollars per barrel. The natural gas futures contract trades in units of one mmBtu and prices are quoted in U.S. dollars per mmBtu. The heating oil futures contract trades in units of 42,000 gallons and prices are quoted in U.S. dollars per gallon. The gasoline futures contract trades in units of 42,000 gallons and prices are quoted in U.S. dollars per gallon.

sample period. Figure 1 indicates that the natural gas futures market is in backwardation for extended periods of time during our sample period.

Figure 1 also clearly illustrates how closely related the markets for crude oil, heating oil, and gasoline are. The correlation between the one-month prices for crude oil and heating oil is 0.99, while the correlation between the one-month crude oil and natural gas contracts is 0.53. For the one-year contracts, these correlations are 0.99 and 0.61, respectively. The correlation between the crude oil and heating oil slopes in the right-side panels is 0.84, and for the crude oil and natural gas slopes, the correlation is 0.51.

Table A1 reports the arithmetic and geometric average of daily simple returns. For example, the geometric average for the one-month futures contract is 24.6 basis point per month, which amounts to just under 3% per year. Average returns on gasoline futures are negative for all maturities, which is due to the sample period. The pattern of futures returns as a function of maturity is not monotonic in any of the four markets.

Table 1 reports summary statistics for option contracts with different maturities in the crude oil, natural gas, heating oil, and gasoline markets. As with the underlying futures markets, the liquidity of options also generally decreases as the time to maturity increases. Implied volatility decreases as a function of time to maturity in all markets, similar to the patterns in index option and equity option markets. Table 2 reports summary statistics for these markets as a function of moneyness. OTM options have the largest number of option contracts trading (on average) in all four markets. OTM options also have the highest trading volume and open interest in all four markets.

Tables 1 and 2 clearly indicate that the crude oil option market has the highest liquidity, followed by the natural gas markets. Heating oil and gasoline option markets are much less liquid.

Figure 2 plots the time series of ATM option-implied volatility in the four energy markets. The left-side panels plot the one-month implied volatility. The right-side panels plot the difference between the one-month and one-year implied volatilities. As indicated by Table 2, implied volatilities with shorter maturities are significantly higher on average in all four markets. They are also more volatile, as suggested by the volatility of implied volatility in Table 1. Figure 2 also confirms that implied volatility in the natural gas market is significantly higher compared to the other three markets. Table 2 indicates that the level of implied volatility is similar in the crude oil, heating oil, and gasoline markets. These implied volatilities are, in fact, highly correlated. For the one-month contract on the left-side of Figure 2, the correlations between crude oil implied volatility and implied volatility in the heating oil and gasoline markets are 0.91 and 0.92, respectively. The correlation between the implied volatility in the natural gas market and that in the other three markets is substantially lower. For instance, the correlation between implied volatility in crude oil and natural gas is 0.42. This is not surprising, because the gasoline and heating oil markets are very narrowly related to the crude oil market. While the market for natural gas is of course also heavily impacted by the crude oil market, it is a relatively large market that is heavily influenced by some supply and demand factors that are distinct from the main factors that determine fluctuations in the crude oil market.

Figure 3 graphs the energy markets' implied volatility as a function of moneyness for contracts with maturities of one week, one month, three months, and six months. The graphs in the left column use call options. The graphs in the middle column use put options. In the right-side column, we only use OTM contracts in order to study the moneyness patterns with the most liquid contracts. The graphs in the right-side column suggest that any inference from the two other columns is not likely to be very distorted by liquidity concerns.

Consistent with the findings in Figure 2, Figure 3 shows that short-term implied volatility is higher for all moneyness intervals and all markets. The patterns in implied volatility as a function of moneyness in Figure 3 are possibly the most important stylized facts in the option literature. In Figure 3, ATM options have the lowest implied volatility in all four energy markets. This is a pattern that is also present in several other markets. Figure 3 also shows that the levels of ITM and OTM implied volatility are not very dissimilar, for both calls and puts. When the implied volatility plot is symmetric around the minimum (ATM) implied volatility, it is referred to as the option smile. Several important markets exhibit such a smile, such as the U.S. index option market pre-1987 and currency markets pre-2007. For other markets, the pattern is asymmetric, often with OTM puts having the highest implied volatility. This pattern is usually referred to as the smirk. The post-1987 U.S. index option market is an important example of an option market that is characterized by a smirk. In fact, the smirk in the U.S. index option market is very pronounced, while other major option markets, such as equity options, display a much less pronounced smirk. The existing literature often explains the cross-sectional differences in these patterns through differential exposure to systematic risk.

While the four energy markets in Figure 3 do not exhibit a very pronounced smirk when compared to index option markets, the pattern is also certainly not symmetric. It is important to note that once again, the natural gas market is the outlier and the patterns for the crude oil market, the heating oil market, and the gasoline market are very similar. Specifically, the volatility smiles for short-term options are flatter in the natural gas market. Confirming the findings of Table 2, Figure 3 also indicates that the average implied volatility in the natural gas market is higher than that in the other three markets. Moreover, for ATM natural gas options, the difference between implied volatilities for different maturities is much wider than that in the other three markets. Recall from Table 1 that the natural gas futures option market is the second-most liquid market of the four. It is therefore unlikely that these patterns are due to liquidity and/or measurement problems.

3 Option Returns

In this section, we first present stylized facts on call and put energy option returns. We then discuss the implications of these stylized facts for economic theory, and their relation to the existing literature.

3.1 Returns on Call and Put Energy Options

We compute hold-to-maturity returns on futures options, as follows:

$$R_{t,\,T}^{call} = \max(F_T - K, 0)/C_{t,\,T} - 1, \tag{3.1}$$

$$R_{t,\,T}^{put} = \max(K - F_T, 0)/P_{t,\,T} - 1, \tag{3.2}$$

where F_T is the futures price at the maturity T.[6] Let $F_{t,\,T}$ denote the underlying futures price at time t with time to maturity $\tau = T - t$, then $F_T \equiv F_{T,\,T} \equiv S_T$ because the futures price converges to the underlying spot price at expiration. Furthermore, K is the option strike price, and $C_{t,\,T}$ and $P_{t,\,T}$ are prices of European style call and put options with time to maturity $T - t$ at time t.[7]

We compute hold-to-maturity option returns for a set of fixed times to maturity. We follow the implementation of Bliss and Panigirtzoglou (2004), and determine a target observation date for horizons of one week, one month, two months and so on up to one year by subtracting the appropriate number of days (weekly horizon) or months (monthly and one-year horizon) from the expiration date. If there are no options traded on the target observation date, the nearest options trading date is determined. If this nearest trading date differs from the target observation date by no more than three days for weekly horizons or four days for monthly and 1-year horizons, that date is substituted for the original target date. If no sufficiently close trading date exists, that expiration is excluded from the sample for that horizon.

Panel A of Table 3 reports average monthly returns on call and put options as a function of maturity. We sum over all available contracts for the maturities mentioned above. Returns that are significantly different from zero are indicated in bold. We now discuss the implications

[6] Crude oil futures expire on the third business day before the 25th calendar day (or the business day right before it if the 25th is not a business day) of the month that precedes the delivery month. Natural gas futures expire the third last business days of the month prior to the delivery month. Heating oil futures and Gasoline futures expire on the last business day of the month preceding the delivery month.

[7] For crude oil, heating oil, and gasoline, options expire three business days before the termination of trading in their corresponding futures. The natural gas options expire one business day before the termination of trading in the natural gas futures. For notational simplicity, we use the same maturity, T, for futures and options on the corresponding futures.

of these return patterns for economic theory. Recall that Coval and Shumway (2001) show that the derivative of the expected option returns with respect to the strike price is proportional to the negative value of the covariance between a stochastic discount factor and the underlying price. Therefore, if the stochastic discount factor is monotonically decreasing, then option returns increase as the strike price increases. Moreover, we expect call returns to be positive on average, while we expect put returns to be negative on average. Panel A of Table 3 indicates that in our sample, both put and call returns are on average negative for crude oil, heating oil, and gasoline, but on average positive for natural gas.

Put returns also conform more to the theory when inspecting average returns as a function of the strike price. Panel A of Table 3 indicates that crude oil, heating oil, and gasoline puts closely follow Coval and Shumway's (2001) prediction that option returns should increase as a function of the strike price. In the natural gas put market, the pattern as a function of the strike price is not monotonic, but the more liquid OTM puts confirm the theoretically expected increasing return pattern. However, in the call markets, the theoretically anticipated pattern is resoundingly rejected in all four markets. The most stringent rejections occur for the more liquid out-of-the-money calls.

Panel B of Table 3 reports return patterns as a function of maturity. For crude oil, heating oil, and gasoline, the average monthly returns on longer-maturity options exceed average returns on shorter-maturity options. In fact, the weighted average returns across the seven moneyness intervals are negative for short maturities and positive for longer maturities. For example, the average return for calls with one-month maturity in the crude oil market is -15.4% per month, while that for calls with six-month maturity is 5.1% per month. The average return for puts with one-month maturity in the crude oil market is -3.4% per month, while that for puts with six-month maturity is 4.1% per month. The overall return average for calls and puts among all maturities and seven moneyness buckets is -11.0% and -8.6% per month, respectively. Option returns for heating oil and gasoline have a similar pattern. The market for natural gas is once again the outlier.

Figure 4 provides additional detail by graphing returns as a function of moneyness for two maturities, one month and six months. The red triangles represent the results for the entire sample, corresponding to the results in Table 3. Figure 4 shows that the average returns on calls are highly negative for short maturities in the crude oil, heating oil, and gasoline markets. These negative returns are mostly obtained for OTM calls.

It is likely that there are structural breaks in energy markets during our sample period. Specifically, there is an intense debate about whether financialization and/or capital inflows from index funds have impacted pricing and volatility in commodity markets over the past decade. Figure 4 therefore also reports results for three sub-samples: before 2005, from January 2005 to December 2008, and from January 2009 to August 2016. Since gasoline option data starts from 2004, we only differentiate two sample periods in this market, before and after December 2008. The first sub-sample period, before 2005, is typically thought of as a steadily increasing commodity market with low volatility. The second sample, from January 2005 to December 2008, ends with the financial crisis and saw an increase in the financialization of commodity markets and higher volatility. In this period, energy futures prices increase rapidly, but they are very volatile. The last sample period, from January 2009 to August 2016, covers the period after the financial crisis when energy futures prices decrease.

Figure 4 shows that call option returns for the first sample period, before 2005, are the highest among the three subsample periods. Call option returns for the third sub-sample, after 2009, are the lowest. For put option returns, the first sub-sample period is characterized by relatively lower average returns. Put option returns for the second sub-sample period are higher. This pattern is consistent across all four energy markets. Figure 4 also indicates that the differences in average returns across different subsamples are bigger for OTM options than for ITM options for all markets and all maturities. This is consistent with theory, due to the higher exposure to the underlying and higher sensitivity to market conditions (Cao and Huang, 2007).

3.2 Discussion

We conclude that the patterns in Table 3 seemingly conflict with priors from economic theory, both with respect to the sign of average returns and the return patterns as a function of moneyness and maturity. The stylized facts in call and put natural gas markets differ from those in the other three markets. However, given the high degree of integration between the crude oil, heating oil, and gasoline markets, and the resulting high correlation between the underlying futures, as evident from Figure 1, the similarities in the return patterns in these option markets are partly by design and technology-driven. We now discuss some potential explanations for these patterns.

Average option returns result from a variety of factors, and of course the expected return on the underlying security is the most important one. Realized average futures returns differ from expected returns, especially in small samples. It is therefore possible that the differences in option returns are related to the fact that in our sample period, prices in the natural gas market evolved differently from those in the crude oil market, as is evident from Figure 1. However, it is unlikely that these differences in the returns on the underlying explain the different patterns in option returns. First, while the patterns in gasoline puts and calls in Table 3 and Figure 4 are similar to those in the crude oil and heating oil markets, note that these averages are due to a much shorter sample period (2005-2016) due to data availability. Table A1 shows that the average return on the gasoline futures contract is negative during this period, while the returns on crude oil and heating oil futures are positive during the longer 1990-2016 sample. Second, Figure 4 shows that in the 1990-2004 sub-sample, one-month natural gas call option returns are positive on average and increase as a function of the strike price, which differs from the patterns in the other three markets in that period. Recall that Figure 1 shows that during this period, the returns on the underlying futures markets are very similar in the three markets for which data are available. We conclude that the differences in patterns in average option returns in Table 3 and Figure 4 are not likely to be driven by differences in the returns on the underlying futures markets.

The finding in Table 3 that call option returns decrease with strike prices and put option returns increase with strike prices has also been documented in the index option literature, see for instance Broadie, Chernov, and Johannes (2009), Bakshi, Madan, and Panayotov (2010), Bondarenko (2014), and Christoffersen, Heston and Jacobs (2013). Hu and Jacobs (2019) document similar patterns in the cross-section of equity option returns and Christoffersen and Pan (2019) for the crude oil market.

Chaudhuri and Schroder (2015) show that a strictly decreasing pricing kernel is equivalent to expected returns increasing in the strike for the log-concave class of option trading strategies. Christoffersen, Heston, and Jacobs (2013) argue that U-shaped pricing kernels can explain these patterns and that such pricing kernels can be supported by preferences defined on variance. Bakshi, Madan, and Panyotov (2010) show that only claims with payouts on the upside are capable of discerning between U-shaped and monotonically declining kernels. They show that expected returns on index call options with strikes higher than a threshold are negative and decreasing in the strike, which supports a U-shaped pricing kernel. Bakshi, Madan, and Panyotov (2010) show that U-shaped pricing kernels can be supported by a model with heterogeneity in beliefs.

How to interpret the finding that energy option return patterns as a function of moneyness are consistent with patterns in the index option market? In index option markets, the underlying asset is the overall stock market, and therefore the prices of risk implied by these markets can be thought of as defining a pricing kernel defined on aggregate wealth. For other option markets, such as those defined on underlying stocks or commodities, the implied prices of risk do not have the same implications. However, the finding that return patterns are similar to those in equity markets is very interesting. For instance, it could be interpreted as suggesting that institutional issues or demands factors common to the index option, and commodity markets explain these factors, rather than the representative agent's pricing of market risk. Alternatively, the pricing of commodity options may reflect the pricing of exposure to market risk.[8] Or the finding of a U-shaped pricing kernel common to both index and commodity

[8] Bhardwaj, Gorton, and Rouwenhorst (2015) find that the correlation between commodity futures returns

markets may indicate that downside and upside risks are priced similarly in both markets, without any reference to the pricing of market risk.

The finding in Table 3 that option returns increase as a function of maturity in the crude oil, gasoline, and heating oil markets is also consistent with existing findings in the index option literature. Bakshi and Kapadia (2003a) argue that a negative variance risk premium suggests that an increase in the variance increases the risk-neutral drift of the variance process, raises option prices, and decreases option returns. Because volatilities are lower at longer horizons, option returns increase. Table 1 indicates that in all four energy markets we study, implied volatility is lower for longer-maturity options, which may therefore explain some of the patterns as a function of maturity in Table 3. However, our findings for the natural gas market do not support this channel.

In summary, average return patterns on calls and puts in the crude oil, heating oil and gas markets are similar and distinct from those in the natural gas market. While differences in underlying returns certainly drive some option return differences over time, the differential option return patterns in the natural gas market cannot be explained by the fact that the natural gas market evolved differently after 2010. In general, while option returns of course fluctuate as a function of the returns in the underlying, the patterns as a function of moneyness and maturity are quite stable over time. These patterns display many similarities with existing findings from index option markets. They are also consistent with some theories that have been proposed to explain patterns in index option markets, such as negative variance risk premia and U-shaped pricing kernels.

4 Returns on Straddles and Delta-Hedged Gains

4.1 Straddle Returns

with equities after 2005 has increased significantly and is about 0.5 during the 2005-2014 period.

Returns on straddles and zero-beta straddles are interesting because they reduce the exposure to underlying price risk. A straddle consists of buying a call option and a put option with the same strike price, K. From (3.1) and (3.2), we compute the return on straddles as follows:

$$R^{straddle}_{t,\,T} = \frac{\max(F_T - K, 0) + \max(K - F_T, 0)}{C_{t,\,T} + P_{t,\,T}} - 1. \tag{4.1}$$

A zero-beta straddle can be computed by weighting the call options and put options appropriately so that the beta of the straddle is zero. Consider

$$R^{zero-beta}_{t,\,T} = \frac{\theta \max(F_T - K, 0) + (1-\theta)\max(K - F_T, 0)}{\theta C_{t,\,T} + (1-\theta) P_{t,\,T}} - 1. \tag{4.2}$$

The weight parameter θ is obtained by solving the following problem, which gives a zero-beta portfolio:

$$\theta \beta^c_{t,\,T} + (1-\theta)\beta^p_{t,\,T} = 0. \tag{4.3}$$

where $\beta^c_{t,\,T}$ and $\beta^p_{t,\,T}$ are betas for calls and puts, which are estimated as follows. For each option maturity and moneyness interval, we run the following regressions for the full sample

$$R^{call}_{t,\,T} - R^f_{t,T} = \alpha^c_T + \beta^c_T(R^{Fut}_{t,\,T} - R^f_{t,T}) + \varepsilon^c_{t,\,T}, \tag{4.4}$$

$$R^{put}_{t,\,T} - R^f_{t,T} = \alpha^p_T + \beta^p_T(R^{Fut}_{t,\,T} - R^f_{t,T}) + \varepsilon^p_{t,\,T}, \tag{4.5}$$

where $R^{Fut}_{t,\,T}$ is the hold-to-maturity futures returns and $R^f_{t,T}$ is the hold-to-maturity risk-free rate. Solving equation (4.3) gives

$$\theta = \frac{-\beta^p_{t,\,T}}{\beta^c_{t,\,T} - \beta^p_{t,\,T}}. \tag{4.6}$$

Table 4 reports monthly returns on straddles and zero-beta straddles. Generally, average returns on straddles increase as a function of maturity. Returns are often negative at short maturities, especially for ATM straddles, and positive at long maturities, except for gasoline options. These results are consistent with Dew-Becker, Giglio, and Kelly (2019), who investigate a broad cross-section of commodity markets and report that straddles generally earn negative returns at short maturities, while average returns are positive for at longer maturities.

The average returns on zero-beta straddles as a function of maturity is similar to the straddle returns. At the one-week and one-month maturities, most of the zero-beta straddle returns are negative. For example, the average return on the zero-beta ATM straddle with one-week maturity in the crude oil market is -3.5% per month. Coval and Shumway (2001) find that the zero-beta, ATM straddles with the second nearest maturity produce an average loss of approximately -13% per month. The zero-beta straddle returns increase with maturity and become positive for maturities longer than three months. For example, the average return on the zero-beta ATM straddle with six-month maturity in the crude oil market is 13.9% per month. The zero-beta straddles returns at the long maturities are on average higher than the straddle returns. The average returns on zero-beta straddles across all maturities are positive except for the gasoline market. Overall, these results suggest that there are important risk factors other than the risk of the underlying futures market in energy option markets, just as in index option markets.

4.2 Delta-Hedged Gains

We compute delta-hedged gains on energy options. Broadie, Chernov, and Johannes (2009) discuss three approaches to implement option delta-hedging strategies. We use the Black (1976) model delta evaluated at the option's implied volatility. The deltas for calls and puts in the Black (1976) model are:

$$\delta^c_{t,\,T} = \exp(-r_f \cdot \tau) N(d_1), \tag{4.7}$$

$$\delta^p_{t,\,T} = \exp(-r_f \cdot \tau)(N(d_1) - 1). \tag{4.8}$$

where $N(\cdot)$ is the cumulative density function of a normal distribution, and $d_1 = \frac{\ln(F_{t,\,T}/K) + 0.5\sigma^2\tau}{\sigma\sqrt{\tau}}$.

Following Bakshi and Kapadia (2003a), we define the daily-rebalanced delta-hedged option gain for a call option and a put option portfolio over a period $[t, T]$ as

$$\begin{aligned}
\Pi_{t,\,T}^{c} &= \max(F_T - K, 0) - C_{t,\,T} - \sum_{n=1}^{\tau} \delta_{t+n-1,\,T}^{c}(F_{t+n,\,T} - F_{t+n-1,\,T}) \qquad (4.9)\\
&\quad - \sum_{n=1}^{\tau} r_f (C_{t,\,T} - \delta_{t+n-1,\,T}^{c} F_{t+n-1,\,T}),\\
\Pi_{t,\,T}^{p} &= \max(K - F_T, 0) - P_{t,\,T} - \sum_{n=1}^{\tau} \delta_{t+n-1,\,T}^{p}(F_{t+n,\,T} - F_{t+n-1,\,T}) \qquad (4.10)\\
&\quad - \sum_{n=1}^{\tau} r_f (P_{t,\,T} - \delta_{t+n-1,\,T}^{p} F_{t+n-1,\,T}).
\end{aligned}$$

Bakshi and Kapadia (2003a) show that if volatility risk is priced in a stochastic volatility model, the expected delta-hedged gain is determined by the price of volatility risk and the vega of the option. As vega is always positive, Bakshi and Kapadia (2003a) model show that a negative (positive) variance risk premium implies that the expected delta-hedged gains are negative (positive). They report that most of the delta-hedged gains for S&P 500 index calls with the two nearest maturities are negative, and in their sample period, the ATM delta-hedged gains for S&P 500 calls with the nearest maturity is approximately -0.26\$.

We find that energy delta-hedged gains are negative for short maturities, but that delta-hedged gains are positive for longer maturities. Table 5 shows results for one-month and six-month maturities. At the one-month maturity, the ATM delta-hedged gains for crude oil calls is -0.123\$. This is smaller than Bakshi and Kapadia's (2003a) findings for the index option market, but our sample period is of course different. Bakshi and Kapadia (2003b) report that for a sample of 25 equity options, the average delta-hedged loss divided by the value of the underlying security is about half of the corresponding loss for the index.

5 The Energy Variance Risk Premium

The variance risk premium is defined and computed in various ways in the literature, depending on how the physical variance is treated. We compute the variance risk premium as the difference between the conditional expectation of the variance under the physical and the risk-neutral (RN) measures:

$$VRP_{t,\,T} = E_t^P[Var_{t \to T}] - E_t^Q[Var_{t \to T}], \tag{5.1}$$

where $VRP_{t,\,T}$ denotes the variance risk premium at time t for futures and options contracts mature at time T.

To compute the risk-neutral (RN) variance, we follow Bakshi, Kapadia, and Madan (2003):

$$E_t^Q[Var_{t \to T}] = \exp(r_f \tau) \cdot [\int_{F_{t,\,T}}^{\infty} \frac{2[1 - \ln(\frac{K}{F_{t,\,T}})]}{K^2} C_{t,\,T}(K)dK + \int_{0}^{F_{t,\,T}} \frac{2[1 + \ln(\frac{F_{t,\,T}}{K})]}{K^2} P_{t,\,T}(K)dK]. \tag{5.2}$$

The measure of physical variance is subject to choices of data frequency and length. Given the data available, we calculate the physical variance using daily futures returns. Each month on the futures expiration date, we calculate the annualized physical variance as the realized variance using the previous τ-day daily futures returns,

$$Var_{t \to T} = \frac{252}{\tau} \sum_{n=1}^{\tau} [\ln(\frac{F_{T-n+1,\,T}}{F_{T-n,\,T}}) - u_{t,\,T}]^2, \tag{5.3}$$

where $u_{t,\,T}$ is the realized average (from t to T) of daily futures returns with maturity T. Depending on the choice of $\tau = T - t$, we have realized variance at different times to maturity. For example, for one-month realized variance, we set $\tau = 21$; and for six-month realized variance, we set $\tau = 126$.

Given that the physical variance is realized while RN variance is forward-looking, we need to estimate an ex-ante physical variance. The forecasting models in Drechsler and Yaron (2011) and Bekaert and Hoerova (2014) regress the time series of realized variance on it lags and the RN variance. We adopt Drechsler and Yaron (2011)'s method and regress the log realized variance on the log values of lagged realized variance and the RN variance. Specifically, for each time to maturity τ, we consider the following monthly regression,

$$\log(Var_{t\to T}) = \alpha + \beta \log(Var_{t-\tau\to t}) + \gamma \log(E_t^Q[Var_{t\to T}]) + \epsilon, \tag{5.4}$$

Using the properties of log-normality and the parameters estimated in (5.4), we calculate the expected realized variance as

$$E_t^P[Var_{t\to T}] = \exp[\alpha + \beta \log(Var_{t-\tau\to t}^P) + \gamma \log(E_t^Q[Var_{t\to T}]) + \frac{1}{2}\sigma_\epsilon]. \tag{5.5}$$

For each of the four energy markets, we compute a monthly variance risk premium with the time to maturity between one month and twelve months. Table 6 reports the time-series averages for all maturities. Variances are expressed in annualized form, so a variance of 0.04 corresponds to 20% yearly volatility. Table 6 shows that the average variance risk premium at the one-month maturity is positive and significant in the crude oil, natural gas, and heating oil markets. For maturities longer than one month, variance risk premia are consistently negative, and they are large for maturities of three months or more. For the six month-maturity, variance risk premia in the natural gas and heating oil markets are almost always negative with occasional positive outliers. In the crude oil market, we find the variance risk premium at the one-month maturity is 0.006, which accounts for 4.41% of the risk-neutral variance. The crude oil variance risk premium at the six-month maturity is -0.015, which accounts for -15.64% of the risk-neutral variance. Figure 5 plots the time series for the one-month and six-month variance risk premia. Variance risk premia are highly time-varying in all markets.

Table 6 indicates that variance risk premia in the natural gas market are much larger than in the other three markets. The natural gas variance risk premium at the six-month maturity is -0.051, representing -31.66% of the risk-neutral variance. Figure 5 indicates that the six-month variance risk premium also contains very large negative outliers.

In index and equity markets, the variance risk premium is mostly negative. Kang and

Pan (2019) use tick data in the crude oil market with maturities of 21, 63, and 126 business days, and also find find negative variance risk premia. Prokopczuk et al. (2017), using a daily sample with a similar starting date but ending in 2011, report that 60-day variance risk premia in energy markets are on average negative. Our results confirm the negative variance risk premium at the long maturities but indicate that the sign changes at shorter maturities. The next step is to use intraday futures data to verify the robustness of the variance risk premium estimates.

6 Conclusion

We investigate option returns on four energy commodities. We find that both put and call returns are on average negative for crude oil, heating oil, and gasoline, but positive for natural gas. For crude oil, heating oil, and gasoline puts, option returns increase as a function of the strike price. In the call option market, the theoretically anticipated pattern is resoundingly rejected in all four markets. We also examine returns on straddles, zero-beta straddles, and delta-neutral portfolios. ATM straddle and zero-beta straddle returns are negative for one-month maturities, but at maturities longer than three months, they earn positive returns. Delta-hedged gains are negative at short maturities, but some delta-hedged gains are positive at longer maturities. However, variance risk premiums are positive at the one-month maturity and negative at longer maturities.

We conclude that several of these patterns in energy option returns are consistent with existing findings in index option markets. A potential explanation for these patterns is the U-shaped pricing kernel.

Several findings require further investigation. First, return patterns in the natural gas options market are different from those in the other three energy markets. We need to verify if this can be explained by the distinct features of this market, such as the higher volatility in the natural gas market or the natural gas futures term structure in our sample, and how these return patterns are related to the structure of the implied volatility smile. Second, we need to further investigate the robustness of our findings on the variance risk premium and the relation with delta-hedged gains and zero-beta straddles. Finally, given that option returns are highly non-normal, we need to assess the statistical significance of our results using more robust methods.

References

[1] Bakshi, G. and Kapadia, N., 2003a. Delta-Hedged Gains and the Negative Market Volatility Risk Premium. The Review of Financial Studies 16 (2), 527–566.

[2] Bakshi, G. and Kapadia, N., 2003b. Volatility Risk Premium Embedded in Individual Equity Options: Some New Insights. Journal of Derivatives 11 (1), 45-54.

[3] Bakshi, G., Kapadia, N., and Madan, D., 2003. Stock Return Characteristics, Skew Laws, and the Differential Pricing of Individual Equity Options. The Review of Financial Studies 16 (1), 101-143.

[4] Bakshi, G., Madan, D., and Panayotov, G., 2010. Returns of Claims on the Upside and the Viability of U-Shaped Pricing Kernels. Journal of Financial Economics 97, 130-154.

[5] Barone-Adesi, G. and Whaley, R., 1987. Efficient Analytic Approximation of American Option Values. Journal of Finance 42, 301-320.

[6] Bekaert, G. and Hoerova, M., 2014. The VIX, the Variance Premium and Stock Market Volatility. Journal of Econometrics 183, 181-192.

[7] Bhardwaj, G., Gorton, G. B., Rouwenhorst, K.G., 2015. Facts and Fantasies about Commodity Futures Ten Years Later. Working Paper, National Bureau of Economic Research.

[8] Black, F., 1976. The Pricing of Commodity Contracts. Journal of Financial Economics 3, 167-179.

[9] Bondarenko, O., 2014. Why Are Put Options So Expensive? Quarterly Journal of Finance 4, 145-195.

[10] Broadie, M., Chernov, M., and Johannes, M., 2009. Understanding Index Option Returns. Review of Financial Studies 22 (11), 4493-4529.

[11] Cao, C. and Huang, J. Z., 2007. Determinants of S&P 500 Index Option Returns. Review of Derivatives Research 10, 1-38.

[12] Chaudhuri, R. and Schroder, M., 2015. Monotonicity of the Stochastic Discount Factor and Expected Option Returns. The Review of Financial Studies 28 (5), 1462-1505.

[13] Christoffersen, P., Heston, S., and Jacobs, K., 2013. Capturing Option Anomalies with a Variance-Dependent Pricing Kernel. The Review of Financial Studies 26 (8), 1963-2006.

[14] Christoffersen, P., Jacobs, K., and Li, B., 2016. Dynamic Jump Intensity and Risk Premiums in the Crude Oil Futures and Options Markets. Journal of Derivatives 24 (2), 1-30.

[15] Christoffersen, P. and Pan, X., 2019. The State Price Density Implied by Crude Oil Futures and Option Prices. Working Paper, University of Oklahoma.

[16] Constantinides, G. M., Jackwerth, J. C., and Perrakis, S., 2009. Mispricing of S&P 500 Index Options. The Review of Financial Studies 22 (3), 1247-1277.

[17] Constantinides, G. M., Jackwerth, J. C., and Savov, A., 2013. The Puzzle of Index Option Returns. The Review of Asset Pricing Studies 3 (2), 229-257.

[18] Coval, J. D. and Shumway, T., 2001. Expected Option Returns. Journal of Finance 56 (3), 983-1009.

[19] Dew-Becker, I., Giglio, S., and Kelly, B., 2019. Hedging Macroeconomic and Financial Uncertainty and Volatility. Working Paper, Northwestern University.

[20] Drechsler, I. and Yaron, A., 2011. What's Vol Got to Do with It. Review of Financial Studies 24, 1-45.

[21] Erb, C. B. and Harvey, C. R., 2006. The Strategic and Tactical Value of Commodity Futures. Financial Analysts Journal 62 (2), 69-97.

[22] Gorton, G. B., Hayashi, F., and Rouwenhorst, K. G., 2012. The Fundamentals of Commodity Futures Returns. Working Paper, Yale University.

[23] Gorton, G. B. and Rouwenhorst, K. G., 2006. Facts and Fantasies about Commodity Futures. Financial Analysts Journal 62 (2), 47-68.

[24] Hu, G. and Jacobs, K., 2019. Volatility and Expected Option Returns. Forthcoming, Journal of Financial and Quantitative Analysis.

[25] Jones, C. S., 2006. A Nonlinear Factor Analysis of S&P 500 Index Option Returns. Journal of Finance 61 (5), 2325-2363.

[26] Ju, N. and Zhong, R., 1999. An Approximate Formula for Pricing American Options. The Journal of Derivatives 7 (2),31-40.

[27] Kang, S. B. and Pan, X. N., 2019. Commodity Variance Risk Premia and Expected Futures Returns: Evidence from the Crude Oil Market. Working Paper, University of Oklahoma.

[28] Keynes, J. M., 1930. A Treatise on Money. Macmillan, London.

[29] Muravyev, D., 2016. Order Flow and Expected Option Returns. Journal of Finance 71 (2), 673-708.

[30] Newey, W. K. and West, K. D., 1987. A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix. Econometrica 55 (3), 703-708.

[31] Ni, S. X., 2009. Stock Option Returns: A Puzzle. Working Paper, Hong Kong Baptist University.

[32] Prokopczuk, M., Symeonidis, L., and Simen, C. W., 2017. Variance Risk in Commodity Markets. Journal of Banking and Finance 81, 136-149.

[33] Trolle, A. B. and Schwartz, E. S., 2009. Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives. Review of Financial Studies 22, 4423-4461.

Table 1. Energy Options: Summary Statistics by Maturity

Panel A. Crude Oil

Maturity	<1m	1m-2m	2m-3m	3m-4m	4m-5m	5m-6m	6m-7m	7m-8m	8m-9m	9m-10m	10m-11m	11m-1y
Price	2.337	3.102	3.814	4.441	4.948	5.418	5.822	6.115	6.435	6.708	7.096	7.544
Implied Volatility	0.425	0.351	0.339	0.332	0.325	0.319	0.311	0.305	0.300	0.293	0.290	0.287
Volatility of IV	0.179	0.122	0.112	0.105	0.098	0.091	0.083	0.080	0.076	0.074	0.075	0.073
No. of Contracts	22	46	44	40	35	30	25	21	18	15	12	10
Open Interest	3739	3280	2237	1740	1478	1328	1226	1183	1143	1101	1085	1152
Trading Volume	645	468	204	120	91	74	69	62	55	51	47	40

Panel B. Natural Gas

Price	0.255	0.382	0.475	0.535	0.582	0.619	0.653	0.673	0.701	0.725	0.748	0.797
Implied Volatility	0.540	0.503	0.481	0.457	0.437	0.419	0.404	0.388	0.377	0.366	0.355	0.350
Volatility of IV	0.188	0.154	0.139	0.121	0.111	0.106	0.101	0.093	0.089	0.084	0.079	0.076
No. of Contracts	31	39	37	34	31	28	26	23	20	18	15	12
Open Interest	1054	762	557	461	404	354	318	280	251	240	240	246
Trading Volume	197	108	54	37	30	23	21	18	16	14	12	10

Panel C. Heating Oil

Price	0.062	0.085	0.105	0.122	0.137	0.149	0.161	0.169	0.174	0.178	0.187	0.211
Implied Volatility	0.359	0.328	0.324	0.319	0.315	0.309	0.304	0.300	0.295	0.291	0.289	0.293
Volatility of IV	0.136	0.103	0.098	0.092	0.089	0.085	0.081	0.080	0.078	0.076	0.077	0.078
No. of Contracts	40	47	41	36	33	29	25	21	17	13	9	5
Open Interest	285	270	212	163	127	101	85	75	67	57	51	57
Trading Volume	27	27	19	12	9	6	4	4	4	3	4	3

Panel D. Gasoline

Price	0.073	0.104	0.131	0.147	0.158	0.169	0.186	0.208	0.229	0.240	0.256	0.274
Implied Volatility	0.376	0.347	0.341	0.335	0.332	0.329	0.323	0.325	0.328	0.326	0.334	0.331
Volatility of IV	0.147	0.117	0.111	0.109	0.113	0.112	0.115	0.116	0.112	0.094	0.082	0.056
No. of Contracts	60	60	44	34	25	17	11	7	5	3	2	1
Open Interest	160	154	147	136	116	116	124	137	149	160	173	188
Trading Volume	13	12	8	8	7	6	7	6	5	4	7	5

Notes to Table: We report daily averages of prices, annualized implied volatility (IV), the volatility of IV, the number of contracts traded, open interest, and trading volume for energy options by maturity. The sample period starts on January 2, 1990 for crude oil and heating oil options, October 5, 1992 for natural gas options, and May 15, 2006 for gasoline options. The sample period ends on August 12, 2016.

Table 2. Energy Options: Summary Statistics by Moneyness

Panel A. Crude Oil	Calls							Puts						
Moneyness (K/F)	.86-.90	.90-.94	.94-.98	.98-1.02	1.02-1.06	1.06-1.10	1.10-1.14	.86-.90	.90-.94	.94-.98	.98-1.02	1.02-1.06	1.06-1.10	1.10-1.14
Price	10.277	8.181	6.309	4.789	3.432	2.513	1.871	1.862	2.558	3.514	4.740	6.111	7.900	9.968
Implied Volatility (IV)	0.370	0.343	0.325	0.314	0.313	0.316	0.323	0.350	0.334	0.322	0.316	0.324	0.343	0.367
Volatility of IV	0.121	0.104	0.099	0.101	0.105	0.110	0.118	0.117	0.105	0.102	0.104	0.110	0.120	0.137
No. of Contracts	17	21	25	28	26	25	23	26	27	28	27	20	15	11
Open Interest	1367	1545	1746	1983	2093	2052	1993	2539	2505	2221	1791	1451	1274	1178
Trading Volume	20	29	70	307	319	250	197	263	330	372	307	74	30	18
Panel B. Natural Gas														
Price	1.019	0.837	0.689	0.570	0.472	0.398	0.338	0.264	0.339	0.437	0.549	0.671	0.814	0.968
Implied Volatility (IV)	0.447	0.437	0.432	0.427	0.432	0.440	0.449	0.432	0.429	0.429	0.435	0.451	0.473	0.494
Volatility of IV	0.144	0.134	0.133	0.134	0.136	0.139	0.142	0.135	0.133	0.132	0.133	0.136	0.139	0.146
No. of Contracts	16	20	24	28	28	26	25	25	26	26	25	20	15	12
Open Interest	426	438	443	457	477	477	476	609	582	531	481	446	426	402
Trading Volume	10	16	35	88	82	73	64	85	101	107	99	34	14	9
Panel C. Heating Oil														
Price	0.290	0.231	0.178	0.133	0.098	0.074	0.058	0.046	0.062	0.083	0.113	0.157	0.207	0.258
Implied Volatility (IV)	0.358	0.331	0.310	0.297	0.304	0.318	0.332	0.331	0.315	0.304	0.304	0.321	0.349	0.380
Volatility of IV	0.102	0.088	0.089	0.094	0.096	0.097	0.102	0.096	0.092	0.095	0.100	0.105	0.112	0.125
No. of Contracts	16	22	29	35	32	28	23	24	26	27	23	15	10	6
Open Interest	118	130	145	167	190	195	188	151	166	176	172	161	148	134
Trading Volume	3	4	7	20	21	18	15	11	16	21	24	7	4	4
Panel D. Gasoline														
Price	0.338	0.258	0.188	0.132	0.093	0.070	0.055	0.038	0.056	0.084	0.126	0.179	0.241	0.301
Implied Volatility (IV)	0.394	0.363	0.337	0.324	0.326	0.337	0.356	0.363	0.343	0.333	0.333	0.356	0.391	0.424
Volatility of IV	0.134	0.113	0.112	0.112	0.115	0.118	0.123	0.124	0.116	0.119	0.123	0.132	0.141	0.157
No. of Contracts	11	16	23	29	29	27	23	19	22	23	21	13	8	5
Open Interest	149	154	154	161	155	136	120	141	153	150	141	134	120	96
Trading Volume	2	3	6	16	13	10	7	6	10	12	15	3	1	1

Notes to Table: We report daily averages of prices, annualized implied volatility (IV), the volatility of IV, the number of contracts traded, open interest, and trading volume for energy options by moneyness. The sample period starts on January 2, 1990 for crude oil and heating oil options, October 5, 1992 for natural gas options, and May 15, 2006 for gasoline options. The sample period ends on August 12, 2016.

Table 3. Option Returns

Panel A. Option Returns by Moneyness

Moneyness (K/F)		.86-.90	.90-.94	.94-.98	.98-1.02	1.02-1.06	1.06-1.10	1.10-1.14	Mean
Crude Oil	Calls	**0.026**	0.014	0.022	0.003	-0.017	-0.229	**-0.589**	**-0.110**
		(1.686)	(0.720)	(0.876)	(0.078)	(-0.085)	(-1.143)	**(-9.071)**	**(-2.476)**
	Puts	**-0.362**	**-0.164**	0.018	-0.006	-0.006	-0.003	0.020	**-0.086**
		(-4.213)	**(-1.721)**	(0.159)	(-0.117)	(-0.162)	(-0.114)	(0.789)	**(-2.607)**
Natural Gas	Calls	0.005	0.027	0.006	0.007	0.159	0.323	0.073	**0.094**
		(0.142)	(0.745)	(0.140)	(0.112)	(1.010)	(1.058)	(0.234)	**(1.690)**
	Puts	0.070	0.200	**0.499**	0.058	**0.071**	**0.086**	**0.107**	**0.170**
		(0.470)	(1.125)	**(1.716)**	(1.439)	**(2.093)**	**(2.907)**	**(3.829)**	**(2.468)**
Heating Oil	Calls	**0.052**	**0.060**	0.024	-0.039	**-0.247**	**-0.387**	**-0.395**	**-0.144**
		(3.045)	**(3.136)**	(0.992)	(-1.057)	**(-4.562)**	**(-4.169)**	**(-2.916)**	**(-5.636)**
	Puts	**-0.427**	**-0.421**	**-0.222**	**-0.108**	**-0.066**	**-0.058**	-0.024	**-0.236**
		(-3.140)	**(-4.823)**	**(-2.305)**	**(-2.172)**	**(-1.723)**	**(-1.664)**	(-0.709)	**(-6.382)**
Gasoline	Calls	0.027	**0.074**	0.064	0.093	0.042	-0.371	**-0.921**	**-0.151**
		(0.690)	**(1.642)**	(1.085)	(0.877)	(0.188)	(-1.262)	**(-11.534)**	**(-2.185)**
	Puts	**-1.078**	**-1.048**	**-0.775**	**-0.404**	**-0.193**	-0.098	-0.058	**-0.661**
		(-7.937)	**(-8.697)**	**(-6.500)**	**(-4.251)**	**(-2.226)**	(-1.269)	(-0.807)	**(-13.959)**

Panel B. Option Returns by Maturity

	Crude Oil		Natural Gas		Heating Oil		Gasoline	
Maturity	Calls	Puts	Calls	Puts	Calls	Puts	Calls	Puts
1w	**-0.793**	**-0.866**	0.644	0.832	**-1.236**	**-1.536**	**-0.607**	**-2.262**
	(-1.822)	**(-3.155)**	(0.797)	(1.215)	**(-5.780)**	**(-5.685)**	**(-1.824)**	**(-13.095)**
1m	**-0.154**	-0.034	-0.016	**0.207**	-0.050	-0.082	**-0.112**	**-0.141**
	(-3.770)	(-0.606)	(-0.233)	**(3.062)**	(-0.799)	(-1.217)	**(-2.165)**	**(-1.766)**
3m	0.001	**0.083**	**0.034**	**0.083**	**0.037**	0.012	0.016	-0.020
	(0.045)	**(2.664)**	**(1.720)**	**(4.458)**	**(1.954)**	(0.486)	(0.678)	(-0.428)
6m	**0.051**	**0.041**	**0.024**	**0.047**	**0.041**	**0.048**	**0.028**	**-0.072**
	(3.928)	**(2.045)**	**(1.824)**	**(4.077)**	**(3.833)**	**(2.328)**	**(1.945)**	**(-2.879)**
Mean	**-0.110**	**-0.086**	**0.094**	**0.170**	**-0.144**	**-0.236**	**-0.151**	**-0.661**
	(-2.476)	**(-2.607)**	**(1.690)**	**(2.468)**	**(-5.636)**	**(-6.382)**	**(-2.185)**	**(-13.959)**

Notes to Table: We report hold-to-maturity option returns, expressed on a monthly basis. Returns are computed according to equations (3.1) and (3.2). Panel A reports average returns by moneyness. Panel B reports average returns by maturity. The Newey–West t-statistics are reported in parentheses. Returns that are significantly different from zero at the 10% significance level are in bold.

Table 4. Returns on Straddles and Zero-Beta Straddles

Panel A. Straddle Returns by Maturity

	ATM Options				All Options			
Maturity	Crude	NG	Heating	Gasoline	Crude	NG	Heating	Gasoline
1w	-0.151	-0.102	**-0.503**	**-0.588**	-0.086	-0.015	**-0.290**	**-0.256**
	(-0.904)	(-0.384)	**(-3.705)**	**(-2.888)**	(-1.333)	(-0.140)	**(-4.415)**	**(-2.787)**
1m	-0.006	0.056	-0.022	0.021	-0.023	**0.045**	-0.011	-0.030
	(-0.141)	(1.443)	(-0.567)	(0.384)	(-1.222)	**(1.657)**	(-0.532)	(-0.983)
3m	**0.024**	**0.047**	0.019	0.022	**0.019**	**0.039**	**0.026**	0.007
	(1.717)	**(1.876)**	(0.811)	(0.642)	**(2.081)**	**(3.146)**	**(2.316)**	(0.372)
6m	**0.022**	**0.035**	**0.039**	-0.014	**0.019**	**0.031**	**0.038**	0.002
	(1.941)	**(2.238)**	**(2.022)**	(-0.601)	**(2.771)**	**(3.919)**	**(4.188)**	(0.096)
Mean	-0.004	0.026	**-0.076**	**-0.169**	-0.004	**0.030**	**-0.037**	**-0.085**
	(-0.156)	(0.772)	**(-2.906)**	**(-2.612)**	(-0.505)	**(2.416)**	**(-3.436)**	**(-3.383)**

Panel B. Zero-Beta Straddle Returns by Maturity

	ATM				All Options			
Maturity	Crude	NG	Heating	Gasoline	Crude	NG	Heating	Gasoline
1w	-0.035	-0.009	**-0.120**	--	-0.010	-0.006	**-0.068**	**-0.020**
	(-0.902)	(-0.168)	**(-3.722)**	--	(-0.527)	(-0.150)	**(-4.113)**	**(-0.855)**
1m	-0.007	**0.088**	-0.024	0.021	-0.018	**0.058**	-0.003	-0.021
	(-0.188)	**(1.700)**	(-0.623)	(0.392)	(-0.929)	**(2.104)**	(-0.148)	(-0.664)
3m	0.068	**0.150**	0.055	0.081	**0.065**	**0.131**	**0.089**	0.031
	(1.166)	**(2.337)**	(0.805)	(0.795)	**(2.281)**	**(3.869)**	**(2.565)**	(0.523)
6m	0.139	**0.174**	**0.239**	-0.080	**0.121**	**0.184**	**0.250**	-0.017
	(1.489)	**(2.275)**	**(2.085)**	(-0.565)	**(2.840)**	**(4.260)**	**(4.509)**	(-0.169)
Mean	**0.083**	**0.132**	**0.072**	-0.007	**0.066**	**0.137**	**0.094**	-0.016
	(3.915)	**(6.275)**	**(3.090)**	(-0.190)	**(6.828)**	**(13.584)**	**(8.875)**	(-1.132)

Panel C. Straddle Returns by Moneyness

Moneyness	.86-.90	.90-.94	.94-.98	.98-1.02	1.02-1.06	1.06-1.10	1.10-1.14	Mean
Crude Oil	**0.019**	0.001	0.007	-0.004	**-0.033**	-0.024	0.000	-0.004
	(1.726)	(0.032)	(0.349)	(-0.156)	**(-1.720)**	(-1.290)	(0.020)	(-0.505)
Natural Gas	0.026	**0.054**	0.027	0.026	0.000	0.024	**0.062**	**0.030**
	(0.838)	**(1.850)**	(0.800)	(0.772)	(-0.013)	(0.913)	**(2.307)**	**(2.416)**
Heating Oil	**0.049**	**0.036**	-0.031	**-0.076**	**-0.101**	**-0.084**	**-0.063**	**-0.037**
	(2.464)	**(1.865)**	(-1.144)	**(-2.906)**	**(-3.070)**	**(-2.568)**	**(-1.932)**	**(-3.436)**
Gasoline	0.022	0.044	-0.064	**-0.169**	**-0.201**	**-0.122**	-0.066	**-0.085**
	(0.494)	(0.858)	(-1.040)	**(-2.612)**	**(-2.846)**	**(-1.762)**	(-0.900)	**(-3.383)**

Panel D. Zero-Beta Straddle Returns by Moneyness

Moneyness	.86-.90	.90-.94	.94-.98	.98-1.02	1.02-1.06	1.06-1.10	1.10-1.14	Mean
Crude Oil	**0.087**	**0.104**	**0.106**	**0.083**	0.043	-0.016	-0.015	**0.066**
	(3.658)	**(3.963)**	**(4.409)**	**(3.915)**	(1.525)	(-0.507)	(-0.569)	**(6.828)**
Natural Gas	**0.138**	**0.180**	**0.133**	**0.132**	**0.125**	**0.123**	**0.114**	**0.137**
	(4.112)	**(6.015)**	**(5.434)**	**(6.275)**	**(5.361)**	**(4.855)**	**(4.304)**	**(13.584)**
Heating Oil	**0.192**	**0.158**	**0.134**	**0.072**	0.012	0.008	0.013	**0.094**
	(6.897)	**(6.276)**	**(5.006)**	**(3.090)**	(0.456)	(0.235)	(0.355)	**(8.875)**
Gasoline	0.019	0.031	0.002	-0.007	**-0.084**	**-0.082**	**-0.087**	-0.016
	(0.549)	(0.915)	(0.056)	(-0.190)	**(-1.787)**	**(-1.961)**	**(-1.870)**	(-1.132)

Notes to Table: We report hold-to-maturity returns, expressed on a monthly basis, for energy straddles, computed using equation (4.1), and for zero-beta straddles, computed using equation (4.2). Panels A and B reports results by maturity for ATM options and all options. Panels C and D report results by moneyness. The Newey–West t-statistics are reported in parentheses. Returns that are significantly different from zero at the 10% significance level are in bold. Note that the missing values for the gasoline ATM options with one-week maturity are due to insufficient data to calculate the option betas.

Table 5. Delta-Hedged Gains

	Call Options								Put Options							
	.86-.90	.90-.94	.94-98	.98-1.02	1.02-1.06	1.06-1.1	1.1-1.14	Mean	.86-.90	.90-.94	.94-98	.98-1.02	1.02-1.06	1.06-1.1	1.1-1.14	Mean
Panel A. Crude Oil																
1m	**-0.169**	**-0.135**	**-0.118**	**-0.123**	**-0.171**	**-0.171**	**-0.157**	**-0.151**	**-0.172**	**-0.183**	**-0.186**	**-0.216**	**-0.257**	**-0.183**	**-0.128**	**-0.189**
	(-5.220)	**(-3.295)**	**(-2.471)**	**(-2.765)**	**(-3.679)**	**(-2.879)**	**(-3.936)**	**(-8.188)**	**(-5.837)**	**(-4.618)**	**(-3.875)**	**(-4.987)**	**(-5.385)**	**(-2.468)**	**(-1.742)**	**(-9.538)**
6m	**1.068**	0.260	0.322	**0.970**	0.352	0.057	**-0.221**	**0.314**	**0.710**	0.129	-0.013	0.297	-0.195	-0.319	**-0.719**	**0.180**
	(3.869)	(1.271)	(1.579)	**(4.267)**	(1.396)	(0.264)	**(-2.197)**	**(4.047)**	**(4.750)**	(0.644)	(-0.054)	(1.559)	(-0.956)	(-1.100)	**(-3.150)**	**(2.178)**
Panel B. Natural Gas																
1m	**-0.019**	**-0.016**	-0.007	0.005	-0.006	**-0.013**	-0.009	**-0.009**	**-0.026**	**-0.030**	-0.011	-0.005	**-0.019**	**-0.023**	**-0.029**	**-0.021**
	(-2.589)	**(-2.185)**	(-0.938)	(0.645)	(-0.881)	**(-1.681)**	(-1.128)	**(-2.867)**	**(-5.147)**	**(-4.829)**	(-1.496)	(-0.664)	**(-2.815)**	**(-3.401)**	**(-4.116)**	**(-7.434)**
6m	**0.038**	0.018	0.017	0.016	**0.036**	**0.073**	**0.058**	**0.040**	**-0.036**	0.002	0.004	0.005	-0.001	0.021	0.018	-0.008
	(1.923)	(0.856)	(1.148)	(1.241)	**(2.293)**	**(3.538)**	**(4.487)**	**(5.991)**	**(-3.739)**	(0.140)	(0.268)	(0.442)	(-0.056)	(0.797)	(0.613)	(-1.206)
Panel C. Heating Oil																
1m	**-0.001**	**-0.002**	**-0.003**	**-0.003**	**-0.005**	**-0.005**	**-0.004**	**-0.003**	**-0.002**	**-0.003**	**-0.005**	**-0.006**	**-0.005**	**-0.005**	**-0.005**	**-0.004**
	(-1.735)	**(-2.032)**	**(-2.382)**	**(-3.358)**	**(-4.527)**	**(-4.095)**	**(-4.881)**	**(-8.121)**	**(-2.958)**	**(-4.566)**	**(-4.965)**	**(-6.170)**	**(-3.602)**	**(-2.944)**	**(-2.788)**	**(-9.721)**
6m	**0.015**	0.000	0.004	**0.010**	**0.009**	**0.020**	-0.005	**0.006**	**0.017**	0.005	0.000	-0.001	0.008	0.000	-0.009	**0.004**
	(1.745)	(0.033)	(0.714)	**(2.725)**	**(1.728)**	**(2.805)**	(-1.615)	**(3.145)**	**(3.622)**	(1.110)	(0.022)	(-0.351)	(0.998)	(-0.023)	(-0.515)	**(2.132)**
Panel D. Gasoline																
1m	**-0.006**	**-0.003**	-0.001	-0.002	**-0.007**	**-0.010**	**-0.011**	**-0.006**	**-0.004**	**-0.004**	**-0.005**	**-0.005**	**-0.007**	**-0.010**	**-0.012**	**-0.005**
	(-3.916)	**(-1.741)**	(-0.514)	(-1.325)	**(-5.049)**	**(-6.049)**	**(-7.340)**	**(-8.508)**	**(-4.601)**	**(-3.412)**	**(-3.264)**	**(-3.299)**	**(-2.866)**	**(-2.563)**	**(-3.216)**	**(-7.374)**
6m	0.015	0.007	-0.001	0.005	0.000	**-0.011**	**-0.028**	**-0.009**	0.004	-0.004	-0.004	**-0.006**	**-0.017**	**-0.043**	**-0.060**	**-0.008**
	(0.668)	(0.492)	(-0.073)	(1.055)	(-0.099)	**(-2.391)**	**(-6.286)**	**(-3.887)**	(0.816)	(-1.167)	(-1.246)	**(-2.011)**	**(-2.329)**	**(-6.422)**	**(-10.060)**	**(-4.109)**

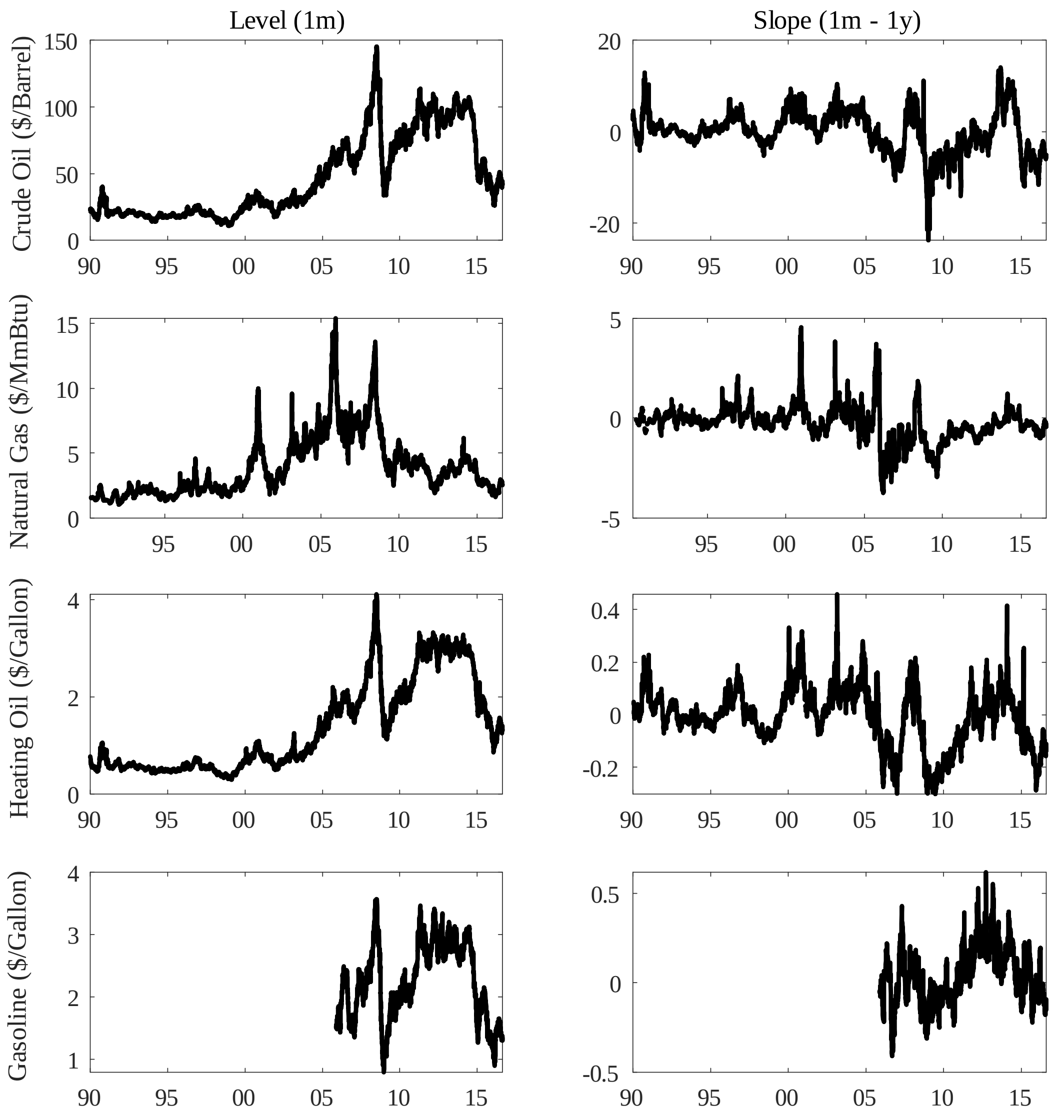
Notes to Table: We report hold-to-maturity daily-adjusted delta-hedged gains in U.S. dollars, computed using equations (4.9) and (4.10). We report the time-series average of the delta-hedged gains within each moneyness category for the one-month and six-month maturities. The Newey–West t-statistics are reported in parentheses. Returns that are significantly different from zero at the 10% significance level are in bold.

Table 6. Variance Risk Premiums

Maturity	1m	2m	3m	4m	5m	6m	7m	8m	9m	10m	11m	1y
Crude Oil	**0.006**	**-0.012**	**-0.016**	**-0.016**	**-0.015**	**-0.015**	**-0.016**	**-0.015**	**-0.014**	-0.013	-0.014	-0.016
	(5.422)	**(-3.240)**	**(-3.032)**	**(-2.448)**	**(-2.310)**	**(-2.367)**	**(-2.614)**	**(-2.099)**	**(-2.023)**	(-1.619)	(-1.546)	(-1.259)
Natural Gas	**0.027**	-0.004	**-0.020**	**-0.038**	**-0.048**	**-0.051**	**-0.047**	**-0.051**	**-0.045**	**-0.042**	**-0.050**	**-0.056**
	(14.690)	(-0.584)	**(-2.593)**	**(-5.060)**	**(-6.813)**	**(-7.496)**	**(-7.728)**	**(-7.937)**	**(-7.395)**	**(-6.852)**	**(-6.902)**	**(-7.103)**
Heating Oil	**0.022**	-0.004	**-0.011**	**-0.014**	**-0.014**	**-0.017**	**-0.018**	**-0.017**	**-0.019**	**-0.017**	**-0.021**	-0.019
	(7.826)	(-1.293)	**(-2.895)**	**(-2.956)**	**(-2.789)**	**(-3.342)**	**(-2.964)**	**(-2.575)**	**(-2.552)**	**(-2.058)**	**(-1.950)**	(-1.015)
Gasoline	0.001	-0.008	-0.006	-0.007	-0.014	-0.026	-0.025	-0.033	--	**-0.024**	--	--
	(0.276)	(-0.787)	(-0.527)	(-0.476)	(-0.742)	(-1.233)	(-1.383)	(-0.925)	--	**(-2.364)**	--	--

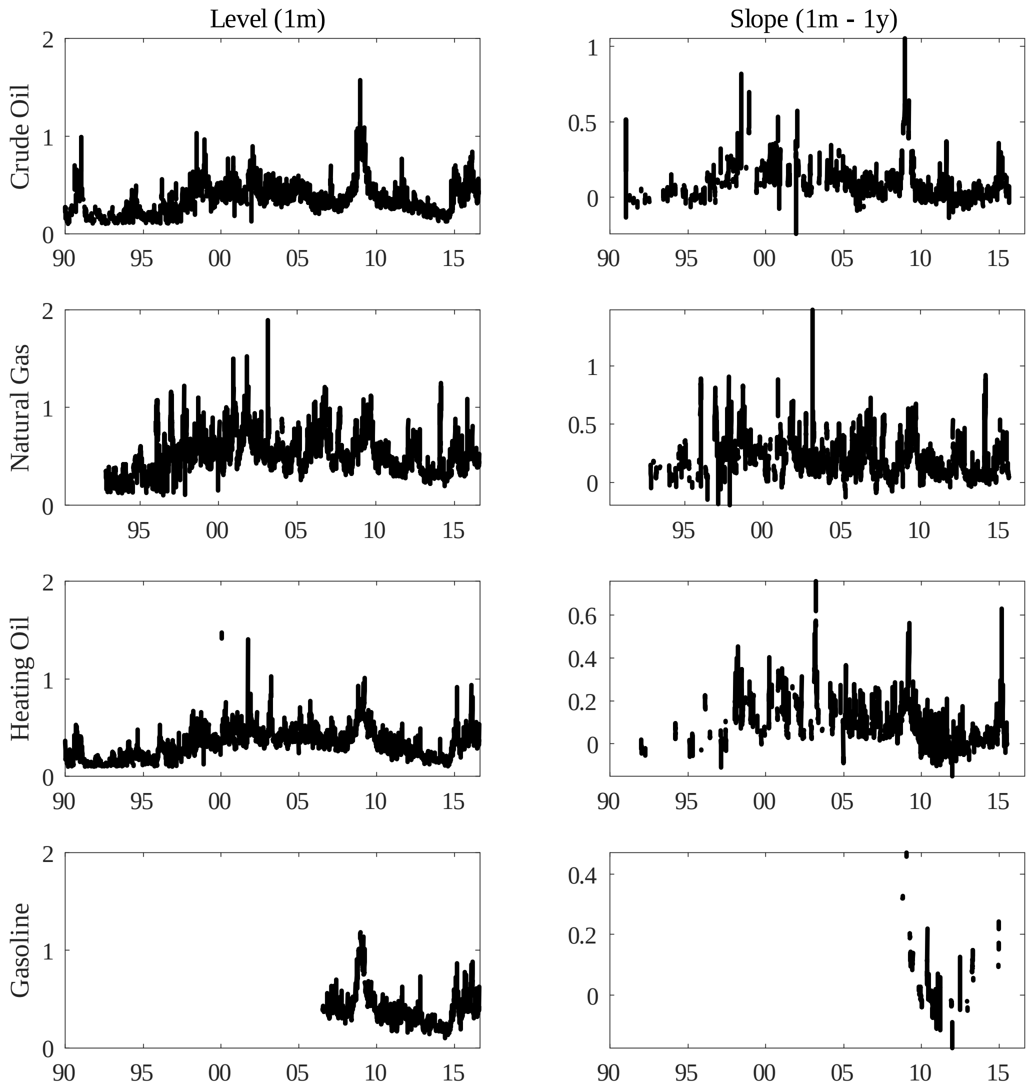
Notes to Table: We report time-series averages of the variance risk premiums for different maturities. Variances are expressed on an annual basis, and the variance risk premiums are computed using the methods in Section 5. The Newey–West t-statistics are reported in parentheses. Returns that are significantly different from zero at the 5% significance level are in bold. Note that the missing values in the gasoline market are due to insufficient data to calculate the variance risk premium.

Figure 1. Futures Prices



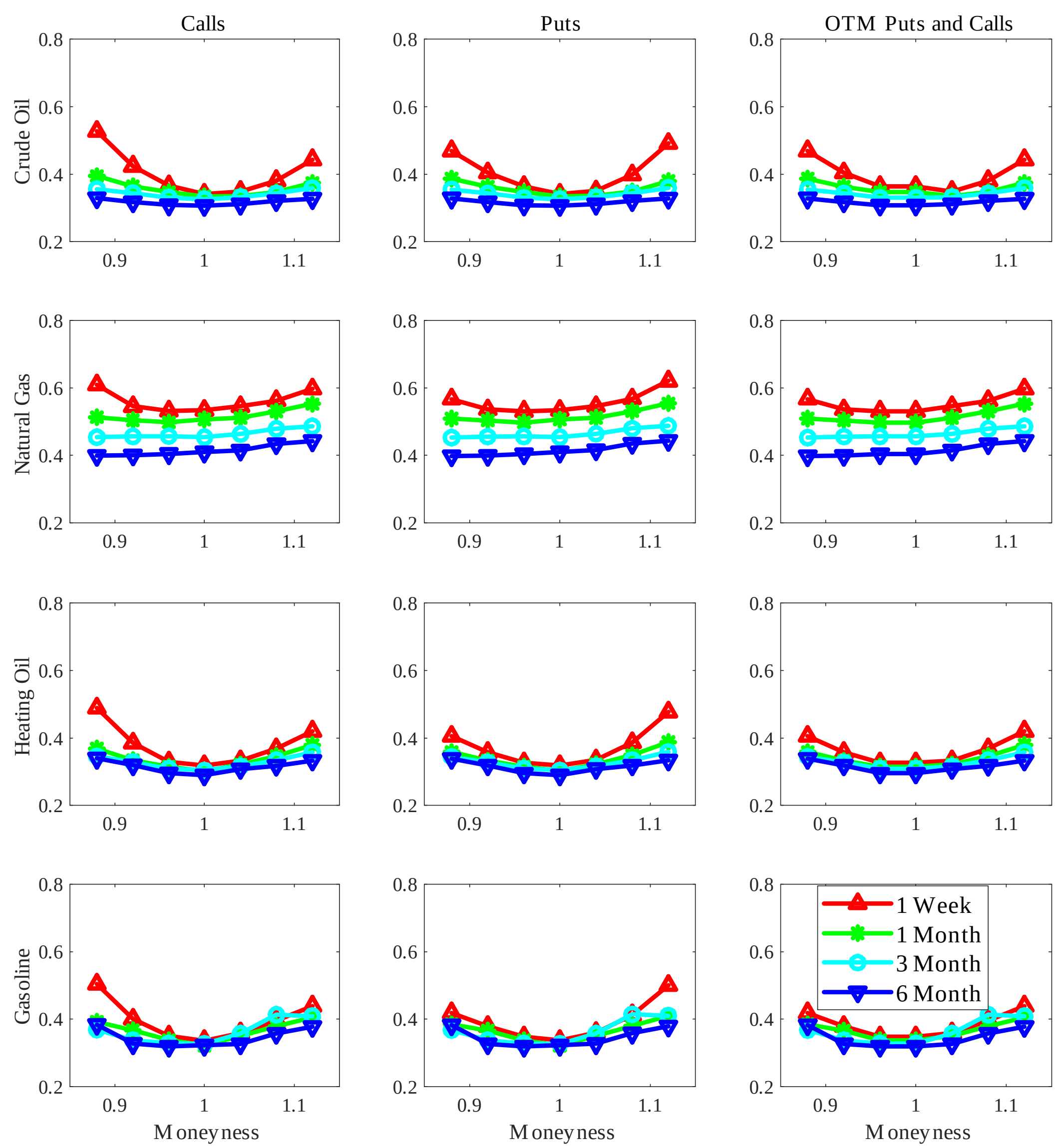
Notes to Figure: The left-side panels plot the time series of the one-month (1m) futures prices. The right-side panels plot the time series of the futures slope, computed as the difference between the one-month (1m) and one-year (1y) futures prices. The sample period starts on January 2, 1990 for crude oil and heating oil options, October 5, 1992 for natural gas options, and May 15, 2006 for gasoline options. The sample period ends on August 12, 2016.

Figure 2. ATM Option-Implied Volatility



Notes to Figure: The left-side panels plot the time series of the one-month (1m) ATM option-implied volatility. The right-side panels plot the time series of the difference between the one-month (1m) ATM implied volatility minus the one-year (1y) ATM implied volatility. Option-implied volatilities are annualized. The sample period starts on January 2, 1990 for crude oil and heating oil options, October 5, 1992 for natural gas options, and May 15, 2006 for gasoline options. The sample period ends on August 12, 2016.

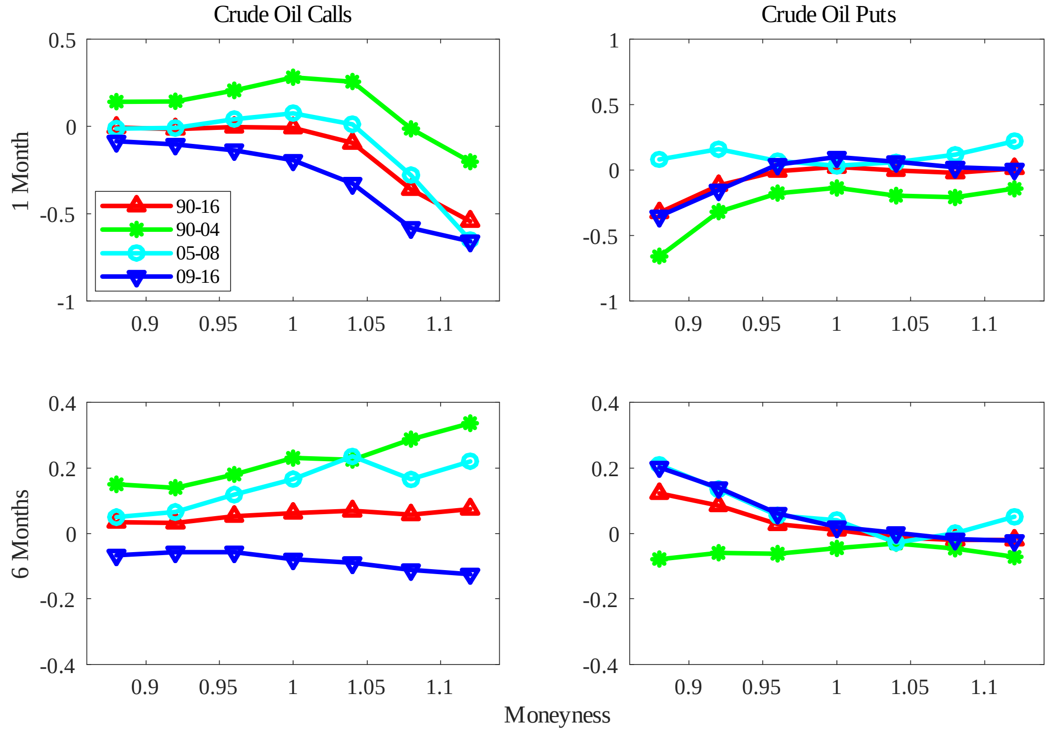
Figure 3. Option-Implied Volatility by Moneyness



Notes to Figure: We plot option-implied volatility as a function of moneyness. Moneyness is defined as the ratio of the strike price over the underlying futures price. Option-implied volatilities are annualized. The sample period starts on January 2, 1990 for crude oil and heating oil options, October 5, 1992 for natural gas options, and May 15, 2006 for gasoline options. The sample period ends on August 12, 2016.

Figure 4. Option Returns

Panel A. Crude Oil



Panel B. Natural Gas

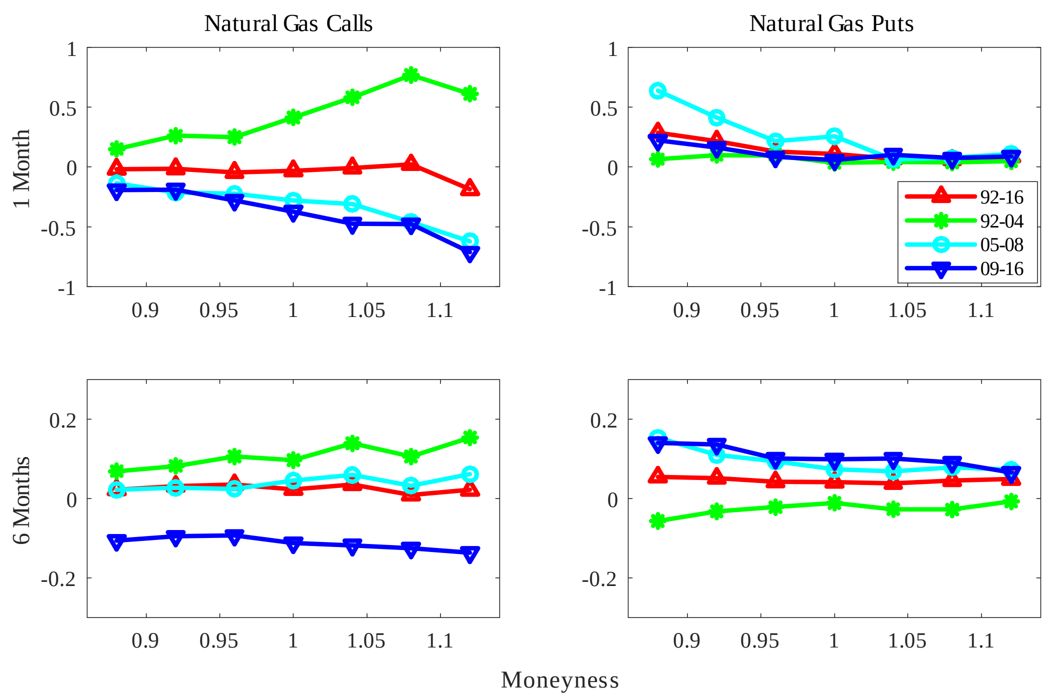
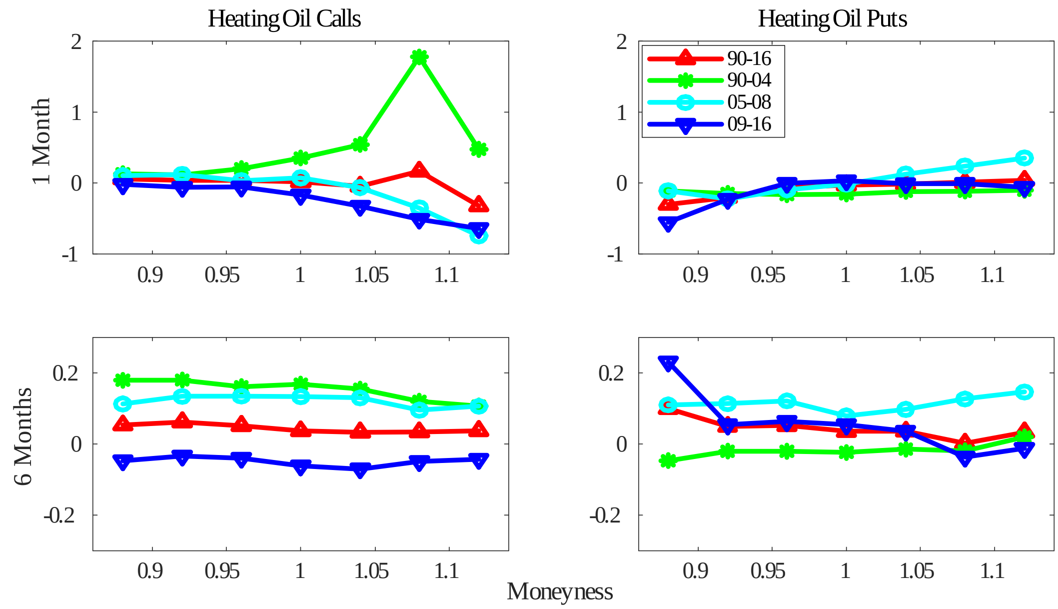
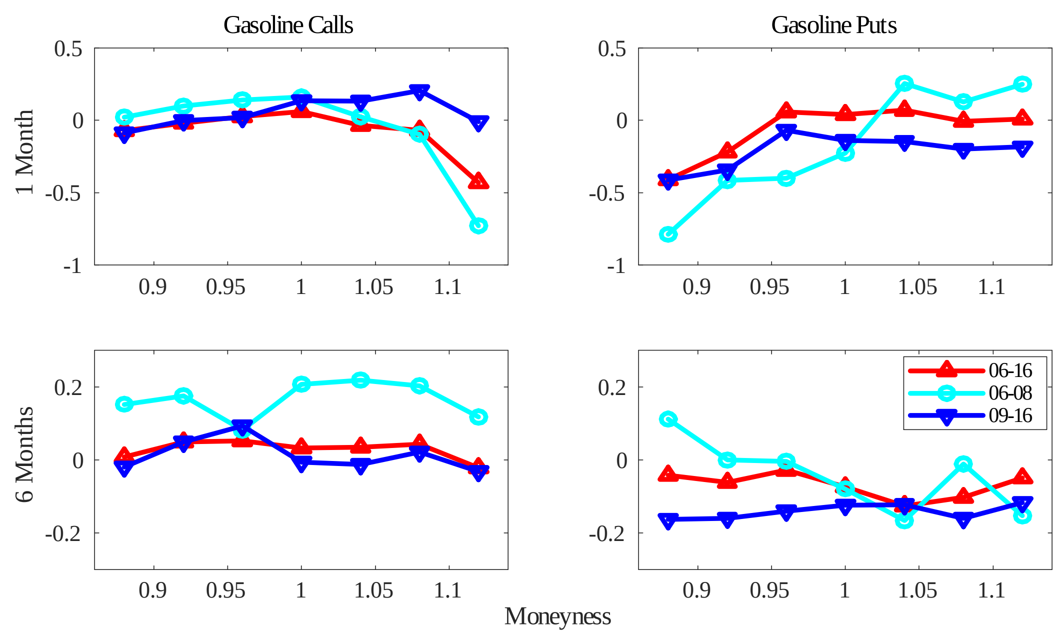


Figure 4. Option Returns (Continued)

Panel C. Heating Oil

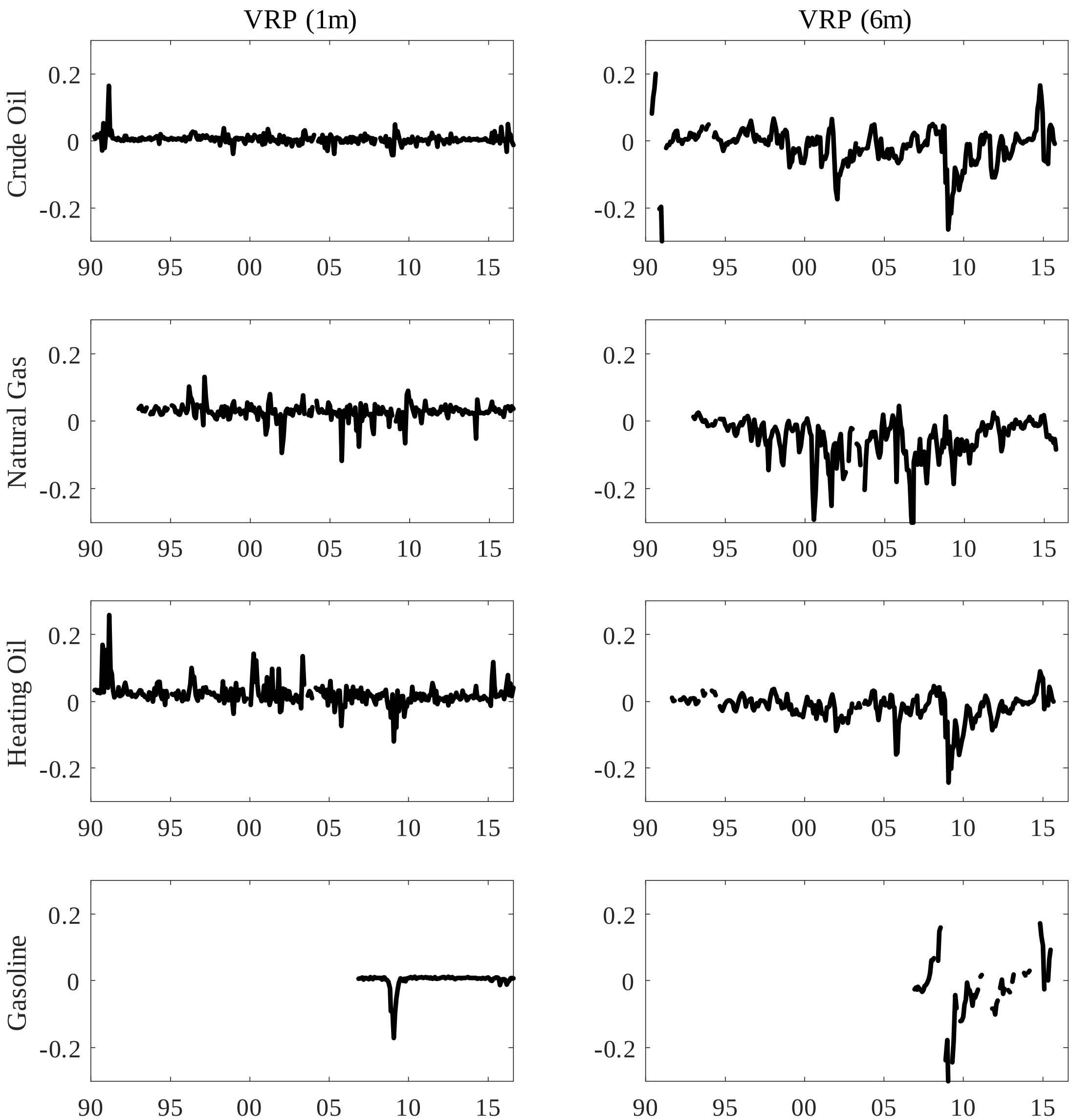


Panel D. Gasoline



Notes to Figure: We plot hold-to-maturity option returns, expressed on a monthly basis, as a function of moneyness for one-month and six-month maturity options. Moneyness is defined as the ratio of the strike price over the underlying futures price. We report results for entire sample period and three subsamples.

Figure 5. Variance Risk Premiums



Notes to Figure: We plot the time series of the variance risk premiums for the one-month (1m) and six-month (6m) maturities. Variance risk premiums are annualized and computed based on the method described in Section 5.

Appendix

Table A1. Energy Futures: Summary Statistics by Maturity

Panel A. Crude Oil

Maturity	<1m	1m-2m	2m-3m	3m-4m	4m-5m	5m-6m	6m-7m	7m-8m	8m-9m	9m-10m	10m-11m	11m-1y
Price	46.821	46.954	47.032	47.063	47.057	47.043	46.992	46.951	46.890	46.841	46.786	46.744
Ret - Arithmetic (%)	0.870	0.733	0.660	0.657	0.618	0.606	0.586	0.583	0.556	0.563	0.548	0.543
Ret - Geometric (%)	0.246	0.236	0.227	0.264	0.256	0.266	0.265	0.278	0.265	0.284	0.280	0.284
Open Interest (OI)	154052	143015	74252	52074	41555	35679	31244	27956	24992	23062	20416	18530
Trading Volume (TV)	143980	81080	29332	15375	9791	7013	5308	4206	3400	2950	2158	1943
TV/OI	0.935	0.567	0.395	0.295	0.236	0.197	0.170	0.150	0.136	0.128	0.106	0.105

Panel B. Natural Gas

	<1m	1m-2m	2m-3m	3m-4m	4m-5m	5m-6m	6m-7m	7m-8m	8m-9m	9m-10m	10m-11m	11m-1y
Price	3.903	3.980	4.043	4.082	4.113	4.138	4.161	4.178	4.187	4.185	4.186	4.220
Ret - Arithmetic (%)	1.443	1.166	1.002	0.862	0.753	0.666	0.588	0.481	0.458	0.469	0.512	0.491
Ret - Geometric (%)	0.183	0.159	0.159	0.193	0.211	0.201	0.150	0.101	0.088	0.123	0.196	0.212
Open Interest (OI)	65184	82601	55578	40406	33776	28586	24844	22059	19847	17522	15331	13791
Trading Volume (TV)	56121	29217	13945	8237	5638	4015	3091	2431	2007	1635	1319	1078
TV/OI	0.861	0.354	0.251	0.204	0.167	0.140	0.124	0.110	0.101	0.093	0.086	0.078

Panel C. Heating Oil

	<1m	1m-2m	2m-3m	3m-4m	4m-5m	5m-6m	6m-7m	7m-8m	8m-9m	9m-10m	10m-11m	11m-1y
Price	1.358	1.360	1.363	1.365	1.367	1.368	1.368	1.368	1.367	1.366	1.365	1.367
Ret - Arithmetic (%)	0.796	0.697	0.673	0.658	0.631	0.624	0.608	0.586	0.575	0.566	0.555	0.584
Ret - Geometric (%)	0.226	0.237	0.269	0.291	0.292	0.303	0.304	0.296	0.297	0.298	0.294	0.327
Open Interest (OI)	39687	46862	26370	18484	14332	11431	9233	7570	6135	4791	3746	2889
Trading Volume (TV)	25244	20177	7793	4356	2769	1821	1275	915	655	452	343	250
TV/OI	0.636	0.431	0.296	0.236	0.193	0.159	0.138	0.121	0.107	0.094	0.092	0.086

Panel D. Gasoline

	<1m	1m-2m	2m-3m	3m-4m	4m-5m	5m-6m	6m-7m	7m-8m	8m-9m	9m-10m	10m-11m	11m-1y
Price	2.250	2.238	2.233	2.228	2.223	2.218	2.214	2.212	2.211	2.211	2.213	2.212
Ret - Arithmetic (%)	0.556	0.270	0.241	0.206	0.138	0.126	0.125	0.237	0.232	0.228	0.237	0.266
Ret - Geometric (%)	-0.083	-0.286	-0.280	-0.278	-0.309	-0.301	-0.286	-0.180	-0.166	-0.168	-0.154	-0.120
Open Interest (OI)	49510	69252	35740	24527	18287	13303	9588	7279	5577	4166	3309	2593
Trading Volume (TV)	37869	34924	16079	8902	5475	3330	2011	1249	791	507	348	245
TV/OI	0.765	0.504	0.450	0.363	0.299	0.250	0.210	0.172	0.142	0.122	0.105	0.094

Notes to Table: We report daily averages of energy futures prices, returns, open interest (OI), and trading volume (TV) for energy futures by maturity. The sample period starts on January 2, 1990 for crude oil and heating oil options, October 5, 1992 for natural gas options, and May 15, 2006 for gasoline options. The sample period ends on August 12, 2016.

The Low Energy Investor: Energy Risks and the Cross Section of Stock Returns*

Virgilio Zurita†

May 24, 2019

Abstract

Energy risks carry systematic effects in the cross-section of equity portfolios and individual stocks. Using a recursive framework, we endogenously derive expected returns from investors' preferences for uncertainty and expectations about distress states of the economy, which we estimate from the crude oil options market. Increasing distress risks decrease firms' energy usage, triggering an amplification mechanism that impact expected returns. We empirically confirm this channel, stocks with lower exposure to energy risks exhibit higher returns months ahead, indicating that investors demand extra compensation to hold these assets. Energy risk exposure remains significant after controlling for stock market, commodity-specific and global risk factors, as well as abnormal media coverage. With the financialization of commodities stock return predictability increases, strengthening the commodity-equity markets link.

JEL Classification: G12, G13

Keywords: energy risks; cross-section of stock returns; options market

*For valuable comments I thank Ethan Chiang, Zhi Da, Bob Dittmar, Hitesh Doshi, Andrea Gamba, Cam Harvey, Kewei Hou, Jianfeng Hu, Kris Jacobs, Haibo Jiang, Tim Loughran, Neil Pearson, Geert Rouwenhorst, Ivan Shaliastovich, Marta Szymanowska, Bohui Zhang and seminar participants at the University of Connecticut, Houston, LSU, Nanyang Tech, SMU, UNSW, and Warwick.

†Hankamer School of Business, Baylor University. Email: virgilio_zurita@baylor.edu, Ph: (254) 710-4650.

1 Introduction

Crude oil is widely considered an important driver of economic activity. Through its use in the production chain, the evolution of the commodity greatly influences business conditions in industrialized economies. Hamilton (2013) finds that spikes in crude oil prices preceded nine out of the ten postwar U.S. recessions, and Chiang, Hughen, and Sagi (2015) document the negative effects of oil risks on the economy and in particular, energy-related stocks.[1] Earlier studies (e.g., Harvey and Siddique (2000), Gomes, Kogan and Zhang (2003)) show that state variables affecting changes in investment opportunities in turn affect a security' covariance with these state variables, and therefore its expected return. A natural question is to what extent the pricing energy risks matter for the predictability of stock returns, and do they capture information beyond common risks factors and characteristics?

In this paper, we develop a dynamic framework where expected returns of a representative oil-intensive firm are derived from investors' optimal decisions and have economic foundations. Expected returns depend on investors' preferences for uncertainty and expectations about bad states of the economy. The first feature uses a combination between investors' risk aversion and willingness to delay investments. The second feature uses the crude oil options market to estimate distress risks, the likelihood of entering a bad state of the economy.[2] We show that the pricing of energy risks do matter for explaining the return of assets. Not all energy risks are alike, the risks of spikes prove economically significant, beyond volatility and skewness risk, a finding in line with the empirical work of Hamilton (2009). Stocks with relatively low energy risk beta exhibit higher subsequent returns, and the exposure to these risks becomes more relevant with the financialization of commodity markets (e.g., Cheng and Xiong (2014), Henderson, Pearson and Wang (2015)).

[1] Table A.1 reports the top net users and providers of crude oil, with the U.S. economy leading the first group, and countries with a history of geo-political instability leading the second group. Bailey and Chan (1993) and Hamilton (1983, 2003) document the interaction between commodity risks and the economy.

[2] We use interchangeable the terms energy risk, distress risk and distress state probability to define the likelihood of entering a distress state of the economy.

The low energy investor, who is long a portfolio of stocks with low energy beta and short a portfolio of stocks with high energy beta, generates an annualized return of 16.9% and its statistical significance exceeds the Harvey, Liu, and Zhu (2016) threshold. We compute the sensitivity of individual stock returns to energy risks (ER) and find strong economic and statistical significance in the cross-section. We sort energy related stocks into decile portfolios, and compute the out of sample portfolio returns, conditioning on past ER beta information. We find that the cross sectional significance of ER beta is not affected after controlling for well-established risk factors and characteristics.

ER beta seems to capture additional information on top of stock market and commodity-specific factors such as the factors in Fama and French (1993) and Carhart (1997) as well as the commodity futures variance, skewness, basis, and open interest (see for example Chabi-Yo, Doshi and Zurita (2018), Kang, Rouwenhorst and Tang (2017), Hong and Yogo (2012), and Szymanowska, de Roon, Nijman and van den Goorbergh (2014)). Moreover, we cannot attribute the performance of the strategy to the limited attention of investors to media coverage about energy risks. We compute several textual analytics to identify and isolate information specific to energy risks. We find that the results of the strategy in times of media coverage above and below trend remain largely unchanged.

Investors care about the uncertainty of events over which the future return distribution occurs. Since the future return distribution is influenced by the state of the economy, distress risks endogenously affect an investor's utility function, the pricing kernel, and expected returns. Harvey (2017) notes the importance of risk factors derived from first principles. Our findings are consistent with Bloom (2009) and Drechsler (2013), who show that economic uncertainty is a relevant state variable proxing for investment opportunities.

Our empirical results confirm the preference-based explanation for energy risks. Due to their negative ER beta, individual stocks in decile 1 correlate negatively with increasing energy risks, the risk of entering into a bad economic state. Hence, investors demand extra compensation in the form of higher expected return to hold these stocks with lower energy risk exposure. On the other hand, with their positive ER beta, the returns of individual stocks in decile 10 correlate positively with increases in energy risks. Since stocks with positive ER beta would be viewed as relatively safer assets at times of increased economic uncertainty, investors are willing to pay higher prices for these stocks and accept lower returns.

We study the performance of the long-short strategy over time and find that it increases with the financialization of commodities in the early 2000s, marked by the increase in investment inflows to the sector (Tang and Xiong (2012) and Singleton (2014)), which suggests a link between trading in the energy derivatives market and the pricing of energy related stocks. Several authors link this increased correlation to the trading of hedge funds holding position in both markets (see for instance Buyuksahin and Robe (2014) and Cheng and Xiong (2014)). Our results complement Basak and Pavlova (2016) and Goldstein and Yang (2018), who show that financialization strengthens the commodity-equity market co-movement.

Our stylized economy features a production technology that is oil intensive, and the usage of oil for production purposes is affected by the likelihood of entering a distress state. As the risk of a distress scenario increases, the total usage of oil decreases, triggering an amplification mechanism across the economy, which ultimately impacts on expected returns. Intuitively, as the distress state probability increases, returns on investments become riskier, and therefore investors demand a higher compensation for risk. We show that this compensation for risk is positive when investors exhibit preferences for an early resolution of uncertainty.

Our paper contributes to the literature on the links between energy markets, the economy, and stock returns. Earlier studies (e.g., Chen, Rolls and Ross (1986), Driesprong, Jacobsen, and Maat (2008), Ferson and Harvey (1993), Jones and Kaul (1996)) document the effects of oil price changes on the economy and stock markets. Sockin and Xiong (2015) show that prices of key industrial commodities can serve as signals for the strength of the economy. Gao, Hitzemann, Shaliastovich, and Xu (2017) and Ready (2018) model the effects of oil volatility risks on stocks. Using crude oil option prices, Christoffersen and Pan (2017) empirically confirm the importance of oil risks in the cross section of stock returns. Chiang, Hughen, and Sagi (2015) find that a latent oil volatility factor negatively loads on the return of energy related stocks, indicating the systematic nature of energy risks. Hamilton (2003, 2009) finds that asymmetric oil shocks help explain economic activity.[3] Motivated by these findings, we develop a stylized model where the expected return of a representative oil intensive firm is endogenously affected by distress risks, proxied by the probability of drastic increases in crude oil prices. We empirically find that the sensitivity of stocks to energy risks is significant in the cross section of energy related stocks, and is robust to stock market exposure and additional risk factors and characteristics.

The structure of the paper is as follows. Section 2 introduces the model. Section 3 describes the data and methodologies. Section 4 provides an analysis of distress risks and the pricing kernel. Section 5 investigates the relation between ER beta and the cross section of stock returns. Section 6 concludes.

2 The Model

To build intuition for the empirical analysis, we first describe the stylized economic framework and the modeling of the distress state probability.

We consider a representative agent with Epstein-Zin (1989) recursive preferences, with utility V_t over consumption C_t

$$V_t = \left[(1-\delta)\, C_t^{1-\frac{1}{\psi}} + \delta \left[E_t \left(V_{t+1}^{1-\gamma} \right) \right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \tag{2.1}$$

where δ is the time discount factor, γ is the relative risk aversion, and ψ is the intertemporal

[3] Consistent with this argument, Baumeister and Kilian (2016) find no effects of sharp declines in crude oil prices for the U.S. economy.

elasticity of substitution (IES).

In this economy the production technology is oil intensive

$$Y_t = A_t^{1-\eta} O_t^{\eta} \tag{2.2}$$

where A_t embeds the normal shock, and O_t embeds the recessive shock.

This technology is similar to Sockin and Xiong (2015) where the dynamics of the model are driven by A, the productivity factor. In our model, the key implications come from the dynamics of O, the energy usage from production purposes.

Let $A_t = \exp(a_t)$. Changes in the log of productivity are normally distributed with mean μ and standard deviation σ

$$\Delta a_t = \mu + \sigma \epsilon_t \tag{2.3}$$

This specification is assumed for analytical tractability, since homogeneity of degree 1 in the factors allows the value function to be normalized by the productivity process (see for example Bloom (2009)).

The total usage of oil for production evolves according to

$$\Delta O_{t+1} = \begin{cases} (\phi + I_t) \text{ with probability } (1 - \pi_t) \\ (\phi + I_t)(1 - b) \text{ with probability } \pi_t \end{cases} \tag{2.4}$$

where ϕ is a positive parameter and π_t is the probability of entering a distress state in the next period. If a distress state materializes, then the total usage of oil for production decreases by a factor b. Otherwise, the economy is only affected by normal, symmetric, shocks. Ex-ante, the risk of a distress scenario adds uncertainty to the investment process and is measured by the distress state probability.[4]

[4]Harvey and Siddique (2000) and Lettau, Maggiori and Weber (2014) discuss the importance of investors'

The stochastic discount factor is given by

$$M_{t+1} = \delta \left(\frac{C_{t+1}}{C_t} \right)^{-\frac{1}{\psi}} \left[\frac{V_{t+1}}{E_t \left(V_{t+1}^{1-\gamma} \right)^{\frac{1}{1-\gamma}}} \right]^{\frac{1}{\psi}-\gamma} \tag{2.5}$$

Distress risks affect expected returns through the continuation utility component (in brackets) when investors exhibit preferences for the resolution of uncertainty $\left(\gamma \neq \psi^{-1}\right)$.

Recursive preferences and the homogeneity property allow us to express optimal investments in terms of distress risk

$$i_t = \frac{e^{\mu+\sigma\varepsilon_{t+1}} \left[(1-\pi_t) + \pi_t (1-b)^{1-\gamma} \right]^{\frac{\psi-1}{1-\gamma}} \delta^{\psi} \left[E_t f\left(\theta_{t+1}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{\psi-1}{1-\gamma}} - \phi}{1 + \left[(1-\pi_t) + \pi_t (1-b)^{1-\gamma} \right]^{\frac{\psi-1}{1-\gamma}} \delta^{\psi} \left[E_t f\left(\theta_{t+1}\right)^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{\psi-1}{1-\gamma}}} \tag{2.6}$$

where i_t denotes investments normalized by a_t and $f\left(\theta_{t+1}\right)$ includes state variables beyond the current period. See Appendix A for details on the solution of the model.

Note the relation (2.6) between investments and distress risks is negative when investors have preferences for the early resolution of uncertainty. The dynamics of the model also depend on the evolution of the market's perception of distress states. The probability of entering a distress state of the economy in next period π_t follows the process

$$\ln \pi_t = (1-\rho) \ln \pi + \rho \ln \pi_{t-1} + \varepsilon_t \tag{2.7}$$

We follow the vast empirical evidence surveyed in Hamilton (2013) and use market expectations of large crude oil price increases to proxy for the probability of a distress state of the economy. Using crude oil options on futures prices we estimate a forward looking measure of distress risk.

concerns about bad economic times and asymmetric risks for explaining the return of assets.

Let $F(x) = \int_{-\infty}^{x} f(x)\, dz$ denote the cumulative distribution function. Following Breeden and Litzenberger (1978), we twice differentiate the value of the option B with maturity T and interest rate r with respect to the strike price K to obtain the risk neutral distribution to the option prices $\frac{\partial F}{\partial K} = e^{rT} \frac{\partial^2 B(K,T)}{\partial K^2}$. Given that we are interested in accurately estimating large oscillations, we estimate the tails of the distribution and complete the entire density function using a generalized extreme value distribution (Figlewski (2008) and Linn, Shive and Shumway (2018)). See Appendix B for derivation details.

3 Data and Methodologies

In this section we describe the data sources, parameter calibration, and empirical methodologies.

We use the empirical evidence relating oil spikes and bad economic times and estimate the likelihood of entering a bad economic state in next period using the entire cross section of crude oil option prices. Singleton (2014) emphasizes the importance of accounting for agents' expectations in explaining movements in commodity prices. Moreover, Sockin and Xiong (2015) note that investors observe crude oil prices in a timely fashion but observe quantity variables such as global oil production with delay.

We obtain data on crude oil futures and options from the Chicago Mercantile Exchange (CME group, formerly NYMEX). The sample period is from October 1, 1990 to May 30, 2014. Market liquidity determines the start of the sample. Futures contracts expire on the third business day prior to the 25th calendar day (or the previous business day before if the 25th is not a business day) of the month that precedes the delivery month. Options on futures contracts expire three business days prior to the expiration date of futures.

We convert American option prices into European option prices following Trolle and

Schwartz (2009), who use the methodology of Barone-Adesi and Whaley (1987). We discard observations with Black (1976) implied volatility less than 1% or greater than 200%. In addition, we discard observations with prices less than $0.01 and contracts violating standard no-arbitrage constraints. We fit a spline function of 4th order to the implied volatilities to compute a dense set of interpolated volatilities. We then convert them to call prices.

We use data on aggregate economic indicators from the Federal Reserve Bank of St. Louis (FRED). We obtain open interest data on crude oil futures from the U.S. Commodity Futures Trading Commission (CFTC). We download the data on market wide factors from Kenneth French's website. The data on energy usage is from the Energy Information Administration (EIA). Figure 1 shows the usage of oil for the U.S. economy in thousand barrels per day. We adjust for seasonalities using the U.S. Census Bureau's X-12-ARIMA methodology. The figure indicates that energy usage is highly sensitive to economic downturns.

[Insert Figure 1 Here]

We obtain energy related stocks from CRSP using the four-digit SIC code classifications 1311, 1381, 1389, 2911, and 5172. We compute the sensitivities of stocks to stock market factors and commodity specific characteristics. We use the Fama and French (1993) and Carhart (1997) for stock related factors. For commodity specific variables, we compute the crude oil basis, open interest, variance risk premium, and skewness risk premium. Commodity Basis (BAS) is the log difference between the two-month maturity futures contract and the one-month maturity futures contract and adjusted by days between contracts. Commodity Open Interest Growth (OPI) is the monthly growth rate of open interest. Open interest is the total of all futures contracts entered into and not yet offset as reported by the CFTC. We follow Bakshi, Kapadia, and Madan (2003) and compute the commodity variance and skewness under the risk neutral measure. In line with the literature, we proxy for the physical risk measure using the lagged realized variance and skewness using a 60-day rolling window. We define variance risk premia (VRP) as the difference between the commodity variance under the physical measure and variance under the risk neutral measure. Similarly, we define skewness risk premia (SRP) as the difference between the commodity skewness under the physical measure and skewness under the risk neutral measure. Appendix D provides further details.

We focus the empirical analysis on energy related U.S. stocks and therefore calibrate the model to the U.S. economy Our goal is to reproduce the aggregate moments of the U.S. economy along with the observed dynamics between distress risks and business conditions. Table 1 reports the parameters of the benchmark model. These are broadly in line with the literature reviewed in Section 1.

[Insert Table 1 Here]

In the household sector, the time discount factor δ is set to 0.99. Reasonable parameterizations for risk aversion (γ) are between 1 and 10. Campanale, Castro, and Clementi (2010) argue that risk aversion values above this range are economically implausible. Consistent with the literature, we set the baseline value of γ to 8.

There is a large literature on the magnitude of the intertemporal elasticity of substitution parameter (IES) ψ. Vissing-Jorgensen and Attanasio (2003) and Guvenen (2006) estimate the IES to be in excess of 1, while Campbell (1999) estimates the IES to be below 1. Bansal and Yaron (2004) argue that low estimates of IES are based on a model without time-varying volatility. The implications of our model with time-varying distress risks are in line with the long-run risk literature, where IES larger than 1 is necessary to explain key features of asset markets (see for instance Bansal, Kiku, Shaliastovich, and Yaron (2013)).

We solve the model for different values of IES to match the relation between investments and distress risks observed in the data. Note that when risk aversion exceeds the reciprocal of IES, investors prefer early resolution of uncertainty of consumption. The level parameter

ϕ is to 0.01. We set the share of oil usage in total production η is set to 0.6, in line with Baker (2014). We set the parameter b to 0.084, the average decrease in oil usage during U.S. recessions reported by the Energy Information Administration. Distress risks correspond to the probability of large crude oil prices increases (more than 50%). We estimate the risk neutral density and compute π_t for each month using the right tail of the distribution. In (2.7) we set to $\rho = 0.97$ and $\pi = 0.048$, the corresponding autocorrelation and historical average for π_t.

Despite the simplicity of the model, this setup allows us to generate plausible moments for consumption, investment, and income, along with the observed dynamics between investments and distress risks. See Appendix C for details.

In the next two sections we present the main empirical results from the model.

4 The Distress Risk – Pricing Kernel Link

In this section we study the relation between distress risk, energy risks estimated from the cross section of crude oil options prices, and the stochastic discount factor. Figure 2 shows the time series for the distress state probability from October 1990 to May 2014. This measure determines the likelihood of entering a distress scenario in the following period as reflected by the crude oil options market. The time series exhibit a mean of 0.0484 and standard deviation of 0.0109. Its autocorrelation coefficient is 0.97.[5] The figure shows that distress risks are time-varying, increasing before or at the onset of periods of economic or geo-political distress. Once the market recovers from these periods of turmoil, the distress state probability recedes.

[5] Using a different approach, Kelly and Jiang (2014) estimate a tail risk measure from the cross-section of equity prices and find that persistence for this measure of risk is a necessary condition for generating significant predictability of returns.

[Insert Figure 2 Here]

Note that in our model, distress risks affect expected returns endogenously, and value-maximizing behavior of investors leads to an impact of distress risks on investment decisions. To highlight the main intuition for our results, we next analyze the relation between distress risks, investments, and the pricing kernel.

How do distress risks relate to investments? Intuitively, an increase in the distress state probability implies higher uncertainty about future returns on investments and therefore requires for a higher compensation for risk. We thus expect that, as the likelihood of entering a distress state of the world increases, total investments will consequentially decrease.[6]

In the model, the relation between distress risks and investments is determined by the agent's preferences for the timing of the resolution of uncertainty. In turn, the timing preference is determined by the relation between the relative risk aversion parameter (γ) and the intertemporal elasticity of substitution parameter (ψ). Investors have a preference for an early resolution of uncertainty when the relative risk aversion is higher than the reciprocal of the IES $\left(\gamma > \psi^{-1}\right)$. We show in Appendix A that in this case, as the distress state probability increases, investment decreases. While there is no clear consensus about the exact value for the relative risk aversion, conventional values in the asset pricing literature set this parameter above 1.[7] Regarding the intertemporal elasticity of substitution parameter, several authors argue for a parameter value larger than 1.[8] Finally, note that in the power utility case, a relative risk aversion parameter above 1 implies an IES parameter below 1.

In order to compare our results to the data, we solve the model for different values for the timing preference of the resolution of uncertainty. We set the benchmark specification for the relative risk aversion at 8, in line with the asset pricing literature, and then solve the model for

[6] This mechanism is also consistent with the empirical findings of Da, Huang, and Yun (2017), who document a negative relation between industrial electricity usage and future stock returns.

[7] See Campanale, Castro, and Clementi (2010) and references therein.

[8] See for instance Bansal, Kiku, Shaliastovich, and Yaron (2013).

different values of the IES. Figure 3 shows the unconditional correlation between the distress state probability and investments as a function of the data and the model. Consistent with the intuition of the model, we observe that in the data this relation is negative and around -20 percent. Figure 3 also shows the correlation between the distress state probability and investments for the model under different preferences for the resolution of uncertainty. The figure shows that, as the IES increases, the correlation between the distress state probability and investment decreases. Starting from a low value for an IES of 0.25, which implies a positive correlation of 0.61, correlation decreases in magnitude and change signs reaching its minimum value of -0.38 for an IES value of 3. Note that our benchmark calibration sets the IES at 2 and generates a correlation of -0.26, in line with the data.

It is important to note that investments increase with distress risks when the IES is below unity, a counterintuitive results. The lack of economic intuition is consistent with consumption-based models using power utility preferences. In these models, a decrease in expected growth generates an increase in asset prices, at odds with the data. The inclusion of recursive preferences resolves this problem (see for example Bansal and Yaron (2004) and Hasseltoft (2012)).

[Insert Figure 3 Here]

These results imply that agents care not only about the riskiness of the environment, but also about the timing for the resolution of uncertainty. For the model to reproduce the sign and level of the relation between distress risks and investments in the data, a preference for an early resolution of uncertainty is required.

We next study another important implication of time-varying distress risks. Since these risks reflect the likelihood of entering into bad periods of the economy, it is natural to think about their relation with the stochastic discount factor. Intuitively, as the probability of entering a bad state of the economy in the next period increases, the marginal utility of consumption increases as well. This interpretation is consistent with the relation between wealth and marginal utility of endowment based models. In these models, low levels of wealth (bad states of the economy) carry high levels of marginal utility of consumption.

In Figure 4, we plot the pricing kernel as a function of increasing probabilities of being in a distress state in the following period. The figure shows that when the likelihood of entering a distressed state is low, the pricing kernel is consequently low. As this likelihood increases, so does the pricing kernel, reflecting worse states of the economy and lower real rates. The seven probabilities imply stochastic discount factors of 0.9896, 0.9898, 0.9902, 0.9911, 0.9937, 0.9979 and 1.0048, which in turn imply real interest rates of 1.05, 1.03, 0.99, 0.90, 0.64, 0.21 and -0.48 percent respectively. These real rate values are economically plausible and intuitive. Note that the average three-month real interest rate (the three-month U.S. Treasury Bill deflated by the Consumption Price Index) during last three recessions in our sample is -0.42 percent. In addition, the results from our model are in line with the empirical findings of Linn, Shive and Shumway (2018), who obtain a monotone relation between the pricing kernel and economic states.

[Insert Figure 4 Here]

Overall, these empirical findings highlight the interesting features of the model. A preference for the early resolution of uncertainty generates the correlation between investments and distress risks observed in the data. Moreover, these time-varying risks are positively related with the pricing kernel.

5 Energy Risks Sensitivities and Stock Returns

In this section we discuss the main empirical results of the paper. We first analyze the importance of ER beta in the cross sectional pricing of energy related stocks. We then investigate on the time series of the long-short strategy and its relation to commodity financialization, as well as investors attention to crude oil news coverage.

5.1 The Cross Section of Stock Returns

We compute the exposure of individual stocks to energy risks using monthly rolling regressions of excess stock returns on the one-month-ahead ER

$$R_{i,t} = \alpha_{i,t} + \beta_{i,t}^{ER} ER_t + \beta_{i,t}^{MKT} MKT_t + \beta_{i,t}^{SMB} SMB_t + \beta_{i,t}^{HML} HML_t + \beta_{i,t}^{MOM} MOM_t + \varepsilon_{i,t}$$

We use a 60-month fixed window estimation. The first set of ER betas are obtained using the sample from September 1990 to September 1995. We then use these monthly ER betas to predict the cross-sectional stock returns in the following month (October 1995). We repeat this approach until April 2014. Each month, we form decile portfolios by sorting individual stocks based on their ER betas (β^{ER}). Decile 10 contains stocks with the highest β^{ER} during the previous month, while decile 1 contains stocks with the lowest β^{ER} during the previous month. The difference portfolio (High minus Low) results from holding a long position in the high beta ER portfolio P10 and a short position in the low beta ER portfolio P1. Table 2 reports the average monthly returns for portfolios sorted by ER betas. Column 2 reports the average ER betas for the decile portfolios, and columns 3 to 7 present the average excess returns and the alphas on the equal-weighted portfolios.

[Insert Table 2 Here]

Moving from decile 1 to decile 10, there is significant cross-sectional variation in the average values of β^{ER}. The average ER beta increases from -1.18 to 1.01. The average return difference between decile 10 (high- β^{ER}) and decile 1 (low- β^{ER}) is -1.53% per month with a six-lags Newey and West (1987) $t-$statistic of -3.50. This result indicates that stocks in the lowest β^{ER} decile generate 18.36% higher annual returns compared to stocks in the highest β^{ER} decile. In addition to the average raw returns, Table 2 presents the magnitude and statistical significance of the risk- adjusted returns (alphas) from four different factor models. In column 4, α is the intercept from the regression of the excess portfolio returns on a constant and the excess market return (MKT). Column 5 to 7 incrementally augment the model factors with a size factor (SMB), a book-to-market factor (HML), and a momentum factor (MOM). In all cases, the risk adjusted returns remain significant.

In column 4, the alpha from the long-short strategy α_{MKT} decreases from 2.66% to 1.25% per month, when moving from the lowest to the highest β^{ER} decile. The difference in alphas between the high- β^{ER} and low- β^{ER} portfolios is -1.41% per month (or -16.92% per annum) with a Newey-West $t-$statistic of -2.86. Next, we investigate the source of the 16.92% annualized risk-adjusted return difference between the high- β^{ER} and low- β^{ER} portfolios. Is it due to outperformance by low- β^{ER} stocks, underperformance by high- β^{ER} stocks, or both? For this, we focus on the economic and statistical significance of the risk-adjusted returns of decile 1 versus decile 10. As reported in Table 2, the risk adjusted returns α of decile 1 (low- β^{ER} stocks) and decile 10 (high- β^{ER} stocks) are significantly positive. Hence, we conclude that the significantly negative alpha spread between high- β^{ER} and low- β^{ER} stocks is due to both the outperformance by low- β^{ER} stocks and the underperformance by high- β^{ER} stocks.

The next three columns in Table 2 present similar alpha results from alternative factor models. The alphas α_2, α_3, and α_4 decrease almost monotonically when moving from the lowest to the highest β^{ER} decile. The difference in alphas between the high- β^{ER} and low- β^{ER} portfolios is α_2= -1.45% per month ($t = -2.96$), α_3 = -1.38% per month ($t = -2.78$), and α_4 = -1.38% per month ($t = -2.72$) for the combined four-factor model. This indicates that after controlling for the well-known market, size, book-to-market, and momentum factors, the return difference be tween the high- β^{ER} and low- β^{ER} stocks remains negative and statistically significant.

These results are consistent with a well-established literature that distinguishes risk and uncertainty. Due to their negative ER betas, the returns of individual stocks in decile 1 correlate negatively with increases in economic uncertainty, hence uncertainty-averse investors would demand extra compensation in the form of higher expected return to hold these stocks with negative β^{ER} . On the other hand, with their positive ER betas, the returns of individual stocks in decile 10 correlate positively with increases in economic uncertainty. Since stocks with positive β^{ER} would be viewed as relatively safer assets at times of increased economic uncertainty, investors are willing to pay higher prices for these stocks and accept lower returns. Note that in our model, expected returns depend on investors expectations of bad economic times, as well as their preferences for the resolution of uncertainty.

The ER betas in Table 2 are for the portfolio formation month and, not for the subsequent month over which we measure average returns. Investors may pay high prices for stocks that have exhibited high ER beta in the past in the expectation that this behavior will be repeated in the future. These expectations are natural derivations from the model in Section 2, but we now test them empirically. To this end, we examine the persistence of β^{ER} by running firm-level cross-sectional regressions of β^{ER} on lagged β^{ER} and lagged cross-sectional predictors. Specifically, for each month in the sample we run a regression across firms of 6-month to 60-month-ahead β^{ER} $(\beta_{i,t}^{ER})$ on the lagged β^{ER} $(\beta_{i,t}^{ER})$ and 8 lagged control variables defined in Section 3.

In Table 3, the second row reports the average cross-sectional coefficients on $\beta_{i,t}^{ER}$ from the multivariate cross-sectional regressions. In the regression of 6-month-ahead β^{ER} on lagged β^{ER} , the coefficient is positive, quite large, and extremely statistically significant. In other words, stocks with high β^{ER} also tend to exhibit similar features in the following 6 months. We also investigate the persistence of β^{ER} for one to five years ahead. The last five rows in Table 3 show that β^{ER} remains highly persistent up to four years into the future. These results indicate that the estimated historical ER betas successfully predict future ER betas and hence are good proxies for the true conditional betas. These results show that the ER betas are not simply characteristics of firms that result in differences in expected returns, but proxies for a source of economic uncertainty.

[Insert Table 3 Here]

So far we have tested the significance of the ER beta as a determinant of the cross-section of future returns at the portfolio level. This portfolio-level analysis does not account for information in the cross-section due to aggregation, and does not allow to control for multiple effects or factors simultaneously. Consequently, we now examine the cross-sectional relation between the ER beta and expected returns at the stock level using the Fama and MacBeth (1973) regressions. We present the time-series averages of the slope coefficients from the regressions of one-month-ahead stock returns on the ER beta with and without control variables. The average slopes provide standard Fama-MacBeth tests for determining which explanatory variables on average have nonzero premiums. We implement monthly cross-sectional regressions for the following econometric specification and incremental versions

$$R_{i,t+1} = \lambda_{0,t} + \lambda_{1,t}\beta_{i,t}^{ER} + \lambda_{X,t}\beta_{i,t}^{X} + \varepsilon_{i,t+1}$$

where $R_{i,t+1}$ is the realized excess return on stock i in month $t+1$, $\beta_{i,t}^{ER}$ is the ER beta of stock i in month t, and $\beta_{i,t}^{X}$ is a collection of stock sensitivities observable at time t for stock i.

Table 4 reports the time-series averages of the slope coefficients and the Newey-West $t-$statistics in parentheses. The univariate regression results reported in the first column indicate a negative and statistically significant relation between the ER beta and the cross-section of future stock returns. The average slope from the monthly regressions of realized returns on $\beta_{i,t}^{ER}$ alone is -0.51 with a Newey-West $t-$statistic of -2.01, and it is -0.62 ($t =$

-2.56) after controlling for stock and commodity characteristics. Its statistical significance is consistent with Harvey (2017) and Harvey, Liu, and Zhu (2016), who argue for a lower threshold for a risk factor developed from first principles, as is in the case of ER beta. The second column in Table 4 controls for the market beta ($\beta_{i,t}^{MKT}$) and the average slope on $\beta_{i,t}^{ER}$ remains negative and highly significant, whereas the average slope on $\beta_{i,t}^{MKT}$ is statistically insignificant.

[Insert Table 4 Here]

Columns 3 to 9 incrementally include additional sensitivities. Column 9 includes the full set of sensitivities. The beta with respect commodity open interest remains significant but economically small. Hong and Yogo (2012) document that the open interest of a broad set of commodities helps explain economic variables, while Boons, de Roon, and Szymanowska (2014) find similar result for stocks. Note that in all 9 cases beta ER remains economically and statistically significant.

To determine the economic significance of this average slope coefficient, we use the average values of the ER betas in the decile portfolios. Table 2 shows that the difference in $\beta_{i,t}^{ER}$ values between average stocks in the first and tenth deciles (β_{High}^{ER} - β_{Low}^{ER}) is 2.19. If a stock were to move from the first to the tenth decile of $\beta_{i,t}^{ER}$, what would be the change in that stock's expected return? The average slope coefficient of -0 .62 on $\beta_{i,t}^{ER}$ in Table 3 represents an economically significant decrease of 1.35% per month (-0.62×2.19) in the average stock's expected return for moving from the first to the tenth decile of $\beta_{i,t}^{ER}$.

We next assess the cross sectional relations between ER beta and stock returns after controlling for alternative sorting sensitivities. We use sensitivities to stock market factors (Fama and French (1993) and Carhart (1997)), as well as commodity specific variables including variance risk premium, skewness risk premium, basis, and open interest. To study these interaction we implement a bi-variate independent sort analysis.

We divide the sample into three groups based on ER beta (percentiles 25 and 75), and two groups based on the alternative beta (percentile 50). We thus generate six portfolios each time-period and compute the average for each group as well as the difference group (top minus bottom).

Table 5 reports the results from the double sorts using ER beta. The top panels (Panels A to E) report the double sorts using ER beta and stocks exposure to stock markets factors MKT, SMH, HML and MOM. The bottom panels (Panels F to H) report the double sorts using ER beta and stocks exposure to commodity specific characteristics VRP, SRP, BAS, and OPI. For ease of exposition, we focus on the top and bottom groups for each of the two measures as well as the difference between the top and bottom groups ($ER_{Diff.}$). We also report the risk adjusted return α_4 from the regression of the difference group on to the four factor model of Fama and French (1993) and Carhart (1997). In all panels, rows correspond to different groups of ER beta, given a group of the alternative beta. Columns correspond to different groups of the alternative beta, given a group of ER beta.

[Insert Table 5 Here]

The economic and statistical significance of ER beta is present in sub-groups. Panel A shows that returns from the ER beta strategy remain positive and statistically significant in the high MKT beta group with an ER beta difference portfolio monthly return of -1.21% and $t-$statistic of -3.1. The risk adjusted return shows an alpha of -1.21% monthly ($t = 2.79$). The significance of the difference group for ER beta decreases in the low MKT beta group, suggesting the performance of the beta ER strategy is relatively stronger in good economic times. Stronger results for the ER beta strategy are obtained when double sorting by sensitivities to SMB, HML, and MOM in Panels B to D. These results suggest that the long-short beta ER strategy is significant regardless of the size, value, and momentum of stocks. Panels E to H of Table 5 double sort by commodity specific variables estimated from the crude oil options and futures markets. In panel E, the risk adjusted return of the beta ER difference portfolio is -1.08% monthly ($t = -3.05$), suggesting the performance of the β^{ER} strategy increases in times of high uncertainty in the crude oil market. Similar results are obtained in Panels F to H, with the strategy losing significance when double sorting by crude oil futures basis. Yang (2013) finds that commodity basis helps explain the cross section of commodities and is related to investment shocks.

Overall these findings indicate that the well-known cross-sectional effects cannot explain the significant performance of the β^{ER} strategy. Note that when controlling for different groups of beta, most alternative sorting measures lose significance. Interestingly, the difference group (top minus bottom) of the β^{ER} strategy remains economically and statistically significant after controlling for the different groups of the alternative sorting measure.

5.2 Stock Returns, Energy Risks, and Commodity Financialization

We investigate the time series behavior the trading strategy and compute the cumulative returns of the long-short strategy. In Figure 5, we plot the cumulative returns and as well as the commodity open interest, the number of crude oil contracts outstanding at each point in time.

[Insert Figure 5 Here]

A recent literature studies the large inflow of investment capital to commodity derivatives since the early 2000s (see for instance Tang and Xiong (2012) and Henderson, Pearson and Wang (2015)). Buyuksahin and Harris (2011) document a change in two key market indicators since 2002. Total open interest in crude oil futures contracts and the ratio of positions between commercial and non-commercial investors drastically increased since 2002.

Figure 6 shows the time series for the total open interests in crude oil futures reported by the CFTC. The figure confirms the findings of previous studies. Between 1990 and 2001, the figure displays a stable pattern with an average open interest of 399,039 contracts. Between 2002 and 2012, it shows an increasing trend, with an average open interest of 1,068,368 contracts. This represents an increase of 168 percent between both periods.

Several interest results emerge. First, despite mostly positive throughout the entire sample, the performance increases since the early 2000s. This period corresponds to drastic increases in trading in the commodity derivatives markets, which is confirmed by the increase in open interest as shown in Figure 5, suggesting there is a stronger interlink between the crude oil derivatives market and the oil related companies since the financialization of commodity markets (Basak and Pavlova (2016)). The results are also consistent with a time-varying market segmentation between equity and commodity markets (Acharya, Lochstoer and Ramadorai (2013) and Bessembinder (1992)).

Second, given the strong economic and statistically significance of the long-short strategy, it is natural to ask what can empirically explain it. Do stock market, commodity specific, or global risk factors explain the returns of the ER beta strategy? If these factors can explain the spikes' risks embedded in the commodity market, standard asset pricing tests imply that the intercept equals zero. We thus regress the long-short portfolio returns on different risk factors. These include the stock market factors MKT, SMB, HML, and MOM, along with the commodity specific factors VRP, SRP, BAS, and OPI. We also include factors related to the overall economy proxied by changes in the VIX index, U.S. Treasury-Eurodollar spread (TED), and economic policy uncertainty index (EPU) of Baker, Bloom, and Davis (2016).

Table 6 reports the results for the univariate and multivariate time-series regressions. The results in Table 6 are consistent with the findings from the cross-sectional analysis in Section 5.1. Columns 1 to 11 report the univariate regressions. In all cases, the intercept of the ER beta long-short portfolio is negative and significant even after controlling for the 11 covariates individually. The intercept remains negative and significant, with an average coefficient of -0.149 and average $t-$statistic of -2.9. The univariate time series regressions suggest the ER beta strategy is marginally statistically related to the market factor and economic policy uncertainty, confirming the systematic nature of energy risks. Column 12 reports the multivariate regression. Only the commodity variance risk premium shows an economically and statistically significant slope. This result is in line with the literature studying the importance of commodity volatility for explaining economic activity and the return of assets (see for example Chiang, Hughen, Sagi (2015) and Gao, Hitzemann, Shaliastovich, and Xu (2017)). Note that including all covariates only explain 1.73% of the total variation in the strategy.

[Insert Table 6 Here]

These results are consistent with the literature on financialization of commodities, and confirm our findings in the cross-sectional analysis of Section 5.1. Tang and Xiong (2012) find an increase in the volatility of commodity futures prices as well as the correlation with equities since the early 2000s.[9] This is in line with the pattern observed in Figure 5. Moreover, results from the double sort analysis confirm these findings. In table 5, the long-short strategy seems to perform particularly well in times of high trading volume in the commodity market (high OPI group in Panel H). In addition, the strategy performs particularly well in periods of high volatility risk in the energy markets as shown in panel E.

5.3 Energy Risk Exposure and Investors Attention

A final concern about the performance of the ER relates to the attention of investors to the arrival of news about crude oil prices. Perhaps the strategy is driven not by a risk based

[9] Christoffersen, Lunde and Olesen (2018) document the empirical non-linear dependencies between commodity and equity markets.

explanation but by the limited attention of investors.[10] Recent studies investigate the effect of media coverage on crude oil prices (see for example Brandt and Gao (2017) and Loughran, McDonald, and Pragidis (2018)). Our goal is different, we ask what are the effects of crude oil news on the performance of the energy risk strategy. If our ER measure, computed using crude oil option prices and energy usage, shows different performance when news coverage is above or below trend, then ER may be just capturing the limited attention of investors to media coverage.

To this end, we search for monthly news covering crude oil prices in the Wall Street Journal, the New York Times, Dow Jones News Wires, and Reuters News Wires. We classify our search into news on crude oil price spikes, and the broader term crude oil price uncertainty. Appendix F describes the methodology. We then compute the spread of monthly news from the 12-month average news count. We examine the performance of the strategy when news coverage is above its annual trend and when it is below its annual trend.

In Table 7, Panel A reports the summary statistics. On average, the number of news referring to crude oil prices spikes or crude oil price uncertainty are 148 and 240 respectively. Between September 1995 and April 2014, the total number of news referring to crude oil spikes is 33,215, and the total number news referring to the broader search, crude oil uncertainty, is 53,794.

[Insert Table 7 Here]

In Panel B, the left column reports the results for the news search about crude oil price spikes, and the right column reports the results for a broader search, namely crude oil price uncertainty. For the case of crude oil price spikes, in times of higher media coverage when the monthly number of news exceeds the historical annual average, the monthly return is 1.57% ($t = -2.04$). Note the return of the strategy when the media coverage is below

[10] Andrei and Hasler (2015) show that investors attention can affect asset pricing.

average is similar (1.50%), which suggest that limited investors attention to crude oil news is not driving the performance of the ER strategy. We obtain similar results when using media coverage about crude oil uncertainty, but the significance of the below average group decreases.

These empirical findings are a by product of our recursive preference framework, where investors prefer to resolve uncertainty earlier rather than later. Our results are consistent with Sicherman, Loewenstein, Seppi and Utkus (2016), who find that investor attention is important in financial markets because attention affects trading, and empirically document the importance of investors preferences for the timing of information revelation.

Overall, using textual analysis to examine the effects of abnormal media coverage of crude oil prices, we find that the performance of the ER strategy remains largely unchanged under different media coverage environments, which provides further support to the risk based explanation for energy risk exposure.

6 Conclusion

To the best of our knowledge, we are the first to provide a channel through which energy risks can have systematic effects in the economy, including energy related stocks. We provide evidence that the pricing of energy risks capture additional and unaccounted information in the cross-section of portfolios and individual stocks. Investors' willingness to delay investments over time and expectations about a distress scenario drive expected returns. We estimate distress risks using crude oil options prices, and show that the negative relation between investments and distress risks observed in the data can be achieved in the model when investors exhibit preferences for the early resolution of uncertainty.

Our goal is not to implement a horse race in search for a risk factor but rather to provide an economic rationale using a dynamic model to answer why disruptions from the energy sector can have pervasive effects in the economy and individual stocks (Chiang, Hughen, and Sagi (2015)). In his American Finance Association presidential address Harvey (2017) highlights the importance of economically motivated risk factors over alternatives discovered from a purely empirical exercise.

The exposure of stocks to energy risks seems to explain the cross section of returns. We show that stocks in the lowest energy risk beta decile generate 16.9% more annualized risk-adjusted return compared to stocks in the highest risk beta decile. We find that the energy risk strategy is driven by the outperformance of stocks with negative ER betas as well as the underperformance of stocks with positive ER betas. Stocks exposure to energy risks remains strong after controlling for stock market, commodity-specific, and global risk factors. Moreover, the results from the strategy do not change under different media coverage environments. We find that with the financialization of commodity markets, the performance of the long-short strategy increases, suggesting a link between trading in the energy derivatives market and the pricing of energy-related stocks.

Appendix

A Model Solution

We numerically solve the model by value function iteration. The optimization procedure discretizes the dynamics for π using a Markov chain.[11] We maximize the value function (A.1) subject to the resource constraint (A.2), the process for oil usage (A.3), and the technology process (A.4)

$$
\begin{aligned}
V\left(O_t, \pi_t, a_t\right) &= \max_{C_t, I_t} \left\{ \left[\left(1-\delta\right) C_t^{1-\frac{1}{\psi}} + \delta \left[E_t \left(V_{t+1}^{1-\gamma} \right) \right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \right]^{\frac{1}{1-\frac{1}{\psi}}} \right\} && \text{(A.1)} \\
C_t + I_t &= A_t^{1-\eta} O_t^{\eta} && \text{(A.2)} \\
\Delta O_{t+1} &= \left(\phi + I_t\right) \text{ w.p. } \left(1-\pi_t\right) && \text{(A.3)} \\
&= \left(\phi + I_t\right)\left(1-b\right) \text{ w.p. } \pi_t && \\
a_{t+1} &= a_t + \mu + \sigma \epsilon_{t+1} && \text{(A.4)}
\end{aligned}
$$

We normalize the problem by the technology process and write the value function in terms of the state variables $(o,\, \pi)$ and subject to the oil usage process

$$
v\left(o_t, \pi_t\right) = \left\{ \begin{array}{c} \left(e^{(\mu+\sigma)} - i_t\right)^{1-\frac{1}{\psi}} + \delta \left[\left(\phi + i_t\right)^{1-\gamma} \right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \times \\ \left[\left(1-\pi_t\right) + \pi_t \left(1-b\right)^{1-\gamma} \right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \left[E_t v\left(o_{t+1}, \pi_{t+1}\right) \right]^{\frac{1-\frac{1}{\psi}}{1-\gamma}} \end{array} \right\}^{\frac{1}{1-\frac{1}{\psi}}} \qquad \text{(A.5)}
$$

When investors are indifferent about the timing for the resolution of uncertainty $\left(\gamma = \frac{1}{\psi}\right)$,

[11] See Judd (1991) and Richter, Throckmorton, Walker (2013) for an analysis on the numerical solution method. See Kopecky and Suen (2010) and Tauchen and Hussey (1991) for a treatment of markov chains for persistent processes.

optimality conditions indicate that increments in π lead to increments in i

$$i_t = \frac{e^{\mu+\sigma\varepsilon_{t+1}}\left[(1-\pi_t)+\pi_t(1-b)^{1-\gamma}\right]\delta^{\psi}E_t f(\theta_{t+1})-\phi}{1+\left[(1-\pi_t)+\pi_t(1-b)^{1-\gamma}\right]\delta^{\psi}E_t f(\theta_{t+1})} \tag{A.6}$$

Conversely, when $\left(\gamma > \frac{1}{\psi}\right)$, an increment in π leads to a reduction in the optimal i, which corresponds with (2.8) in Section 2.

B Estimation Procedure for Distress State Probabilities

The value of a call option B with underlying asset value S, strike price K, maturity T, and risk-free rate r is given by

$$\begin{aligned} B(S,K,T) &= e^{-rT}\int_0^{\infty}(S-K)^{+}Q(S,S_T,T)\,dS && \text{(B.1)} \\ &= e^{-rT}\int_K^{\infty}(S-K)\,Q(S,S_T,T)\,dS && \text{(B.2)} \end{aligned}$$

We relate option prices B to state price densities Q by taking first derivatives in (B.2) with respect to the strike price

$$\begin{aligned} \frac{\partial B(K,T)}{\partial K} &= e^{-rT}\left[-(K-K)\,Q(K,T)+\int_K^{\infty}-Q(S,S_T,T)\,dS_T\right] && \text{(B.3)} \\ &= -e^{-rT}\left[1-\widetilde{Q}(K,T)\right] && \text{(B.4)} \end{aligned}$$

where the second derivative corresponds to the state price density $e^{rT}\frac{\partial^2 B(K,T)}{\partial K^2}$ in Section 2.

We extend the left and right tails of the empirical density using a Generalized Extreme Value (GEV) distribution (Figlewski, 2008). The GEV density is governed by the shape (ξ), location (μ) and scale (σ) parameters

$$f(\xi,\mu,\sigma) = \frac{1}{\sigma}\left[1+\xi\left(\frac{S_T-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}-1} e^{-\left[1+\xi\left(\frac{S_T-\mu}{\sigma}\right)\right]^{-\frac{1}{\xi}}} \tag{B.5}$$

Given the strike price K corresponding to the quantile α of the density, we select an inner (α_0) and outer (α_1) value for the right tail.[12] We require that the cumulative probability in the tail for the empirical and GEV densities must equal. In addition, we require for the empirical and GEV densities to have the same curvature in the overlapping area

$$
\begin{aligned}
F\left(K\left(\alpha_0\right)\right) &= \alpha_0 && \text{(B.6)}\\
f\left(K\left(\alpha_0\right)\right) &= \widetilde{f}\left(K\left(\alpha_0\right)\right) && \text{(B.7)}\\
f\left(K\left(\alpha_1\right)\right) &= \widetilde{f}\left(K\left(\alpha_1\right)\right) && \text{(B.8)}
\end{aligned}
$$

where f denotes the GEV density, and $\widetilde{f}$ denotes the estimated density.

C Economic Implication of the Model

We calibrate the model parameters to the first moments of aggregate economic indicators and report the results for the second moments. Table A.2 reports the unconditional moments for income, consumption, and investments. Our sample period is from October 1990 to May 2014.

[Insert Table A.2 Here]

The table reports that the annualized volatility of income is 2.4 percent in the model and 2.3 percent in the data. Similarly, the volatility of consumption is 2.4 percent in the model and 2.2 in the data, while the volatility of investments is 7.4 percent in the model and 8.1 in the data. The correlations with income are also in line with the data. Namely, consumption correlates with income at 0.75 both in the model and the data, and investments correlates with income at 0.66 in the model and 0.86 in the data.

[12] A similar procedure is implemented for the fitting of the left tail.

D Risk Neutral Variance and Skewness

We follow Bakshi, Kapadia, and Madan (2003) to estimate the variance and skewness of the risk-neutral density function of individual securities. The risk neutral variance ($VAR_t^{\mathbb{Q}}$) and skewness ($SKE_t^{\mathbb{Q}}$) at time t for a τ-maturity contract are given by

$$VAR_t^{\mathbb{Q}} = e^{r\tau} V_t(\tau) - \mu_t(\tau)^2 \tag{D.1}$$

$$SKE_t^{\mathbb{Q}} = \frac{e^{r\tau} W_t(\tau) - 3\mu_t(\tau)^2 e^{r\tau} V_t(\tau) + 2\mu_t(\tau)^3}{\left[e^{r\tau} V_t(\tau) - \mu_t(\tau)^2\right]^{\frac{3}{2}}} \tag{D.2}$$

where $\mu_t(\tau) = e^{r\tau} - 1 - e^{r\tau} V_t(\tau)/2 - e^{r\tau} W_t(\tau)/6 - e^{r\tau} X_t(\tau)/24$ and r is the risk free rate. Bakshi, Kapadia, and Madan (2003) show that one can express the τ-maturity price of a security that pays the quadratic, cubic, and quartic return on the base security as

$$V_t(\tau) = \int_{F_t}^{\infty} \frac{2\left(1 - \ln\left(\frac{K}{F_t}\right)\right)}{K^2} C_t(\tau; K)\, dK + \int_0^{F_t} \frac{2\left(1 - \ln\left(\frac{K}{F_t}\right)\right)}{K^2} P_t(\tau; K)\, dK \tag{D.3}$$

$$W_t(\tau) = \int_{F_t}^{\infty} \frac{6\ln\left(\frac{K}{F_t}\right) - 3\ln\left(\frac{K}{F_t}\right)^2}{K^2} C_t(\tau; K)\, dK + \int_0^{F_t} \frac{6\ln\left(\frac{K}{F_t}\right) - 3\ln\left(\frac{K}{F_t}\right)^2}{K^2} P_t(\tau; K)\, dK \tag{D.4}$$

$$X_t(\tau) = \int_{F_t}^{\infty} \frac{12\ln\left(\frac{K}{F_t}\right)^2 - 4\ln\left(\frac{K}{F_t}\right)^3}{K^2} C_t(\tau; K)\, dK + \int_0^{F_t} \frac{12\ln\left(\frac{K}{F_t}\right)^2 - 4\ln\left(\frac{K}{F_t}\right)^3}{K^2} P_t(\tau; K)\, dK \tag{D.5}$$

where (D.3)-(D.5) are the time t prices of $\tau-$maturity quadratic, cubic, and quartic contracts, respectively. $C_t(\tau; K)$ and $P_t(\tau; K)$ are the time t prices of European calls and puts written on the underlying asset with strike price K and expiration τ periods from time t.

E Media Coverage

We search for financial news pertaining to crude oil risks. Importantly, we restrict our search to articles where words referring to crude oil risks are within 5 words of distance. A preliminary inspection suggest that the location of the three arguments within the news article matters, and that without a nearest neighbor algorithm restriction the results bring news that are not specifically related to our target.[13] We drop news when this restriction is violated. Because of this restriction, manual validation reveals that our success rate upwards 90% in selecting specialized news. Given the importance of crude oil in the general news media, we find that using an unrestricted location algorithm yields results that are unrelated to crude oil prices.

We use Factiva to search for two groups of news: Crude oil price spikes, and crude oil price uncertainty. Our source includes the Wall Street Journal, the New York Times, Dow Jones News Wires, and Reuters News Wires. Based on our initial manual search, the following set of words encompass most of the financial news articles, and adding additional words does not change our overall results.

We define the spike set of related words: spike*, large increase, shoot up, much higher, climb*, drastic increase, massive increase.

We define the uncertainty set of related words: risk*, volatil*, fear*, uncertain*.

Terms that end with * search for words sharing the first characters (e.g.: spike* searches for spike, spikes, spiked).

[13] Manual inspection reveals that using simpler word connectors such as AND, OR, or using longer word window yields results that are not specifically related to crude oil risks.

References

[1] Acharya, Viral, Lochstoer, Lars, and Tarun Ramadorai. 2013. “Limits to arbitrage and hedging: Evidence from commodity markets”, *Journal of Financial Economics*, 109, 441-465.

[2] Andrei, Daniel and Michael Hasler. 2015. “Investor Attention and Stock Market Volatility”, *Review of Financial Studies*, 28, 1, 33-72.

[3] Bailey, Warren, and K. C. Chan. 1993. “Macroeconomic Influences and the Variability of the Commodity Futures Basis”, *Journal of Finance,* 48, 2, 555–573.

[4] Bakshi, Gurdip, Kapadia, Nikunj, and Dilip Madan. 2003. “Stock return characteristics, skew laws, and the differential pricing of individual equity options”, *Review of Financial Studies,* 16, 101–143.

[5] Bansal, Ravi, and Amir Yaron. 2004. “Risks for the Long Run: A Potential Explanation of Asset Pricing Puzzles”, *Journal of Finance,* 59, 1481–1509.

[6] Bansal, Ravi, Kiku, Dana, Shaliastovich, Ivan, and Amir Yaron. 2014. “Volatility, the Macroeconomy and Asset Prices”, *Journal of Finance,* 69, 6, 2471–2511.

[7] Barone-Adesi, Giovanni, and Robert Whaley. 1987. “Efficient Analytic Approximation of American Option Values”, *Journal of Finance,* 42, 301-320.

[8] Baumeister, Christiane, and Lutz Kilian. 2016. “Lower Oil Prices and the U.S. Economy: Is This Time Different?”, Brookings Papers on Economic Activity, 287-336.

[9] Basak, Suleyman and Ana Pavlova. 2016. “A Model of Financialization of Commodities”, *Journal of Finance*, 71, 4, 1511—1556.

[10] Bessembinder, Hendrik. 1992. “Systematic risk, hedging pressure, and risk premiums in futures markets”, *Review of Financial Studies*, 5, 637-667.

[11] Bessembinder, Hendrik, and Kalok Chan. 1992. “Time Varying Risk Premia and Forecastable Returns in Futures Markets”, *Journal of Financial Economics*, 32, 169-193.

[12] Black, Fischer. 1976. “The Pricing of Commodity Contracts”, *Journal of Financial Economics,* 3, 167-179.

[13] Bloom, Nicholas. 2009. “The Impact of Uncertainty Shocks”, *Econometrica*, 77, 3, 623–685.

[14] Boons, Martijn, de Roon, Frans, and Marta Szymanowska. 2014. "The Price of Commodity Risk in Stock and Futures Markets", Working Paper.

[15] Brandt, Michael, and Lin Gao. 2017. "Macro Fundamentals or Geopolitical Events? A Textual Analysis of News Events for Crude Oil", Working Paper.

[16] Breeden, Douglas, and Robert Litzenberger. 1978. "Prices of State Contingent Claims Implicit in Options Prices", *Journal of Business*, 51, 621–651.

[17] Büyükşahin, Bahattin, and Jeffrey Harris. 2011. "Do Speculators Drive Crude Oil Futures Prices?", *Energy Journal*, 32, 167-202.

[18] Büyükşahin, Bahattin, and Michel Robe. 2014. "Speculators, commodities and cross-market linkages", *Journal of International Money and Finance*, 42, 38-70.

[19] Carr, Peter, and Jiming Yu. 2012. "Risk, Return, and Ross Recovery", *Journal of Derivatives,* 20.

[20] Carhart, Mark. 1997. "On Persistence in Mutual Fund Performance", *Journal of Finance,* 52, 1, 57–82.

[21] Chabi-Yo, Fousseni, Doshi, Hitesh, and Virgilio Zurita. 2018. "Never a Dull Moment: Entropy Risk in Commodity Marketss", Working Paper.

[22] Chen, Nai-Fu, Rolls, Richard, and Stephen Ross. 1986. "Economic Forces and the Stock Market", *Journal of Business*, 59, 3, 383-403.

[23] Cheng, Ing-Haw, and Wei Xiong. 2014. "The Financialization of Commodity Markets", *Annual Review of Financial Economics*, 6, 419-441.

[24] Chiang, I-Hsuan Ethan, Hughen, Keener, and Jacob Sagi. 2015. "Estimating Oil Risk Factors Using Information from Equity and Derivatives Markets", *Journal of Finance*, 70, 2, 769–804.

[25] Christoffersen, Peter, Lunde, Asde, and Kasper Olesen. 2018. "Factor Structure in Commodity Futures Return and Volatility", *Journal of Financial and Quantitative Analysis*, forthcoming.

[26] Christoffersen, Peter, and Xuhui Pan. 2017. "Oil volatility risk and expected stock returns", *Journal of Banking and Finance*, 1-22.

[27] Da, Zhi, Huang, Dayong, and Hayoung Yun. 2017. "Industrial Electricity Usage and Stock Returns", *Journal of Financial and Quantitative Analysis*, 52, 1, 37-69.

[28] Drechsler, Itamar. 2013. "Uncertainty, time-varying fear, and asset prices", *Journal of Finance,* 68, 1843–1889

[29] Driesprong, Gerben, Jacobsen, Ben, and Benjamin Maat. 2008. "Striking oil: Another puzzle?", *Journal of Financial Economics*, 89, 307-327.

[30] Edwards, Franklin, and Michael Canter. 1995. "The Collapse of Metallgesellschaft: Unhedgeable Risks, Poor Hedging Strategy or Just Bad Luck?", *Journal of Futures Markets*, 15.

[31] Epstein, Larry, and Stanley Zin. 1989. "Substitution, Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: A Theoretical Framework", *Econometrica*, 57, 4, 937-969.

[32] Fama, Eugene, and Kenneth French. 1993. "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics,* 33, 1,3–56.

[33] Ferson, Wayne, and Campbell Harvey. 1993. "The Risk and Predictability of International Equity Returns", *Review of Financial Studies,* 6, 3, 527–566.

[34] Figlewski, Stephen. 2008. "Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio", Volatility and Time Series Econometrics: Essays in Honor of Robert Engle. Oxford University Press.

[35] Gao, Lin, Hitzemann, Steffen, Shaliastovich, Ivan, and Lai Xu. 2017. "Oil Volatility Risk", Working Paper.

[36] Goldstein, Itay and Liyan Yang. 2018. "Commodity Financialization and Information Transmission", Working Paper.

[37] Gomes, Joao, Kogan, Leonid, and Lu Zhang. 2003. "Equilibrium Cross Section of Returns", *Journal of Political Economy*, 111, 4, 693-732.

[38] Hamilton, James. 1983. "Oil and the Macroeconomy Since World War II", *Journal of Political Economy,* 91, 228-48.

[39] ———. 2003. "What Is an Oil Shock?", *Journal of Econometrics*, 113, 2, 363-98.

[40] ———. 2009. "Causes and Consequences of the Oil Shock of 2007-2008", Brookings Papers on Economic Activity, 215-261.

[41] ———. 2013. "Historical Oil Shocks", Handbook of Major Events in Economic History, Routledge International Handbooks, Part III-21.

[42] Harvey, Campbell. 2017. "Presidential Address: The Scientific Outlook in Financial Economics", *Journal of Finance*, 77, 4, 1399-1440.

[43] Harvey, Campbell, Liu, Yan, and Heqing Zhu. 2016. " ...and the cross-section of expected returns", *Review of Financial Studies*, 29, 5-68.

[44] Harvey, Campbell and Akhtar Siddique. 2000. "Conditional skewness in asset pricing tests", *Journal of Finance*, 55,1263-1295.

[45] Hasseltoft, Henrik. 2012. "Stocks, Bonds, and Long-Run Consumption Risks", *Journal of Financial and Quantitative Analysis*, 47, 309-332.

[46] Henderson, Brian, Pearson, Neil, and Li Wang. 2015. "New Evidence on the Financialization of Commodity Markets", *Review of Financial Studies*, 28, 5, 1285-1311.

[47] Hong, Harrison, and Jeremy Stein. 2007. "Disagreement and the Stock Market", *Journal of Economic Perspectives,* 21, 2, 109–128.

[48] Hong, Harrison, and Motohiro Yogo. 2012. "What Does Futures Market Interest Tell Us about the Macro-economy and Asset Prices?", *Journal of Financial Economics*, 105, 3, 473-490.

[49] Jones, Charles , Gautam Kaul. 1996. "Oil and the stock markets", *Journal of Finance*, 51, 2, 463–491 .

[50] Judd, Kenneth. 1998. "Numerical Methods in Economics", MIT Press, Cambridge.

[51] Kelly, Bryan, and Hao Jiang. 2014. "Tail Risk and Asset Prices", *Review of Financial Studies,* 27, 10, 2841-2871.

[52] Kilian, Lutz. 2008. "Exogenous Oil Supply Shocks: How Big Are They and How Much Do They Matter for the U.S. Economy?", *Review of Economics and Statistics*, 90, 216-240.

[53] Kopecky, Karen, and Richard Suen. 2010. "Finite State Markov-chain Approximations to Highly Persistent Processes", *Review of Economic Dynamics,* 13, 701-714.

[54] Lettau, Martin, Maggiori, Matteo, and Michael Weber. 2014. "Conditional risk premia in currency markets and other asset classes", *Journal of Financial Economics*, 114, 197-225.

[55] Linn Matthew, Shive, Sophie and Tyler Shumway. 2018. "Pricing kernel monotonicity and conditional information", *Review of Financial Studies*, 31, 2, 493-531.

[56] Loughran, Tim, McDonald, Bill, and Ioannis Pragidis. 2018. "Assimilation of Oil News into Prices", Working Paper.

[57] Ready, Robert. 2017. "Oil prices and the stock market", *Review of Finance* (forthcoming).

[58] Richter, Alexander, Throckmorton, Nathaniel, and Todd Walker. 2013. "Accuracy, Speed and Robustness of Policy Function Iteration", *Computational Economics*, 1-32.

[59] Sicherman, Nachum, Loewenstein, George, Seppi, Duane, and Stephen Utkus. 2016. "Financial Attention", *Review of Financial Studies*, 29, 4, 863-897.

[60] Singleton, Kenneth. 2014. "Investor Flows and the 2008 Boom/Bust in Oil Prices", *Management Science,* 60, 2, 300–318.

[61] Sockin, Michael, and Wei Xiong. 2015. "Informational Frictions and Commodity Markets", *Journal of Finance*, 70, 5, 2063-2098.

[62] Szymanowska, Marta, de Roon, Frans, Nijman, Theo, and Rob van den Goorbergh. 2014. "An Anatomy of Commodity Futures Risk Premia", *Journal of Finance*, 69, 1, 453–482.

[63] Tang, Ke, and Wei Xiong. 2012. "Index Investment and the Financialization of Commodities", *Financial Analysts Journal,* 68, 54-74.

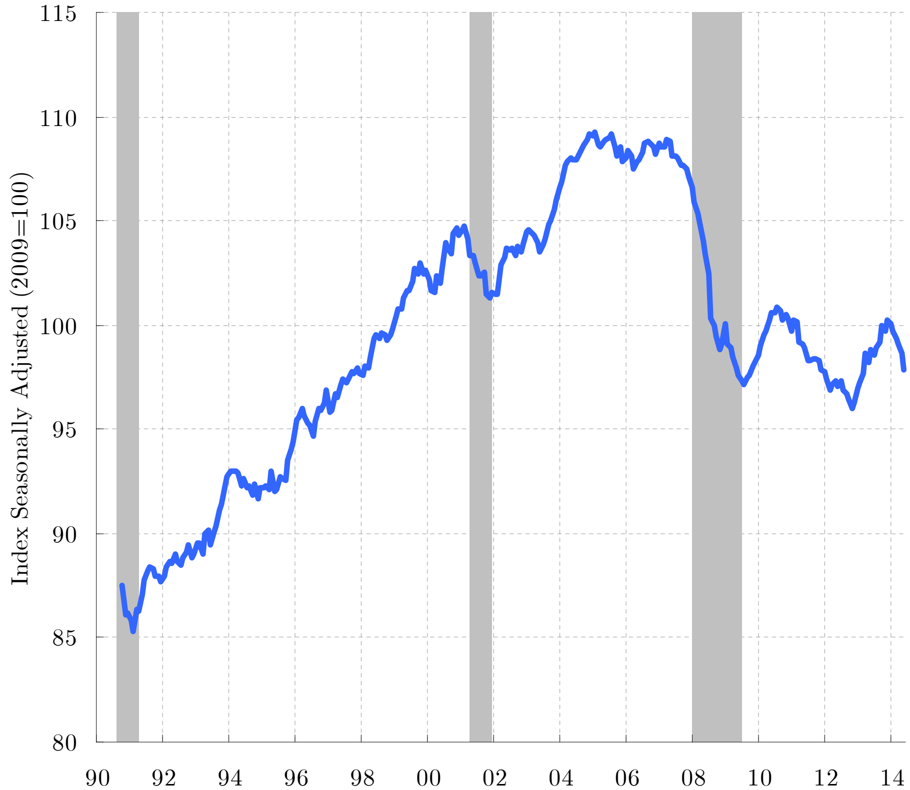
[64] Tauchen, George, and Robert Hussey. 1991. "Quadrature-based Methods for Obtaining Approximate Solutions to Nonlinear Asset Pricing Models", *Econometrica*, 59, 317–396.

[65] Trolle, Anders, and Eduardo Schwartz. 2009. "Unspanned Stochastic Volatility and the Pricing of Commodity Derivatives", *Review of Financial Studies*, 22, 4423-4461.

[66] Vissing-Jørgensen, Annette, and Orazio Attanasio. 2003 "Stock-Market Participation, Intertemporal Substitution, and Risk-Aversion", *American Economic Review*, 93, 2, 383-391.

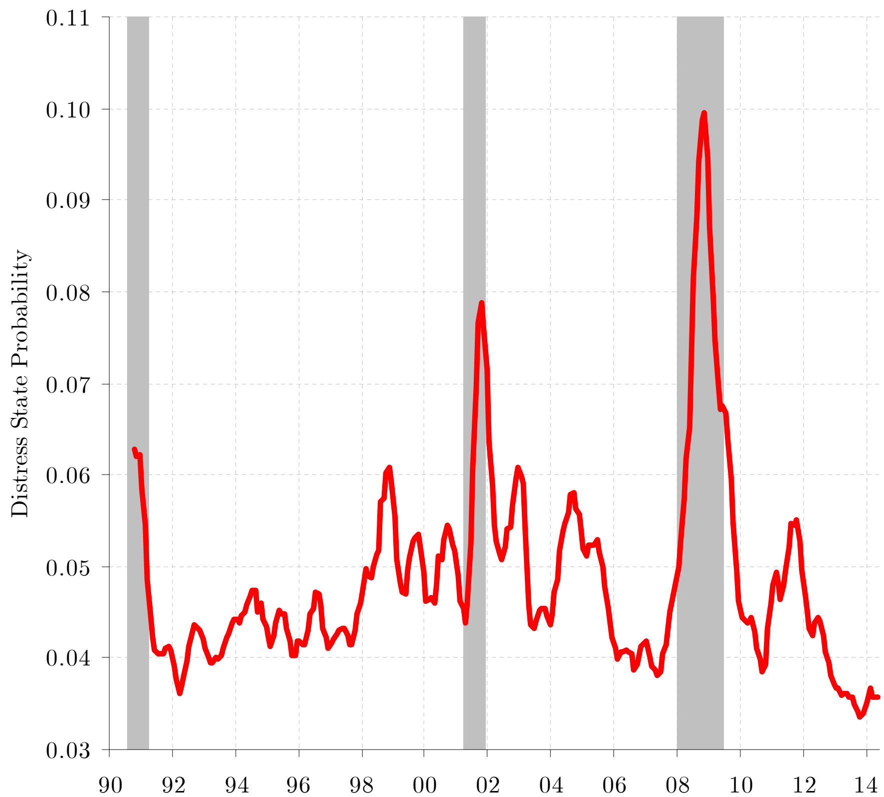
[67] Yang, Fan. 2013. "Investment Shocks and the Commodity Basis Spread", *Journal of Financial Economics*, 110, 164-184.

Figure 1: Energy Usage



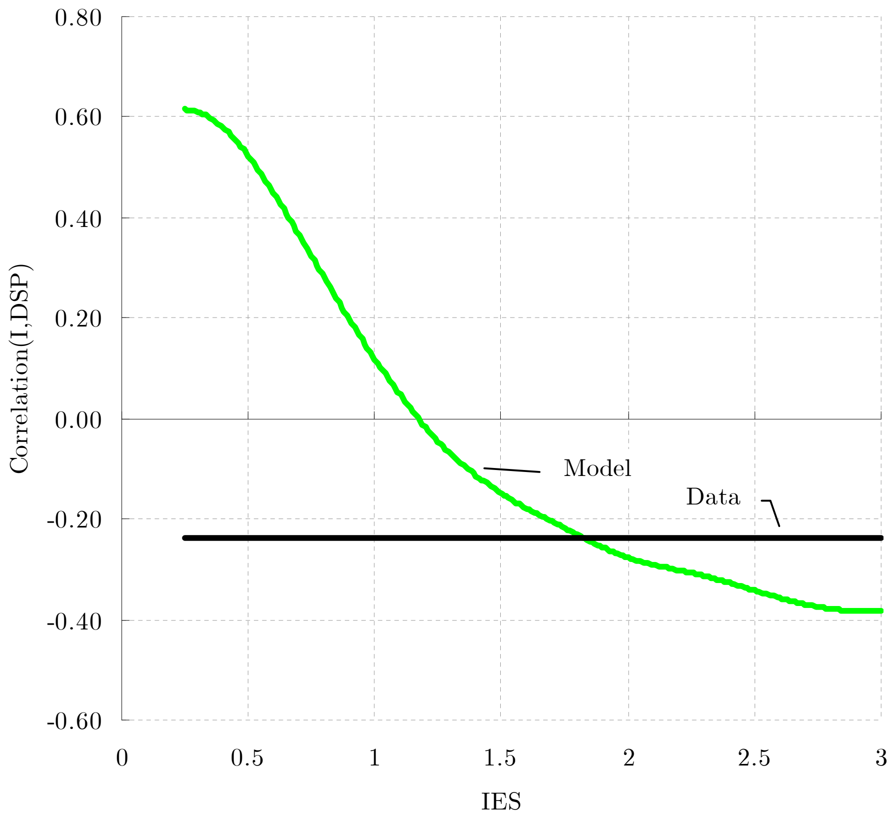
Notes to Figure: We plot the total usage of oil, in thousand barrels per day, from the Energy Information Administration (EIA). The time series is seasonally adjusted and indexed to 2009 (2009=100). The vertical axis displays the index values. U.S. recessions in gray bars from the National Bureau of Economic Research (NBER). The sample period is from October 1990 to May 2014.

Figure 2: Distress State Probabilities



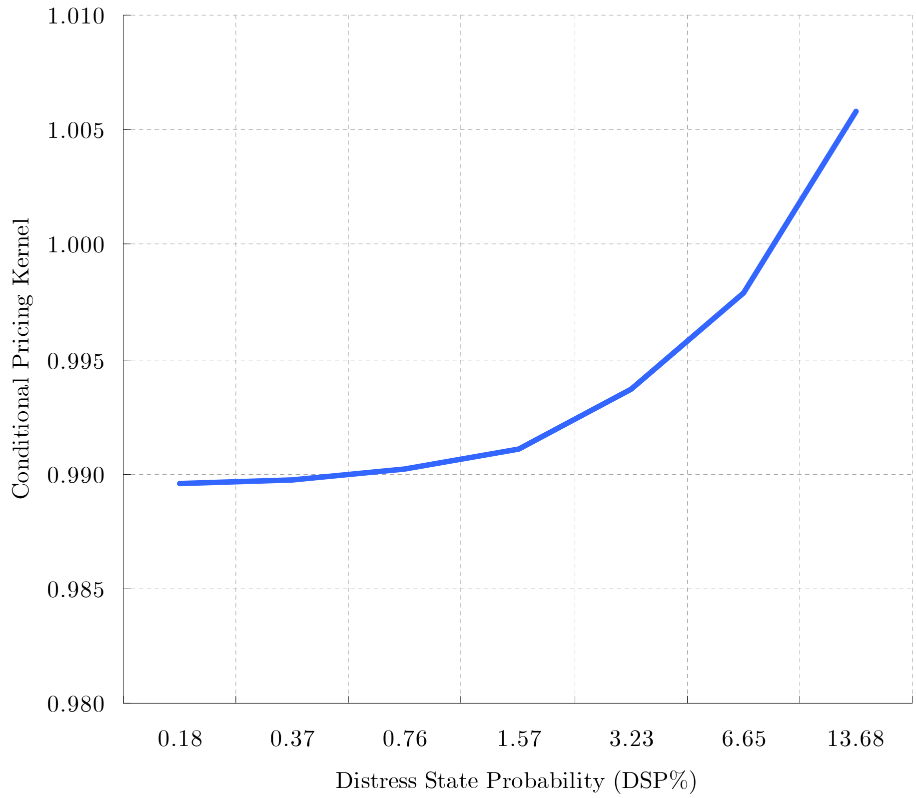
Notes to Figure: We plot the time series for the conditional probability of entering a distress state in the next period. U.S. recessions in gray bars from the National Bureau of Economic Research (NBER). The sample period is from October 1990 to May 2014.

Figure 3: Distress State Probability and Investment Correlation



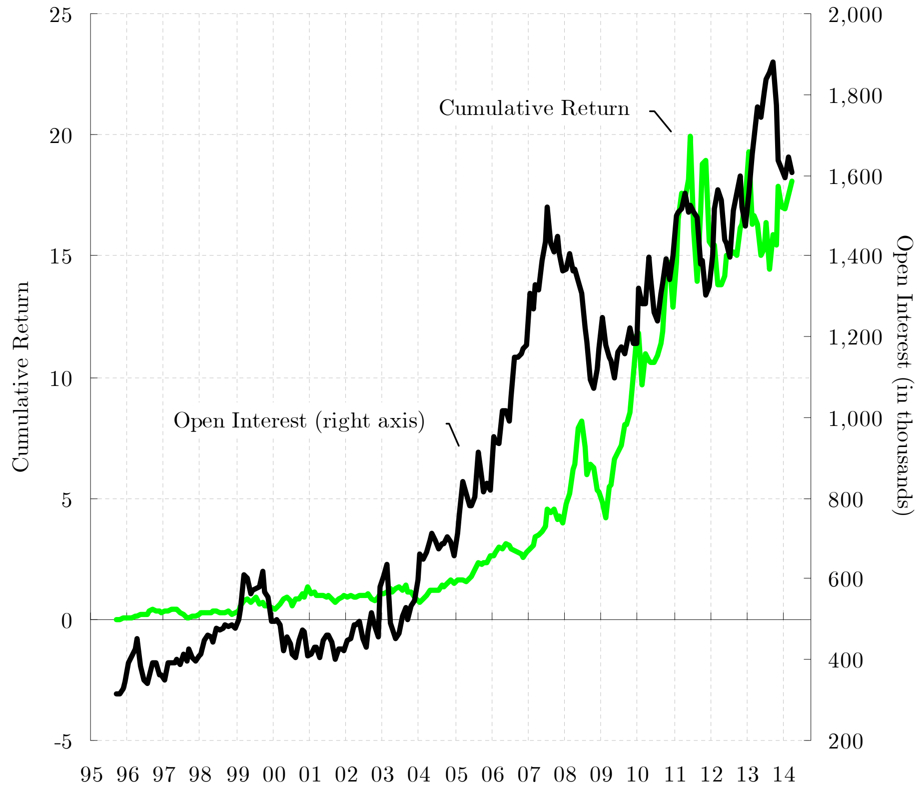
Notes to Figure: We plot the correlation between the distress state probability (DSP) and investments (I) as a function of the intertemporal elasticity of substitution (IES), given a relative risk aversion parameter of 8 (benchmark specification). The vertical axis measures correlation. The horizontal axis displays the intertemporal elasticity of substitution parameter.

Figure 4: Conditional Pricing Kernel



Notes to Figure: We plot the conditional pricing kernel for different likelihoods of entering a recessive state in the next period. The vertical axis measures the conditional pricing kernel. The horizontal axis displays the distress state probability (DSP). Distress state probability $\times 100$.

Figure 5: ER Beta Strategy and Open Interest



Notes to Figure: We plot the cumulative returns of the long-short ER beta strategy and the crude oil open interest of futures contracts (right axis) from September 1995 to April 2014.

Table 1: Model Parameter Calibration

Parameter	Value
δ	0.99
γ	8
ψ	2
ϕ	0.01
η	0.6
b	0.084

Notes to Table: We report the parameters for the benchmark calibration of the model. The second row reports the time discount factor parameter (δ). The third row reports the relative risk aversion parameter (γ). The fourth row reports the intertemporal elasticity of substitution parameter (IES). The fifth row reports the level parameter (ϕ). The sixth row reports the share of oil usage in total production (η). The seventh row reports oil usage reduction factor (b).

Table 2: Portfolio Returns

Decile	β^{ER}	Return	α_{MKT}	α_2	α_3	α_4
High	1.01	0.0133	0.0125	0.0121	0.0109	0.0130
		(2.55)	(2.21)	(2.17)	(2.00)	(2.49)
9	0.42	0.0176	0.0170	0.0170	0.0160	0.0174
		(3.69)	(2.81)	(2.82)	(2.79)	(3.02)
8	0.24	0.0139	0.0133	0.0133	0.0125	0.0136
		(3.09)	(2.19)	(2.19)	(2.11)	(2.27)
7	0.13	0.0165	0.0157	0.0157	0.0146	0.0166
		(3.65)	(2.77)	(2.78)	(2.77)	(3.09)
6	0.03	0.0147	0.0141	0.0141	0.0132	0.0151
		(3.37)	(2.50)	(2.53)	(2.47)	(2.74)
5	-0.07	0.0179	0.0172	0.0173	0.0165	0.0184
		(4.02)	(3.13)	(3.13)	(3.02)	(3.35)
4	-0.18	0.0173	0.0168	0.0167	0.0153	0.0172
		(3.53)	(2.88)	(2.86)	(2.68)	(2.90)
3	-0.31	0.0196	0.0183	0.0182	0.0170	0.0193
		(3.60)	(2.65)	(2.66)	(2.59)	(2.90)
2	-0.53	0.0226	0.0216	0.0215	0.0206	0.0225
		(3.97)	(2.87)	(2.87)	(2.80)	(2.97)
Low	-1.18	0.0286	0.0266	0.0266	0.0247	0.0268
		(4.77)	(3.70)	(3.67)	(3.53)	(3.78)
High-Low		-0.0153	-0.0141	-0.0145	-0.0138	-0.0138
		(-3.50)	(-2.86)	(-2.96)	(-2.78)	(-2.72)

Notes to Table: We report the monthly excess returns by portfolio deciles. We report the risk adjusted returns (alphas) from the regression onto the incremental risk factors. We report in parentheses the Newey-West corrected $t-$statistics.

Table 3: Persistence of Energy Risk Beta

	Months Ahead					
	6	12	24	36	48	60
Intercept	-0.0204	-0.0106	-0.0228	-0.0717	-0.1161	-0.1221
	(-1.31)	(-0.54)	(-0.77)	(-2.05)	(-2.67)	(-2.56)
β^{ER}	0.8529	0.7223	0.4817	0.2501	0.1151	0.0465
	(39.03)	(22.68)	(10.94)	(5.79)	(3.10)	(1.69)
β^{MKT}	2.696	2.918	3.937	8.679	12.674	12.762
	(1.65)	(1.31)	(1.42)	(2.94)	(3.05)	(2.25)
β^{SMB}	-0.9414	-2.4700	-3.3874	-2.3614	1.7958	4.1322
	(-0.90)	(-1.67)	(-1.54)	(-1.14)	(0.76)	(1.57)
β^{HML}	-0.5315	-3.1330	-5.6746	-5.2804	-1.3567	1.7546
	(-0.57)	(-2.03)	(-2.12)	(-2.06)	(-0.63)	(0.55)
β^{MOM}	0.4958	0.6658	-1.2707	-5.3394	-8.8983	-5.1913
	(0.30)	(0.35)	(-0.40)	(-1.79)	(-2.45)	(-1.37)
β^{VRP}	0.0027	0.0291	0.0764	0.1340	0.1352	0.1288
	(0.07)	(0.86)	(1.65)	(1.62)	(1.89)	(2.47)
β^{SRP}	0.5044	1.3046	1.8460	0.9361	-0.2077	-0.5017
	(0.88)	(1.95)	(2.56)	(1.24)	(-0.22)	(-0.61)
β^{BAS}	-0.0002	-0.0002	0.0003	0.0007	0.0003	0.0000
	(-0.65)	(-0.40)	(0.65)	(1.41)	(0.59)	(0.03)
β^{OPI}	0.0167	0.0232	-0.0238	-0.0810	-0.1062	-0.1227
	(0.97)	(0.89)	(-0.80)	(-2.47)	(-3.20)	(-2.55)

Notes to Table: We report the persistence of ER beta onto lagged cross sectional predictors for 6 to 60 months. We report in parentheses the Newey-West corrected t-statistics.

Table 4: Cross Sectional Firm Level Regressions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Intercept	0.018	0.014	0.014	0.014	0.014	0.014	0.014	0.013	0.011
	(3.31)	(2.90)	(3.22)	(2.93)	(3.13)	(3.23)	(3.37)	(3.10)	(2.71)
β^{ER}	-0.510	-0.512	-0.478	-0.486	-0.671	-0.565	-0.614	-0.578	-0.624
	(-2.01)	(-2.02)	(-2.01)	(-2.16)	(-2.81)	(-2.46)	(-2.42)	(-2.44)	(-2.56)
β^{MKT}		0.484	0.151	0.225	0.015	-0.163	-0.096	-0.015	-0.008
		(1.87)	(0.51)	(0.64)	(0.04)	(-0.45)	(-0.27)	(-0.04)	(-0.02)
β^{SMB}			0.418	0.414	0.339	0.330	0.307	0.181	0.119
			(2.26)	(2.09)	(1.71)	(1.57)	(1.48)	(0.81)	(0.53)
β^{HML}				0.069	-0.246	-0.196	-0.159	0.007	0.030
				(0.33)	(-0.95)	(-0.78)	(-0.64)	(0.03)	(0.12)
β^{MOM}					-0.803	-0.514	-0.426	-0.327	-0.344
					(-2.15)	(-1.45)	(-1.16)	(-0.84)	(-0.83)
β^{VRP}						0.013	0.015	0.012	0.009
						(2.69)	(2.64)	(2.09)	(1.35)
β^{SRP}							0.040	0.019	-0.014
							(0.62)	(0.25)	(-0.17)
β^{BAS}								0.000	0.000
								(-0.81)	(-0.97)
β^{OPI}									0.009
									(2.37)
$R^2_{Adj.}$	1.94%	4.22%	5.34%	6.37%	7.24%	8.10%	9.13%	10.12%	11.00%

Notes to Table: We report the incremental Fama-MacBeth cross sectional regressions for individual stocks. The bottom row reports the adjusted $R-$squared. We report in parentheses the Newey-West corrected $t-$statistics.

Table 5: Double Sorts

	Panel A. β^{MKT}		Panel B. β^{SMB}		Panel C. β^{HML}		Panel D. β^{MOM}	
	High	Low	High	Low	High	Low	High	Low
High	0.0150	0.0196	0.0168	0.0149	0.0198	0.0144	0.0130	0.0190
	(2.60)	(4.00)	(2.93)	(3.22)	(3.68)	(2.68)	(2.50)	(3.48)
Low	0.0277	0.0210	0.0246	0.0242	0.0257	0.0218	0.0235	0.0265
	(3.90)	(3.61)	(3.50)	(4.26)	(3.84)	(3.53)	(3.85)	(3.85)
$ER_{Diff.}$	-0.0127	-0.0014	-0.0078	-0.0092	-0.0056	-0.0082	-0.0105	-0.0075
	(-3.10)	(-0.33)	(-2.10)	(-2.30)	(-1.27)	(-2.12)	(-2.89)	(-1.75)
α_4	-0.0121	-0.0009	-0.0076	-0.0084	-0.0082	-0.0092	-0.0118	-0.0062
	(-2.79)	(-0.20)	(-1.95)	(-2.11)	(-1.96)	(-2.15)	(-3.16)	(-1.28)

	Panel E. β^{VRP}		Panel F. β^{SRP}		Panel G. β^{BAS}		Panel H. β^{OPI}	
	High	Low	High	Low	High	Low	High	Low
High	0.0165	0.0169	0.0193	0.0156	0.0190	0.0172	0.0179	0.0158
	(3.03)	(3.23)	(3.44)	(2.95)	(3.81)	(2.45)	(3.34)	(2.90)
Low	0.0248	0.0221	0.0263	0.0211	0.0225	0.0235	0.0281	0.0199
	(3.61)	(3.64)	(3.91)	(3.29)	(3.77)	(3.42)	(4.04)	(3.22)
$ER_{Diff.}$	-0.0086	-0.0049	-0.0091	-0.0064	-0.0035	-0.0059	-0.0102	-0.0041
	(-2.14)	(-1.23)	(-2.14)	(-1.46)	(-0.82)	(-1.36)	(-2.42)	(-1.04)
α_4	-0.0108	-0.0061	-0.0109	-0.0083	-0.0023	-0.0042	-0.0096	-0.0040
	(-3.05)	(-1.61)	(-2.54)	(-2.37)	(-0.38)	(-1.11)	(-2.14)	(-0.97)

Notes to Table: We report the independent double sorts of ER beta and alternative variables. Panels A to D report the double sorts for ER beta along with stock sensitivities to MKT, SMB, HML, and MOM factors. Panels E to H report the double sorts for ER beta along with stock sensitivities to VRP, SRP, BAS, and OPI characteristics. We report the intercept α_4 from the difference group onto the Fama and French (1992) and Carhart (1997) factors. Newey-West corrected t-statistics in parentheses.

Table 6: Energy Risks and Equity, Commodity, and Global Risk Factors

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Intercept	-0.0141	-0.0158	-0.0148	-0.0157	-0.0125	-0.0139	-0.0152	-0.0153	-0.0153	-0.0153	-0.0154	-0.0103
	(-2.84)	(-3.13)	(-2.93)	(-3.14)	(-2.21)	(-2.30)	(-2.91)	(-2.98)	(-3.10)	(-3.05)	(-3.07)	(-1.68)
b_{MKT}	-0.0021											-0.0026
	(-1.88)											(-1.24)
b_{SMB}		0.0020										0.0030
		(1.57)										(2.12)
b_{HML}			-0.0018									-0.0022
			(-1.12)									(-1.32)
b_{MOM}				0.0009								0.0001
				(0.79)								(0.05)
b_{VRP}					0.0510							0.0723
					(1.65)							(2.11)
b_{SRP}						-0.0026						0.0010
						(-0.77)						(0.24)
b_{BAS}							1.9806					7.2142
							(0.21)					(0.84)
b_{OPI}								-0.0032				0.0221
								(-0.04)				(0.28)
b_{VIX}									0.1330			-0.0603
									(1.62)			(-0.34)
b_{TED}										-0.0798		-0.7681
										(-0.04)		(-0.41)
b_{EPU}											0.0003	0.0003
											(1.95)	(1.88)
$R^2_{Adj.}$	1.25%	0.32%	0.21%	0.01%	-0.11%	-0.35%	-0.44%	-0.45%	0.21%	-0.45%	1.64%	1.73%

Notes to Table: We report the time series regressions of the long short energy risk strategy onto alternative covariates. Columns 1 to 11 report the univariate regressions. Column 12 reports the multivariate regression. The bottom row reports the adjusted R^2. We report in parentheses the Newey-West corrected t-statistics

Table 7: Energy Risks and Media Coverage

Panel A. Crude Oil Media Coverage		
	Crude Oil Spikes	Crude Oil Uncertainty
Avg.	148	240
Std. Dev.	162	248
Min.	8	8
Max.	1,345	2,062
Total	33,215	53,794

Panel B. ER Beta Strategy		
	Coverage Above Trend	
	Crude Oil Spikes	Crude Oil Uncertainty
Return	-0.0157	-0.0208
t-Stat.	-2.04	-2.85
Sharpe	-0.22	-0.29
	Coverage Below Trend	
	Crude Oil Spikes	Crude Oil Uncertainty
Return	-0.0150	-0.0112
t-Stat.	-2.33	-1.68
Sharpe	-0.20	-0.15

Notes to Table: Panel A reports the summary statistics for the number of news referring to crude oil spikes and crude oil uncertainty. Panel B reports the results of the ER beta long-short strategy in times of above average and below average media coverage on crude oil spikes and crude oil uncertainty. The sample period is from September 1995 to April 2014.

Table A.1: Top Oil Consumers and Net Exporters

Panel A. Top World Oil Consumers				Panel B. Top World Oil Net Exporters		
	(Thousand Barrels per Day)				(Thousand Barrels per Day)	
Country	2012	2013		Country	2012	2013
United States	18,490	18,961		Saudi Arabia	8,865	8,733
China	9,875	10,303		Russia	7,201	7,249
Japan	4,726	4,531		United Arab Emirates	2,544	2,743
India	3,450	3,509		Kuwait	2,347	2,345
Russia	3,195	3,515		Iraq	2,247	2,289
Brazil	2,997	2,998		Nigeria	2,224	2,070
Saudi Arabia	2,861	2,968		Qatar	1,829	1,847
Germany	2,388	2,403		Iran	1,728	1,322
Korea	2,301	2,324		Angola	1,713	1,756
Canada	2,281	2,431		Venezuela	1,712	1,905

Notes to Table: We report the top ten oil consumers and net exporters from the Energy Information Administration (EIA). Panel A reports the top ten world oil consumers for the years 2012 in the second column and 2013 in the third column (latest available year). Panel B reports the top ten world oil net exporters for the years 2012 in the second column and 2013 in the third column. The units are in thousand barrels per day for both panels.

Table A.2: Economic Moments

	σ_Y	σ_C	σ_I	$\rho_{Y,C}$	$\rho_{Y,I}$
Model	0.024	0.024	0.074	0.748	0.656
Data	0.023	0.022	0.081	0.748	0.863

Notes to Table: We report unconditional aggregate economic moments for the model and the data. The second row reports results from the model. The third row reports results from the data. The second column reports the standard deviation of income, the third column reports the standard deviation of consumption, the fourth column reports the standard deviation of investments, the fifth column reports the correlation between income and consumption, and the sixth column reports the correlation between income and investments. The sample period is from October 1990 to May 2014. All values are annualized.

Corporate ESG Profiles and Banking Relationships*

Joel F. Houston† and Hongyu Shan‡

First Version: February, 2019
This Version: May, 2019

Abstract

In recent years, banks have faced increased internal and external pressure to be more accountable to their customers and to make more socially-responsible lending decisions. Recognizing these pressures, this paper examines if and when banking relationships act as a transmission mechanism for promoting corporate Environmental, Social and Governance (ESG) policies. We show that banks are more likely to grant loans to borrowers with similar ESG profiles, and positively influence the evolution of borrowers' ESG performance over time. Consistent with the theory of creditor control, this influence is more pronounced 1) when banks have relatively better ESG ratings than their borrowers, and 2) when borrowers are bank-dependent. As a disciplinary mechanism, we show that borrowers are more likely to experience costly disruptions in existing lending relationships following negative news coverage on their ESG-related issues. We exploit M&A in the banking industry as a quasi-exogenous shock to the lender's ESG standard to establish causality. Overall, our results suggest that banks have become an effective conduit promoting socially responsible activities among borrowers.

JEL classification: G21, G28, G38, Q50
Keywords: ESG, banking, responsible lending, creditor control

*We thank Iftekhar Hasan, Chris James, Jay Ritter, Guner Velioglu, and Xinyan Yan for comments and suggestions. All remaining errors are our own.

†Corresponding author. Eugene Brigham Chair in Finance, University of Florida, Warrington College of Business, Gainesville, FL 32611, E-mail: joel.houston@warrington.ufl.edu

‡Assistant Professor in Finance, Fordham University, Gabelli School of Business, New York, NY 10023

1. Introduction

Beyond meeting their financial objectives, firms often strive to integrate a wide variety of Environmental, Social and Governance (ESG) goals into their business models (Bénabou and Tirole (2010), Hart and Zingales (2017)). Coincident with these efforts, they have faced growing internal and external pressures to improve performance along various non-financial dimensions including environmental impacts, social welfare, and fair labor practices. While these forces apply to a wide range of firms, banks in recent years have particularly faced increased pressure to be more accountable to their customers and to make more socially-responsible lending decisions. For example, in April 2019, a group promoting gun control released a well-publicized report card ranking banks on their ties to firearm manufacturers and organizations such as the National Rifle Association (NRA).[1] In another recent example, the head of Wells Fargo stepped down shortly after receiving intense criticism during Congressional hearings. This criticism stemmed primarily from the bank's well-documented fake accounts scandal, but Congressional leaders also heavily criticized the bank for making loans to firms operating private prisons and energy pipelines.[2] Despite the heightened interests in the social economic impact of bank lending practise, there is far from a consensus on whether banks could effectively discipline and shape borrowers' ESG activities.[3] Our study presents the first empirical evidence on the interplay between responsible lending and borrower ESG behavior, and informs policy makers regarding the efficacy and real ESG outcomes of bank lending regulation.

In this paper, we propose a novel economic mechanism to explain the propagation of ESG policies through lending relationships. We broadly examine the evolution of borrower ESG profiles for both public and private firms. While there is considerable evidence that bankers may affect their borrowers' policies and investments (Shleifer and Vishny (1997), Chava and Roberts (2008),

[1]See details of the news coverage at www.nytimes.com/2019/04/04/business/gun-control-banks.html

[2]See details of the news coverage at www.nytimes.com/2019/03/28/business/wells-fargo-timothy-sloan.html

[3]The literature on the causes and effects of firm ESG policies is long-standing. Earlier studies have largely focused on the determinants motivating cross-sectional differences in observed levels of ESG ratings, as well as the wealth effects of these policies, with a particular emphasis on the positive impact provided by institutional investors (Starks et al. (2017), Cao et al. (2019), He et al. (2018), Dimson et al. (2015), among others). Given this focus, most studies have concentrated on public firms. Nevertheless, data from RepRisk, a Zurich-based data science company that scans negative ESG news incidents, shows that the number of private firms involved in ESG incidents was six times higher than that of public firms between 2007 and 2018. In fact, the majority of firms that pose ESG risks to society are small, private firms that receive minimal level of public scrutiny from the equity market. In light of these facts, the roles played by critical stakeholders in shaping corporate ESG practice remain under-explored.

Nini et al. (2012)), it remains an open question whether lenders use this leverage to specifically influence borrower ESG policies. One view is that lenders focus primarily on their borrowers' financial performance, and consequently resist costly investments that chiefly benefit other non-bank stakeholders. Friedman (1970) argues that the only responsibility of a firm is to increase its profit. Brammer and Millington (2008) presents evidence consistent with the view that high social responsibility performers score the lowest in short term financial performance.

However, beyond this narrow view, lenders also have many compelling reasons to encourage their borrowers to be socially responsible. There are two mutually non-exclusive channels. The first channel emphasizes the financial exposures of the bank's relationship-specific investments in the existing loans. One obvious reason is that borrowers with high ESG-related risks are more likely to face financial consequences from these policies, which ultimately increase default risk and could theoretically subject the bank to lender liability risk. Consistent with this *financial channel*, some recent research has shown that promoting engagements in ESG issues can reduce firms' downside risk (Hoepner et al. (2018)), and has documented an association between measures of ESG ratings and loan pricing (Chava (2014), Goss and Roberts (2011), Hasan et al. (2017) ,Hauptmann (2017)).

In addition to these direct effects, there is a *reputation channel* which links the borrower ESG issues to the bank's ability to engage future business. In this channel, lending to low ESG borrowers increases the bank's costs dealing with regulatory scrutiny, tarnishes its reputation, and ultimately decreases its opportunity to engage future business. Given that banks are heavily regulated and are often the focus of public condemnation (as highlighted above), they have a strong incentive to reduce negative reputational incidents (both their own and that of their borrowers). For example, after the high school mass shooting in Parkland, Florida that claimed 17 deaths and 17 injures, Bank of America announced that it would stop lending money to gun manufacturers that choose to continue the production of military-inspired firearms for civilian use. Note that the bank's decision is unlikely based on considerations of the default and lender liability risks, given the lucrative nature and liquidity of its clients.[4] The event is hardly an isolated occurrence. Citibank and Wells Fargo have withdrawn their lending to oil firms that are involved in mountaintop removal, a legal and lucrative method of extracting oil but very harmful to environment. This anecdotal evidence collectively demonstrates the importance of expanding the focus beyond default and liability risks

[4]See New York Times, https://www.nytimes.com/2018/04/10/business/bank-of-america-guns.html

when attempting to understand the drivers of banks' lending decisions.[5]

In our analysis, we conduct three broad sets of tests that are designed to explore the extent to which lenders serve as important conduits influencing firm ESG policies. First, we consider whether lenders select borrowers based on their observed ESG profiles, and relatedly whether banks' own ESG standards affect the type of borrowers it works with. Second, we examine whether lenders systematically influence borrower ESG performance over time. Third, we investigate whether borrowers experience disruptions in existing lending relationships following a significant negative shock to their reputation. In each of our tests, we use the RepRisk database to obtain the negative news coverage and ESG ratings of both borrowers and lenders. The database is best suited for our study because its coverage on a wide range of both public and *private* borrowers, and its *outcome-driven* approach. The coverage on private firms is critical when we explore the corporate loan market, where the majority of borrowers are private firms which receive minimal level of scrutiny from the equity market. Also, RepRisk focuses on actual reported ESG related events, which incurred real costs. By contrast, other databases primarily assign ESG ratings based on whether the firm "claim" to enact certain policies that are more discretionary and subject to green-washing bias.[6] Notably, the rating system incorporates not just the number of incidents, but also the severity, reach and novelty of the events to evaluate the firm's reputation exposure to ESG and business conduct risks.

In our first set of results, we show that lenders tend to match with borrowers that have similar ESG profiles. Specifically, we find that a standard deviation decrease in the observed difference between the lender and borrower's RepRisk ratings translates into a 0.23% increase in the likelihood that the two will partner together.[7] This increase is both statistically and economically significant given that the unconditional likelihood of matching between a potential lender and borrower in our sample is only 3%. This finding is robust to the inclusion of a host of control variables including borrower's industry, lender, and year fixed effects, and measures of firm size (Stein (2002), Hubbard

[5]In one of our other projects on the same line of research, we examine how banks benefit from lower reputation risk exposures. We show that the level of bank reputation risk exposure is positively related to the risk-adjusted capital ratios, and cost of capital. See Appendix D and Houston, Shan and Tian (2019) for details.

[6]An increasing number of studies in ESG focus on the real outcomes, instead of discretionary disclosures which are often subject to Green-washing bias. For example, legal and litigation risks (Schiller (2018), and toxic and/or carbon emissions (Bartram et al. (2018), Shive and Forster (2019), Kim and Xu (2017), among others). In the same spirit, RepRisk focuses on real outcomes (externally reported ESG related incidents) which incorporate a broad range of ESG accidents that span across 28 issues. See WRDS RepRisk data manual for details.

[7]We construct both the adjusted and unadjusted difference between the borrower and lender's ESG ratings. Adjustments are made by subtracting the country-sector-month average from the raw ratings to account for comparability across industry, country and time. Our results are robust using both methods.

et al. (2002)), investment opportunities (Ongena and Smith (2001), Gopalan et al. (2011)), bank dependency (Sharpe (1990), Schwert (2018)), prior lending relationships (Bharath et al. (2007)), and debt overhang. Further, our sub-sample analysis on single-lead loans produce nearly identical results to those found in the broader sample. Altogether, our results show that lenders tend to associate with borrowers that share their attitudes regarding ESG-related policies, and/or have similar observed reputations related to ESG issues.

Next, we provide evidence that lenders also influence the evolution of their borrower's ESG profiles. Specifically, borrowers that have lenders with relatively better ESG standard are more likely to realize improved ESG ratings over time. Here we find that a one standard deviation increase in the difference between the borrower and lender(s)' ESG ratings is associated with a 0.70 increase in the borrower's RepRisk rating over a two year window centered on the loan initiation date, which is equivalent to 37% of the standard deviation of the change in ESG for all firms during the same two year window. These results confirm that banks can impact the borrowers' ESG performance in a significant and dynamic manner.

While the demonstrated associations appear to be economically significant and robust to a variety of specifications, establishing direct causation is notoriously challenging. The biggest identification concern relates to disentangling treatment from selection effects. While we believe that banks have a positive impact on the evolution of borrower's ESG performance (*treatment*), a reasonable alternative explanation is that borrowers who expect to improve their ESG standard choose to borrow money from ESG focused banks (*selection*). To alleviate concerns on the potential selection problem and other omitted variable bias, we follow the literature to exploit M&A in the banking industry as a quasi-exogenous shock to the lender's ESG standard (Asker and Ljungqvist (2010), Hong and Kacperczyk (2010), Chen et al. (2015)).[8] In a Diff-in-Diff setting, we examine if the exogenous variation transmits through the *established* lending relationship to affect the evolution of borrower's ESG ratings post the M&A. We apply a wide range of fixed effects on the lender, borrower, industry, and year levels to absorb the remaining unobservable time-invariant heterogeneities across lenders, borrowers and industries, and to preclude the effects of common time trends. In short, we set a high bar to refute conclusion that banks can effectively discipline

[8]We believe that the timing and the decision of bank M&A activities are arguably exogenous to the borrowers' firm-level unobservable characteristics that determine ESG ratings. As noted by prior studies, the bank merger waves were largely driven by regulatory, technological, and competitive changes (Pilloff (2004)).

borrowers' ESG activities.

We also construct a variety of conditional tests to further explore the specific channels in which banks may influence the borrower's ESG evolution. First, we find that lenders have a more profound influence if the borrower is bank dependent (as proxied by whether the firm is unrated, and the intensity of covenants). We also show that bank's influence critically hinges on secured loans where the transfer of control rights and liability exposures is triggered under a significant adverse shock. Second, we find that there is an important asymmetry – banks that have better ESG-related ratings relative to the borrower are more likely to induce borrowers to improve their ESG levels over time. On the other hand, lender's impact on borrower's ESG evolution is indistinguishable from zero if the lender's ESG rating is worse relative to that of the borrower. Finally, we disaggregate the overall ESG rating into its three sub-components (environmental (E), social (S) and governance (G)). Similar to findings in Dimson et al. (2015), we find that banks are more likely to induce their borrowers to improve along the environmental (E) and social (S) dimension. Their influence on their borrower's governance (G) appears to be negligible.

In a third set of results, we document a novel disciplinary mechanism. In the process of lending to a firm, a bank acquires proprietary firm-specific information that is unavailable to non-lenders (Schenone (2009)). Switching lenders is costly for borrowers and is often accompanied by reduction in the availability of credit (Petersen and Rajan (1994)). If the borrower continues to engage in risky ESG practices and are exposed to a greater number of negative news incidents, does it lead to interruptions in the existing lending relationship? We find that borrowers are significantly more likely to observe a shift in lead lender(s) following a negative shock to their ESG-related reputation. More specifically, conditional on obtaining new loan financing within the 12 months period after the end date of the original loan, we find that borrowers are 3-4% less likely to initiate a new loan with the same lead lender(s) if there was a negative ESG-related news incident. Furthermore, we find that these borrowers exposed to negative ESG related news are more likely to shift to lenders with worse RepRisk ratings. We control for both the level, and the change in the borrower's financials including ROA, assets, leverage, and Tobin's Q to make sure that the switch in lending relationship is not driven by fundamental changes in credit and liability risk. To alleviate concerns of omitted variable bias, we utilize *negative* news coverage initiated by *outsiders*, whose timing relative to the loan expiration date is arguably quasi-exogenous and out of the control of corporate insiders.

On balance, our findings clearly demonstrate that the banking system has an important systematic effect on corporate ESG policies. In this regard, we believe our findings make an important contribution to the growing literature on the role of key stakeholders in shaping corporate ESG policies (Shive and Forster (2019), Lins et al. (2017), Starks et al. (2017), Chava (2014), Dimson et al. (2015), Bartram et al. (2018)). Most notably, recent papers by Schiller (2018) and Dai et al. (2018) document that socially conscious customers have taken steps to induce their key suppliers to become more socially responsible. Given the importance of a sound evaluation of efficacy and real effects of responsible lending, it is surprising how little empirical work has been done on this front. Our work aims to fill this gap.

At the same time, our paper also contributes to the vast literature on banking relationships, by highlighting another important factor that influences the choice of lender and the role that lenders play in influencing firm performance and investment decisions (Shleifer and Vishny (1997), Chava and Roberts (2008), Nini et al. (2012), Schwert (2018), among others). In this vein, our work is related to the long standing theories of relationship lending (Sharpe (1990), Berger and Udell (1995), among others) and bank monitoring (Holmstrom and Tirole (1997), Diamond (1991), among others).

The rest of the paper is organized as follows. Section 2 summarizes the data employed in the various tests. Section 3 presents our main results. Section 4 describes sources of endogeneity and our identification strategy. Section 5 offers several robustness tests. Section 6 concludes.

2. Data

2.1. ESG Data

This study employs an event-based outcome measure of firm-level environmental, social, and governance (ESG) profile for both public and private firms using data from RepRisk. The RepRisk data provides a monthly unbroken time-series ESG rating, and coverage on negative ESG news incidents from January 2007 to June 2017.[9] A dedicated team of analysts leverage a combination

[9]Positive ESG events are *not* covered by RepRisk and reported less often in traditional or social media. Part of the reasons can be attributed to the fact that positive news are more likely to be self-reported for branding and marketing purposes, and are subject to greenwashing biases. To the best of our knowledge, we are not aware of the existence of any positive ESG news database. See Li and Wu (2017) for extended discussions on the collection of positive news.

of artificial intelligence and curated human analysis to track a universe of over 95,000 firms globally, among which 82,000 are private firms with no self-reported ESG compliance information. On a daily basis, over 80,000 public sources and stakeholders in 20 languages are screened. Once an incident is identified, analysts conduct additional analysis to (1) confirm that the incident is indeed ESG-related, (2) remove possible duplicate media coverage on the same incident to make sure each risk event only enters once into the RepRisk Platform, and (3) identify the specific nature of the incident, by mapping it to 28 ESG Issues and 45 ESG topics. Each incident is assigned three proprietary scores based on severity (harshness), reach (influence), and novelty (newness). Finally, the RepRisk Index (RRI hereafter) is updated, reflecting the impact of the news incident.

Compared with the widely used annual KLD database (now MSCI ESGSTATS), the RepRisk data is uniquely suited for our study for three reasons. First, the event-based data evaluates the outcome of ESG activities, which can be directly linked to the societal impact of ESG compliance. The KLD data instead relies on the firm's self-reported information which varies largely with the firm's discretionary disclosures related to ESG compliance. RepRisk arguably provides a more objective assessment of the societal impact of each firm over time, because it is more difficult for firms to endogenously manipulate media attention/negative news detection, than to manipulate self-disclosed policy adoptions. Second, the KLD data does not cover private firms, which are predominant in the corporate loan market. Third, RepRisk has unparalleled granularity. It employs a monthly, continuous ESG rating ranging from 0 to 100, while most of the KLD ratings are structured as an annual, indicator variable that equals 0 or 1.

2.2. Banking Data

We obtain loan pricing and contract information from Loan Pricing Corporation's (LPC) Dealscan database, for the sample period from 2007 to 2017. We focus on the loans granted to U.S.-incorporated firms. Dealscan provides characteristics information for each loan such as size, maturity, type, and purpose, as well as information about the outstanding financial covenants and other terms. We hand-match the Dealscan loan data to RepRisk ESG ratings using company names. We use S&P Capital IQ as well as Google search to track the historical names of each company to verify the accuracy of matches.

We study the evolution of borrower ESG ratings over time at the loan level. Specifically, we consider each loan as a relationship between a borrower and a lead lender that finances the loan. We follow the approach that Bharath et al. (2009) used to classify lead lenders for each loan. We classify a lender as lead lender if the "LeadArrangerCredit" field indicates "Yes" or if the "LenderRole" field indicates one of the following: administrative agent, agent, arranger, lead arranger, lead bank. For some loans in our sample we have multiple lead lenders in the syndicate. In these cases, we calculate the equally-weighted average of ESG ratings of lenders in the syndicate.[10]

2.3. M&A Data

From the SDC M&A database, we extract the set of completed merger and acquisitions in the financial industry from 2007 through 2017. The following filters are applied: 1) both the acquirer and the target have SIC codes between 6000 and 6999, 2) the acquirer owned less than 50% of the target bank's shares six months before the transaction and more than 50% of the shares after the transaction, and 3) we exclude deals with missing transaction values.

We then subsequently match the acquirer and the target's names to the lender's names in the Dealscan database. We view the merger as an exogenous shock to the ESG-related standards of the involved lender(s). This setup enables us to determine whether borrower ESG performance evolve differently if their lender(s) undergo an exogenous shift in their ESG standard. The magnitude of the shock depends on the relative size of the acquirer and target (see detailed discussion in section 2). Our final sample consists of 423 treated loans initiated from 2007 to 2017 where the borrower, lender, and acquirer (or target, if the lender is the acquirer) have non-missing RepRisk ESG ratings.

2.4. Financials

After constructing the sample of loans with corresponding deal characteristics as well as borrower and lender ESG ratings, we also incorporate a broad range of firm-level control variables from Compustat. Specifically, we collect the reference firms' financial information from Compustat, for the most recent fiscal year ending within one-year window prior to the loan start date (i.e., lagged). We use the Chava and Roberts (2008) linking file to link loans from Dealscan to firms in Compustat. We then supplement the firm controls with S&P credit ratings.

[10] Alternatively, as a robustness test, we select a unique lead lender for each loan following Ivashina and Kovner (2011). This procedure selects the lead lender that the firm has the strongest relationship by considering the past borrowing history. We present our findings under this alternative approach in the robustness section 5.2.

An important dimension of our study is that it includes both public and private firms. We classify a firm as a public if we we can find a stock price available from the Center for Research in Security Prices (CRSP) for the same fiscal year, and as a private firm otherwise. The list and detailed definitions of required firm- and loan-level variables are provided in Appendix A.

2.5. Summary Statistics

Table 1 presents the summary statistics for our sample of loans and the corresponding borrowers. In our sample, we have 12,495 loans, taken out by 2,407 borrowers and granted by 116 lenders from 2007 to 2017. The median borrower has an ESG rating of zero, which suggests that median firm has no publicly known issues (the lower the ESG rating, the better). The median lender on the other hand has an ESG rating of 17, which indicates that it has some known issues. These differences could be explained by two possible factors: (i) the median bank in our sample is larger than the median borrower, and larger firms are more likely to receive publicity; (ii) financial industry firms often receive more attention and greater scrutiny, especially during our sample period which corresponds to the financial crisis and post crisis periods. Overall, our interest is the relative standing of each borrower and lender within its own industry, as well as the difference in their ESG ratings.

[Insert Table 1 here]

To account for the size as well as credit risk of the borrower we include firm level controls. About 61% of the loans are granted to rated borrowers, and 29% of all loans are granted to investment grade firms. Similarly, we find that 62% of the loans are granted to public firms. These statistics suggest that a significant portion of our sample has limited access to public debt and equity markets.

An important dimension of our analysis is to test the conditions when the lender has stronger influence over the borrowers. Therefore, although we do not have the corresponding accounting information for about 40% of the loans in our sample, we still include these loans in our baseline tests to determine the importance of creditor control in shaping the ESG policies of bank-dependent borrowers. In addition to the existence of credit ratings and outside borrowing options, we use other proxies for the strength of lender control. In particular, we use the number of covenants as a proxy for the strength of lender control (Nini et al. (2012)).

[Insert Figure 1 here]

One of the empirical challenges in ESG studies is the comparability of scores and ratings across industries and years. In Figure 1, we calculate the mean level of RRI of all borrowers in our sample. Figure 1A documents a rising level of RRI over time, partly driven by an increasing number of ESG related news coverage in recent years. Figure 1B shows that the level of ESG exposures vary by industry. Borrowers in Utilities, Energy, and Chemicals on average have a higher level of RRI. We address this issue by subtracting the industry-country-month average RRI from the borrower's raw RRI to obtain the monthly adjusted RRI, which we use as the independent variable. We also include the time, and year fixed effects to mitigate similar concerns on the dependent side.[11]

3. Main Results

3.1. Matching

This section explores whether lenders are more likely to grant loans to borrowers with similar ESG profiles. Following Houston, Lee and Suntheim (2018) and Cai et al. (2012), we run a linear probability model with the adjusted and unadjusted pairwise similarity measure, $Close_{i,j,t-1}$ as the main right-hand-side (RHS) variable.

In constructing the possible lender-borrower pairs, we consider all the unique banks that act at least once as lead lender in year t, and all the unique borrowers that borrow at least once in year t. The match dummy for each possible lending relationship is constructed on the lender-borrower-year level. Borrowers and lenders that never participate in the corporate loan market in year t are not considered potential candidates for the pairing test. This reduces the heterogeneity in the demand and supply side of the loan market by focusing only on the borrowers (lenders) actively seeking (providing) loan financing in year t. The empirical analysis is based on the following Probit specification:

$$Pr(Match_{i,j,t} = 1) = \phi(\alpha + \beta Close_{i,j,t-1} + \gamma X_{i,t-1} + S_j + I_{ffindustry} + \delta_t + \xi_{i,j,t}) \tag{1}$$

In the Probit Model, $\phi(.)$ denotes the cumulative distribution function (CDF) of the standard normal distribution. $Match_{i,j,t}$ is a dummy variable that equals one if the lender j extends a loan

[11] In the robustness test section, we show that our baseline results are *not* quantitatively changed if we use the unadjusted RRI as independent variables.

to the borrower i in year t. $Close_{i,j,t-1}$ is the main explanatory variable that measures the lagged distance between the lender's, i, and the borrower's, j, yearly average RepRisk ratings in year t-1. $X_{i,t-1}$ is the vector of borrower's characteristics that we use as control variables. These variables include the prior lending relationship dummy (prior), size (log assets), book leverage, Tobin's Q, and the investment grade dummy. Note that some banks may lend more than other banks in the syndicated loan market. Likewise, borrowers from certain industries may be more favored by lenders. To address these issues, S_j, $I_{FFindustry}$ and δ_t respectively denote the dummies for lender, the Fama-French 12 industry categories, and year fixed effects. Finally, standard errors are clustered at the lender-borrower pair level.

[Insert Table 2 here]

The results are presented in Table 2 Panel A. In columns 1, 2 and 3, we use the unadjusted ESG difference measure as the independent variable; while in columns 4, 5 and 6, we adjust both the lender and the borrower's RRI with their country-industry-month means to account for heterogeneity across industries and time. Columns 1 and 4 use the whole sample, while columns 2 and 5 focus only on the sub sample of public borrowers whose financials are available. In Columns 3 and 6, we further exclude all borrowers whose RRI are clustered at zero. Some of the borrowers are never exposed to negative ESG-related news, and therefore maintain a consistently zero RRI. Arguably these firms may have other unobservable firm and industry level characteristics that explain their consistently zero ratings.

The key coefficients of interest in all columns are both statistically and economically significant. Take columns 2 and 4 for example, where the coefficient estimates of the ESG difference measure are significant at the 0.1% level. Using marginal effects estimated at the sample means, we show that a standard deviation decrease in the ESG difference is associated with a 0.34% (0.000188 × 18.05) increase in the likelihood of matching. A standard deviation decrease in the adjusted ESG difference is associated with a 0.23% (0.000144 × 16.21) increase in the likelihood of matching. The economic magnitudes are sizable given the unconditional likelihood of matching between a potential borrower and lender is only 3%.

In Panel B, we repeat the analysis in Panel A using the sample of single-lead loans only. This reduces the stacking of multiple matching pairs between the same borrower but different co-lead lenders, which inevitably reduces the standard errors of the coefficient estimates and inflate the Z statistics. We show our results are robust after removing multiple-lead loans.

3.2. *Evolution of Borrower's ESG Performance*

This section explores how corporate ESG policies propagate through lending relationships. We examine the direct impact of banks on the evolution of the borrowers' ESG performance using facility-level data. The empirical analysis is based on the following OLS specification:

$$\begin{aligned} ESG\_Chg_i =& \alpha + \beta ESG\_Diff_{i,j,t-1} + \lambda Lender\_Chg_j + \theta ESG\_Borrower_{i,t-1} \\ &+ \gamma X_{i,t-1} + I_{ffindustry} + \delta_t + \xi_{i,j,t} \end{aligned} \tag{2}$$

For each facility, the change in the borrower's ESG profile (ESG_Chg) is defined as the difference between the borrower's RRI over a two-year window, from one year before (t-1) to one year after the loan initiation date (t+1). The ex-ante difference between the lender and borrower's ESG ratings (ESG_Diff) is defined as the difference between the lender and borrower's RepRisk ESG rating measured one year before the loan initiation date. To alleviate potential concerns about the comparability of ESG scores across industries and years, both the lender and the borrower's RRI have been adjusted by the country-industry-month mean. Lender_Chg controls for the evolution in the lender's ESG rating over the same two-year window.

We realize that the evolution of the borrower's ESG rating is path-dependent. Borrowers with ex-ante *poor* ESG rating are more likely to improve over time, than borrowers with ex-ante *pristine* ESG rating. The control variable, ESG_Borrower, alleviates the concerns for the potential path-dependency problem and is defined as the borrower's RepRisk ESG rating one year before the loan initiation date. Other control variables include the log loan amount, country of syndication USA, the borrower's public status, and the number of covenants in the loan to control for the heterogeneity in size, regulatory environment, managerial myopia, and the credit risk, respectively. $I_{FFindustry}$ and δ_t denote the dummies for the Fama-French 12 industry and year fixed effects. We cluster the standard errors at the borrower level.

[Insert Table 3 here]

These results are presented in Table 3. In columns 1, 2 and 3, we run the regressions with only basic control variables related to the borrower and lender's ESG ratings; in column 4, we include the control variables that are available for both public and private borrowers including loan amount, country of syndication, public dummy, and number of covenants in the loan; in column 5, we further restrict our analysis to a sub-sample consisting of public firms with additional publicly available control variables including size (log assets), book leverage, return on assets (ROA), and Tobin's Q.

The key coefficient of interest, the difference between lender and borrower ESG ratings (i.e. ESG_Diff), is statistically significant at the 1% level in all five columns. The economic magnitude is also sizable. Take column 4 for example. A standard deviation increase in ESG_Diff is associated with a 0.70 (18.39 × 0.038) increase in the borrower's RRI over time, which is equivalent to 37% (0.70/1.89) of the standard deviation of ESG_chg in our sample.

3.3. Cross Sectional Variation in Bank Dependency

In Table 4, we focus on those cases where we expect the lender to have a particularly strong influence on its borrowers. We first consider whether bankers are more able to influence unrated borrowers. Unrated borrowers typically have less access to public financing, which arguably makes them more bank-dependent and more sensitive to holdup problems:

$$\begin{aligned} ESG\_Chg_i =& \alpha + \beta ESG\_Diff_{i,j,t-1} \times I_{dependency} + \varsigma ESG\_Diff_{i,j,t-1} + \tau I_{dependency} + \lambda Lender\_Chg_j \\ &+ \theta ESG\_Borrower_{i,t-1} + \gamma X_{i,t-1} + I_{ffindustry} + \delta_t + \xi_{i,j,t} \end{aligned} \tag{3}$$

where $ESG\_Diff_{i,j,t-1} \times I_{dependency}$ is an interaction term between the lender and borrower ESG difference and our proxies for bank dependency. These proxies include indicators for credit rating, investment grade status, number of covenants, and secured loan status.

In column 1 of Table 4, we find that lenders have greater influence on borrowers' ESG policies if the borrower is unrated. We repeat this test with investment grade versus non-investment grade firms in column 2 of Table 4. We similarly find that lenders have greater influence over the non-investment grade borrowers' ESG policies.

[Insert Table 4 here]

We next explore other proxies of bank dependency. Nini et al. (2012) find that there is a significant change in the management of covenant-violating firms, which suggests that lenders exert their influence on firm decision-making and governance. In addition, Strahan (1999) shows that loans to riskier borrowers – smaller borrowers, borrowers with less cash, and borrowers that are harder to value by outside investors – are more likely to be secured by collateral. We therefore use the number of financial covenants on the loans, and the indicator for secured loans as proxies for the strength of creditors' control over the borrower.

Columns 3 and 4 of Table 4 present our results. We find that firms with financial covenants and firms with secured loans are more likely to be influenced by the lender in their ESG policies. Overall, our results suggest that the influence of lender on the borrower is greater if the borrower is bank dependent.

3.4. Asymmetric Bank Influence

While our baseline results demonstrate that the gap between the lender and borrower's ESG ratings is significantly related to the evolution of a borrower's rating over time. A natural question arises whether the results are symmetric depending on whether the borrower has a higher or lower rating than its lender. One scenario explaining the observed results is that banks with ESG rating that are relatively stronger than that of the borrowing firm take implicit and explicit steps to force the borrower to improve their ratings. Another explanation is that when a bank has a relatively weaker ESG rating than its lender, its failure to nudge the borrowing firm creates an environment where the borrower may feel freer to take actions that ultimately weaken its ESG rating. If the effects are symmetric, both explanations may be equally relevant. Alternatively, the effects may be asymmetric, in which case the results are driven primarily by one of these two scenarios.

To empirically address this issue, we start by sorting the loans into groups where the lender has a better industry-year adjusted ESG rating than the borrower ("Better bank = 1"), and into groups where the lender has a worse ESG rating than the borrower ("Better bank = 0"). Table 5 Panel A presents the results. In columns 1 and 2, we regress the borrower ESG changes for the subsample of loans where the lenders have a better ESG rating. We find that the economic effect of the ESG difference is even greater when the lender has a better ESG rating. This suggests that lenders have a disciplining influence over the borrowers when they have relatively better ESG ratings.

[Insert Table 5 here]

In columns 3 and 4 of Panel A, we run the test for the subsample of loans where the lenders have worse ESG ratings than their borrowers. In these circumstances, we find no evidence of lenders influencing the evolution of their borrowers' ESG ratings. Overall, our findings suggest that while "good" lenders from an ESG perspective may encourage their borrowers to become more socially responsible, "bad" lenders do not induce their borrowers to become less responsible.

Another question is whether the magnitude of the ESG difference is relevant. We next investigate whether the distance between lenders' and borrowers' ESG ratings influences the evolution of borrowers' ESG ratings. To address this question, we partition the sample into quintiles with respect to the ESG difference variable (ESG_Diff). We then investigate the changes in borrowers' ESG for these subsamples. We present the first and fifth quintiles for brevity - smallest and greatest distances, respectively. Panel B of Table 5 shows our results. In columns 1 and 2, where we have the greatest distance between the lender's and borrower's ESG ratings, we observe the greatest economic effect. When the ESG difference is large, we find a significant improvement in borrowers' ESG ratings. With similar intuition, we consider the subsamples in columns 3 and 4, where the distance between the lender's and the borrower's ESG is the smallest. Consistent with the worse lender results, we find no evidence of lender's influence over the borrowers for these subsamples.

Overall, our findings suggest that the lenders' influence over borrowers' ESG ratings is asymmetric. In particular, the magnitude as well as the sign of the distance (ESG_Diff) are strong determinants of the evolution of borrowers' ESG.

3.5. Propagation of Bank Influence along E, S, and G

We believe that creditors may be particularly concerned with some sub-components of the borrower's ESG policies (Dimson et al. (2015)). For example, the governance (G) issues, which by definition include executive pay, internal audits, and other shareholder rights, primarily affect the interests of shareholders. On the other hand, lenders may focus on environmental (E) issues under the lender liability laws. According to Comprehensive Environmental Response, Compensation, and Liability Act (CERCLA), lenders could be held liable for borrower's environmental disasters. In addition to the litigation costs, the cleaning up costs of the pollution become senior to lender's claims. Indeed, Chava (2014) finds that firms with environmental concerns pay higher interest on their loans and have fewer banks participating in their loan syndicates.

With these concerns in mind, we take a closer look at the components of ESG ratings. In addition to the overall ESG rating, RepRisk reports a proportion that represents each of the three factor's contribution to the overall score. From these proportions, we construct specific ratings for environmental risk, social risk, as well as governance risk, by multiplying the overall ESG rating with the proportion of each factor. We then study the changes in each of these three factors for each loan, which are calculated as follows:

$$Chg\_E = RRI_{t+1} \times Environment\ Weight_{t+1} - RRI_{t-1} \times Environment\ Weight_{t-1} \tag{4}$$

$$Chg\_S = RRI_{t+1} \times Social\ Weight_{t+1} - RRI_{t-1} \times Social\ Weight_{t-1} \tag{5}$$

$$Chg\_G = RRI_{t+1} \times Governance\ Weight_{t+1} - RRI_{t-1} \times Governance\ Weight_{t-1} \tag{6}$$

Table 6 presents our results where we separately consider the three components. In columns 1, 2, and 3, we look at the changes in the environmental, social, and governance components of ESG ratings, respectively. We find that banks have the most significant impact on the environmental component of ESG rating, which is consistent with banks having stronger incentives to discourage firms from polluting activities. The second significantly affected factor is the social component of ESG ratings. Finally, we do not find any significant evidence of banks influencing their borrower's governance practices.

[Insert Table 6 here]

In columns 4, 5, and 6 of Table 6, we repeat this test for the sub sample of public firms, once again controlling for other firm characteristics. We similarly confirm that environmental and social factors play a significant role while governance factor does not change. Overall, our results in this section suggest that ESG ratings change more significantly when the lenders have strong incentives to influence the borrowers' ESG policies.

3.6. Negative Reputational News Events and Changes in Banking Relationship

So far, we have documented a positive impact from banks on the borrowers' ESG performance. However, if the borrowers continue to engage in negative ESG practice, does it lead to disruptions in the lending relationship? In this section, we answer this important second-stage question by examining the relationship between the borrower's negative reputational news coverage, and the likelihood of initiating new loan(s) with the same lead lender within 12 months of the original loan's end date.

$$Pr(Same_{i,j,te} = 1) = \phi(\alpha + \beta Num\ Rep\ Event_{i,ts,te} + \gamma X_{i,te-1} + S_{i,j} + I_{ffindustry} + \delta_t + \xi_{i,j,t}) \quad (7)$$

In the Probit Model, $\phi(.)$ denotes the cumulative distribution function (CDF) of the standard normal distribution. $same_{i,j,te}$ is a dummy variable that equals one if at least one of the lead lenders (j) in the original loan extends a new loan to the borrower i within 12 months of the original loan's end date, te. $Num\ Rep\ Event_{i,ts,te}$ is the main explanatory variable that measures the number of months with negative news coverage on the borrower i from the start (ts) to the end (te) dates of the original loan. $X_{i,te-1}$ is the vector of borrower's characteristics that we use as control variables. These variables include the ex-ante level, and change in book leverage, size (log assets), return on assets, and Tobin's Q. $S_{i,j}$ denote additional control variables that include the original loan length (in years), and the investment grade dummy. $I_{FFindustry}$ and δ_t respectively denote dummies for Fama-French 12 industry, and year fixed effects. Finally, standard errors are clustered at the borrower level.

[Insert Table 7 here]

Note that we restrict the regression sample to borrowers who received at least one loan financing within 12 months of the end date of the original loan. This mitigates concerns related to demand side heterogeneities, because we are only looking at borrowers actively seeking new loan financing.

Column 1 and 2 in Table 7 report the results. The coefficient estimates of *Num Rep Event* are statistically significant, and negatively related to the likelihood of retaining the same lead lender. It indicates that borrowers with greater negative news coverage are more likely to switch lead lender(s) after the end date of the original loan, controlling for the length of the original loan. In

Column 3 and 4, we define the dependent variable more restrictively. *Same res* (restrictive) is the dummy variable that turns on if the borrower initiates new loan(s) with exactly the same group of lead lenders within 12 months of the original loan end date. Our main results remain statistically and economically robust to the variation. Finally, in Column 5 and 6, we define *same sgl* (single lead lender) most restrictively, as the dummy variable that turns on if the original loan has a single lead lender, and the borrower initiates new loan(s) with the same lender within 12 months of the original loan end date. The coefficient estimate is not statistically significant, partly driven by the dramatic drop in sample sizes and the power of the empirical tests. However, the economic magnitude is comparable to that estimated from earlier regressions.

4. Source of Endogeneity and Identification

4.1. Source of Endogeneity

We document a direct and positive impact of bank on the evolution of borrower's ESG profile. However, interpreting the result as causal evidence can be confounded by a few endogeneity concerns.

The first source of endogeneity is the reverse causality introduced by selection problems. There are two types of selection problems embedded in our current framework. One is that borrowers with certain level of ex ante ESG rating ($ESG\_borrower_{i,t-1}$) may self-select to borrow from high ESG standard banks. We alleviate this concern by controlling for the borrower's ex ante ESG rating. By holding the borrower's ESG standard constant, we explore how the difference in the bank's ESG standard affect the borrower's ex post improvement in ESG performance. The second type of the selection problem is that borrowers who expect to improve their ESG performance ($ESG\_Chg_{i,t-1,t+1}$) may self-select to borrow from high ESG standard banks. If this is the case, the borrower's ex post ESG evolution reversely leads to the lending relationship with high ESG standard bank.

The second source of endogeneity is omitted variable bias. One such potentially omitted variable is the CEO's awareness/concern related to ESG issues. Specifically, borrowers with CEOs who are sensitive to ESG issues are more likely to improve their ESG performance over time; at the same time, they are also more likely to borrow from high quality and high ESG standard banks. This behavior simultaneously causes variations in both the dependent and independent sides of the regression, which contaminates the causal interpretation of the main results.

4.2. Difference-in-Difference Analysis using Bank Mergers

Following Asker and Ljungqvist (2010), Hong and Kacperczyk (2010), Ergungor et al. (2015) and Chen and Vashishtha (2017), our identification strategy leverages exogenous shocks to the bank's ESG standard arising from bank mergers. Specifically, we examine how borrowers react to exogenous variations in the lead lender's ESG standard. This Diff-in-Diff strategy is best suited to our study for two reasons. First, it helps disentangle the selection and treatment effects, by looking at shocks to lenders in the already established lending relationships. In other words, the shocks take place after the borrower-lender matching is completed. Second, the timing and the decision of bank M&A activities are arguably exogenous to the borrowers' firm-level unobservable characteristics. As noted by prior studies, the bank merger waves were largely driven by regulatory, technological, and competitive changes (Pilloff (2004)).

To further absorb the impact from borrower level and lender level omitted variables, we include the borrower, lender, year, and industry FEs in the following Diff-in-Diff specification:

$$\begin{aligned} RRI_{i,t} =& \alpha + \beta ESG\_Shock_j \times Post_t + \varsigma ESG\_Shock_j + \tau Post_t \\ &+ \gamma X_{i,j} + I_{ffindustry} + \nu_i + \chi_j + \delta_t + \xi_{i,j,t} \end{aligned} \tag{8}$$

This specification represents a Panel OLS regression of the borrower's monthly RepRisk Indexes (RRI) over a 48-month window around the M&A event. The treatment group consists of all loans where the lender is involved in an M&A event within a five-year window after the loan initiation date.[12] We obtain the monthly RRI (if available) from 24 months before to 24 months after the M&A date. ESG_Shock is the exogenous variation to the lender's ESG standard in the merger and acquisition. Post dummy equals one if the date of the monthly RepRisk Index is after the M&A event date. We also include the log loan amount, country of syndication USA, the borrower's public status, and the number of covenants in the loan to control for the heterogeneity in size, regulatory environment, managerial myopia, and the credit risk, respectively. ν_i, χ_j, $I_{ffindustry}$

[12]The median length of loans in our sample is five years. We require that the M&A event happens less than five years ($\leq$ 4 years) since the loan initiation date, to allow at least one year for the exogenous variation to transmit through lending relationship.

and δ_t denote the dummies for borrower, lender, Fama-French 12 Industry and year fixed effects. Finally, standard errors are clustered at the borrower level.

We quantify the magnitude of the shock to the lender's ESG standard by incorporating the size effect. If the lender is the acquiror in the M&A, and the target is extremely small relative to the acquiror, we assume that the shock to the acquiror's ESG standard post-M&A is virtually zero. Empirically, we calculate ESG_Shock_j for the treatment group using the following specification, and assign zero to all control units:

$$ESG\_Shock_j = (RRI_a - RRI_t) \times Size_a/(Size_a + Size_t) \; , if \; the \; lender \; j \; is \; the \; target \tag{9}$$

$$ESG\_Shock_j = -(RRI_a - RRI_t) \times Size_t/(Size_a + Size_t) \; , if \; the \; lender \; j \; is \; the \; acquiror \tag{10}$$

We pair each treated loan one-to-one with a control group. We first require the control unit to be initiated in the same year-month as the treated loan. This guarantees that the DiD inferences are not being driven by time-series dynamics in the syndicated loan market. The second binding requirement is that the borrower and the lender in the control unit must be different from the borrower and the lender in the treated loan. Third, we choose the borrower with the closest ex ante RRI (measured at the time of loan initiation) to that of the borrower in the treated group. This setup ensures that the assignment of treatment vs. control is orthogonal to the main endogenous variable of interest – borrower's historical RRI. Finally, if there are multiple potential control units with the same ex-ante borrower's RRI, we compare and pick the one with the closest ex-ante lender's RRI.

[Insert Table 8 here]

Table 8 reports the balancing test between the ex-ante characteristics of borrowers in the treatment and control groups. Facility date is the loan initiation date. RepRisk Index is measured ex-ante at the facility start date, rather than at the merger and acquisition date. Public refers to the public status of the borrowers. Log (assets) (if publicly available) compares the size of the borrowers between the treatment and control group. The T-statistics of two-side difference tests are reported in parentheses. None of the reported characteristics are statistically different across groups. Finally, we apply both borrower and lender fixed effects in the DiD to absorb any remaining unobservable heterogeneities between the control and treatment units.

[Insert Table 9 here]

Table 9 reports the main results from the DiD analysis. The key coefficient estimate of the interaction term, $ESG\_Shock_j \times Post_t$, is positive and statistically significant at the 1% level. It indicates that shocks to the lender's ESG standard propagates through the lending relationship post-M&A, causing a change in the borrower's ESG performance in the same direction, and in proportion to the magnitude of the exogenous shock to the lender's ESG rating. Significant coefficient of Post captures an upward trend in the RRI over time. In columns 1 and 2, we regress with only control variables. In column 3, we add the lender fixed effects. In column 4, we apply both the borrower and the lender fixed effects, and observe a significant reduction in the explanatory power of the control variables. The coefficient of interaction term remains positive and statistically significant, and our DiD inference is robust to variations in specifications.

5. Robustness Tests

In this section, we conduct several robustness tests to confirm our baseline results related to borrower ESG rating evolution. Section 5.1 calculates the main explanatory variable using the raw, instead of adjusted RRIs. Section 5.2 considers alternative method to define lead lender(s). Section 5.3 examines alternative sampling criteria. Section 5.4 performs analysis on alternative specifications.

5.1. Measuring ESG Rating Difference Between Borrowers and Lenders

In our main specifications, we adjust both the borrower and lender ESG ratings by the industry-year averages. This alleviates the potential concerns about the comparability of ESG ratings across industries and years.

We now investigate whether our results are sensitive to using raw ESG ratings of lenders and borrowers when comparing their relative standing. Table 10, columns (1) presents the results. We find that our results do not change if we do not adjust the borrower and lender ESG ratings by industry-year averages.

[Insert Table 10 here]

5.2. Lender Profile, Single Lead Loans, and Strongest Relationship Lead Lenders

In our main specifications, when there are multiple lead lenders in the loan syndicate, we calculate the lead lenders' ESG rating by taking the average lender ESG rating for each loan. It is not clear which lender dictates the relationship and influences the borrower, therefore we follow this conservative approach. However, would the results change if we were to instead choose one of the lead lenders randomly, or if we have followed an alternative approach? We empirically address this robustness concern in this section.

We first start with a conservative, simplistic approach. We run the baseline estimation for the subsample of loans where there is a single unique lead lender in the loan syndicate. Columns (2) of Table 10 presents the results. We find that our results are unchanged for the subsample of loans where we have a single lender.

Other alternative approaches for choosing lead lenders include choosing the lead lender with the strongest historical relationship with the borrower, or randomly choosing one of the lead lenders as the lead for the loan. We test the former approach as it is more intuitive (Ivashina and Kovner (2011)). We classify a lead lender as the strongest relationship lead lender if the lead lender financed the greatest fraction of loan amount in the past five years before the current loan. Columns (3) shows our results under this alternative approach. We again find that if the lender has a better ESG rating, the borrower's ESG rating is more likely to improve.

5.3. Sample Selection Criteria

A final, sample related robustness check is related to sample selection criteria. In our main regressions we try to keep our sample as large as possible for representativeness as well as statistical power. In this section, we test our findings to the usage of alternative sample selection criteria. We impose the most common filters on banking and syndicated loans literature. Specifically, we require the loans to be USD denominated, and issued by non-financial, non-utility, US-incorporated borrowers. Columns (4) of Table 10 presents our findings under these criteria. We find that our results are in fact stronger under this approach. Overall, the robustness tests confirm our finding that high ESG rating lenders influence their borrowers in their ESG policies.

5.4. Alternative Specification

In our main analysis, we regress the improvement in the borrower's ESG over time, on the ex-ante difference between the lender's and borrower's ESG ratings, while controlling for the borrower's ex-ante ESG standard. This empirical specification views each loan initiation as an "event" and makes sure each facility enters only once into the analysis. The empirical design alleviates concerns on the stacking of sticky ESG scores/ratings in the panel regression.

In this section, we repeat our baseline analysis in Table 3 using levels, rather than changes as main variables of interests using the following specification:

$$\begin{aligned} RRI\ borrower_{t+1} = &\alpha + \beta RRI\ lender_{t-1} + \varsigma RRI\ borrower_{t-1} + \tau Lender\ Chg_j \\ &+ \gamma X_{i,t-1} + I_{ffindustry} + \delta_t + \xi_{i,j,t} \end{aligned} \tag{11}$$

$RRI\ borrower_{t+1}$ is defined as the level of borrower's RepRisk Indexes (rather than the change) one year after the loan initiation date. $RRI\ lender_{t-1}$ is defined as the level of lender's RepRisk Indexes one year before the loan initiation date. $RRI\ borrower_{t-1}$ is defined as the level of borrower's RepRisk Indexes one year before the loan initiation date. We also include the loan amount (in millions), country of syndication USA, the borrower's public status, and the number of covenants in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. In column 4, we perform a sub sample analysis in the public space only, and control for borrower's financials including log assets, book leverage, ROA, and Tobin's Q. Note that the level of observation in this analysis is facility. Each loan only enters once into the regression. And the year t is defined as the year of loan initiation.

Results are reported in Appendix B. In column 1, 2 and 3, the coefficients are statistically significant and economically sizable. Column 4 presents the subsample analysis focusing on public borrowers only. The T statistic of 1.53 is not significant at the 10% level, but the economic magnitude remains comparable to estimates in earlier columns.

6. Conclusion

We demonstrate that banks have an important influence on firm ESG policies. Specifically, we find that banks are significantly more likely to partner with borrowers that have similar ESG ratings. This result suggests that ESG policies influence the construction of bank lending relationships and that different banks have different attitudes towards borrower ESG policies that are at least partly influenced by the bank's own ESG-related policies and experiences.

We also find that banks have an important influence on the evolution of their borrowers ESG levels. Firms that from banks with relatively better ESG reputations are more likely to improve their own ESG levels over time. By contrast, banks with relatively worse reputations are less likely/less able to nudge their borrowers to take steps to enhance their ESG-related investments. We also find that banks are more likely to influence bank-dependent borrowers, and that their influence is predominantly concentrated among the environmental component of the ESG spectrum. In our third set of tests, we find that borrowers are more likely to experience a shift in their lending relationship following an adverse shock to their ESG-related reputation.

All in all, our results clearly demonstrate that the banking system has an important systematic effect on corporate ESG policies. In this regard, we further demonstrate a specific channel in which a key stakeholder can profoundly promote socially responsible decision making.

References

Asker, J., Ljungqvist, A., 2010. Competition and the structure of vertical relationships in capital markets. Journal of Political Economy 118, 599–647.

Bartram, S. M., Hou, K., Kim, S., 2018. Real effects of climate policy: Financial constraints and spillovers. Fisher College of Business Working Paper p. 004.

Bénabou, R., Tirole, J., 2010. Individual and corporate social responsibility. Economica 77, 1–19.

Berger, A. N., Udell, G. F., 1995. Relationship lending and lines of credit in small firm finance. Journal of business pp. 351–381.

Bharath, S., Dahiya, S., Saunders, A., Srinivasan, A., 2007. So what do i get? the bank's view of lending relationships. Journal of Financial Economics 85, 368–419.

Bharath, S. T., Dahiya, S., Saunders, A., Srinivasan, A., 2009. Lending relationships and loan contract terms. Review of Financial Studies 24, 1141–1203.

Brammer, S., Millington, A., 2008. Does it pay to be different? an analysis of the relationship between corporate social and financial performance. Strategic Management Journal 29, 1325–1343.

Cao, J., Titman, S., Zhan, X., Zhang, W. E., 2019. Esg preference and market efficiency: Evidence from mispricing and institutional trading. Available at SSRN 3353623 .

Chava, S., 2014. Environmental externalities and cost of capital. Management Science 60, 2223–2247.

Chava, S., Roberts, M. R., 2008. How does financing impact investment? the role of debt covenants. Journal of finance 63, 2085–2121.

Chen, Q., Vashishtha, R., 2017. The effects of bank mergers on corporate information disclosure. Journal of Accounting and Economics 64, 56–77.

Chen, T., Harford, J., Lin, C., 2015. Do analysts matter for governance? evidence from natural experiments. Journal of Financial Economics 115, 383–410.

Dai, R., Liang, H., Ng, L., 2018. Socially responsible corporate customers. Available at SSRN 3111073 .

Diamond, D. W., 1991. Monitoring and reputation: The choice between bank loans and directly placed debt. Journal of Political Economy 99, 689–721.

Dimson, E., Karakaş, O., Li, X., 2015. Active ownership. Review of Financial Studies 28, 3225–3268.

Ergungor, O. E., Madureira, L., Nayar, N., Singh, A. K., 2015. Lending relationships and analysts forecasts. Journal of Financial Intermediation 24, 71–88.

Friedman, M., 1970. A friedman doctrine: The social responsibility of business is to increase its profits. New York Times 13, 32–33.

Gopalan, R., Udell, G. F., Yerramilli, V., 2011. Why do firms form new banking relationships? Journal of Financial and Quantitative Analysis 46, 1335–1365.

Goss, A., Roberts, G. S., 2011. The impact of corporate social responsibility on the cost of bank loans. Journal of Banking & Finance 35, 1794–1810.

Hart, O., Zingales, L., 2017. Serving shareholders doesnt mean putting profit above all else. Harvard Business Review 12, 2–6.

Hasan, I., Hoi, C. K., Wu, Q., Zhang, H., 2017. Social capital and debt contracting: Evidence from bank loans and public bonds. Journal of Financial and Quantitative Analysis 52, 1017–1047.

Hauptmann, C., 2017. Corporate sustainability performance and bank loan pricing: It pays to be good, but only when banks are too. Saïd Business School WP 20.

He, Y., Kahraman, B., Lowry, M., 2018. Mutual fund voting on environmental and social proposals. Available at SSRN 3284683 .

Hoepner, A. G., Oikonomou, I., Sautner, Z., Starks, L. T., Zhou, X., 2018. Esg shareholder engagement and downside risk .

Holmstrom, B., Tirole, J., 1997. Financial intermediation, loanable funds, and the real sector. Quarterly Journal of Economics 112, 663–691.

Hong, H., Kacperczyk, M., 2010. Competition and bias. Quarterly Journal of Economics 125, 1683–1725.

Hubbard, R. G., Kuttner, K. N., Palia, D. N., 2002. Are there bank effects in borrowers' costs of funds? evidence from a matched sample of borrowers and banks. Journal of Business 75, 559–581.

Ivashina, V., Kovner, A., 2011. The private equity advantage: Leveraged buyout firms and relationship banking. Review of Financial Studies 24, 2462–2498.

Kim, T., Xu, Q., 2017. Financial constraints and corporate environmental policies. In: *HKUST Finance Symposium*.

Li, J., Wu, D. A., 2017. Does corporate social responsibility benefit society? .

Lins, K. V., Servaes, H., Tamayo, A., 2017. Social capital, trust, and firm performance: The value of corporate social responsibility during the financial crisis. Journal of Finance 72, 1785–1824.

Nini, G., Smith, D. C., Sufi, A., 2012. Creditor control rights, corporate governance, and firm value. Review of Financial Studies 25, 1713–1761.

Ongena, S., Smith, D. C., 2001. The duration of bank relationships. Journal of Financial Economics 61, 449–475.

Petersen, M. A., Rajan, R. G., 1994. The benefits of lending relationships: Evidence from small business data. Journal of Finance 49, 3–37.

Pilloff, S. J., 2004. Bank merger activity in the united states, 1994-2003. Federal Reserve Bulletin p. 328.

Schenone, C., 2009. Lending relationships and information rents: Do banks exploit their information advantages? Review of Financial Studies 23, 1149–1199.

Schiller, C., 2018. Global supply-chain networks and corporate social responsibility. In: *13th Annual Mid-Atlantic Research Conference in Finance (MARC) Paper*.

Schwert, M., 2018. Bank capital and lending relationships. Journal of Finance 73, 787–830.

Sharpe, S. A., 1990. Asymmetric information, bank lending, and implicit contracts: A stylized model of customer relationships. Journal of Finance 45, 1069–1087.

Shive, S., Forster, M., 2019. Corporate governance and pollution externalities of public and private firms. Available at SSRN 3339517 .

Shleifer, A., Vishny, R. W., 1997. A survey of corporate governance. Journal of finance 52, 737–783.

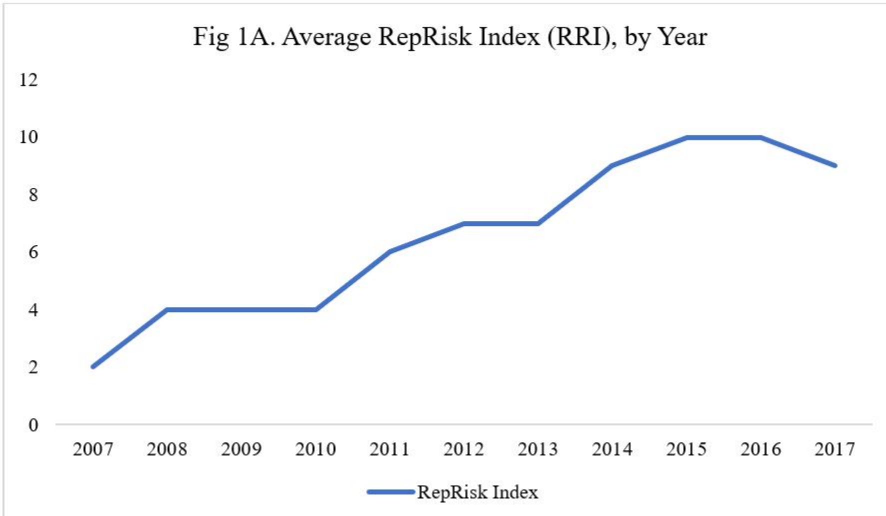
Starks, L. T., Venkat, P., Zhu, Q., 2017. Corporate esg profiles and investor horizons. Available at SSRN 3049943 .

Stein, J. C., 2002. Information production and capital allocation: Decentralized versus hierarchical firms. Journal of Finance 57, 1891–1921.

Strahan, P. E., 1999. Borrower risk and the price and nonprice terms of bank loans. FRB of New York staff report .

Fig. 1. RepRisk Index (RRI) by Year and Industry

The following figures show the mean level of unadjusted RRI by year, and by industry. The sample includes 126 monthly RRI for each borrower in our sample (from Jan 2007 to June 2017). Industry classifications are based on the Fama-French 12 industry classifications.



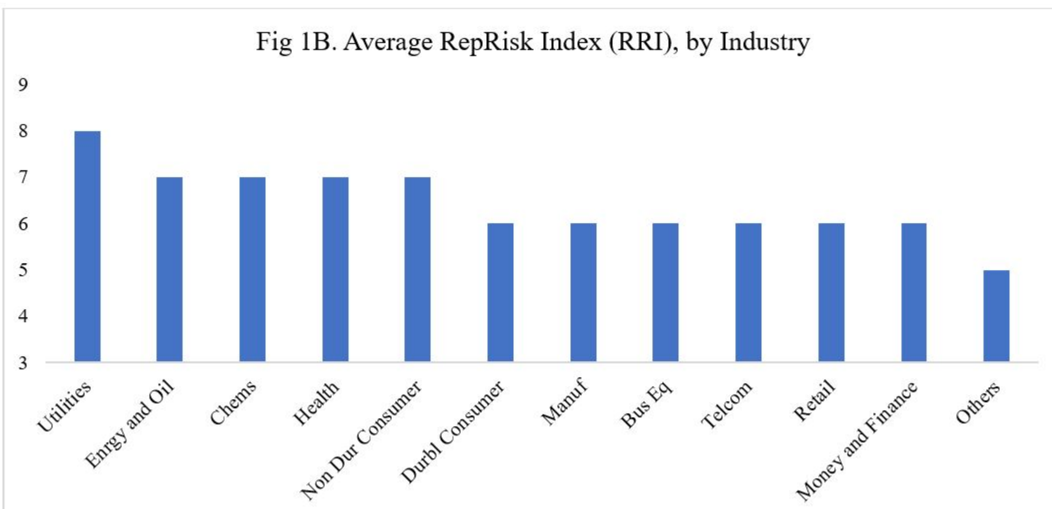


Table 1
Descriptive Statistics

This table summarizes sample statistics. All variables are reported on the loan level (borrower-lender-facility). *Public*, *Secure*, *Rated* and *Investment Grade* are dummy variables. *Log Assets*, *Book Leverage*, *Return on Assets* and *Tobin's Q* are only available for public firms and select private firms through Capital IQ. Detailed variable definitions are available in Appendix A.

Variable	N	Mean	St. dev.	P10	P50	P90
ESG_Chg	12495	1.88	11.46	-12.00	0.00	19.00
ESG_Borrower	12495	7.65	11.74	0.00	0.00	24.00
ESG_Lender	12495	18.60	21.06	0.00	17.00	60.00
Unadjusted ESG_Diff	12495	10.94	23.00	-16.00	1.00	44.00
ESG_Diff	12495	11.39	20.72	-12.00	7.00	40.00
Rated	12495	0.61	0.49	0.00	1.00	1.00
# of Covenants	12495	0.74	1.00	0.00	0.00	2.00
Secure	12495	0.43	0.50	0.00	0.00	1.00
Investment Grade	12495	0.29	0.45	0.00	0.00	1.00
Public	12495	0.62	0.48	0.00	1.00	1.00
Total Assets	8721	22214	87050	699	4858	40293
Book Leverage	8721	0.34	0.23	0.07	0.31	0.62
Return on Assets	8597	0.03	0.10	-0.03	0.04	0.11
Tobin's Q	7660	1.71	0.94	1.00	1.47	2.61

Table 2

Borrower and Lender's Endogenous Matching

The following table reports the Probit regression of the distance between the lender and borrower's ESG indexes, on the likelihood of initiating a loan. *Close* is defined as the absolute value of the lagged difference between the borrower's and lender's (yearly average) RepRisk Indexes. *Close_adj* further adjusts for the country-industry-year average. Match equals one if the lender lends at least once to the borrower in a year. Prior equals one if a loan was previously initiated between the lender-borrower pair in the past 4 years (from year t = -4 to t = -1). In Panel A column 1, we report the basic regression with lender, industry and year fixed effects, and clustering of standard errors on the lender-borrower pair level. Panel A Column 2 shows that our results are robust in the public space, where we control for borrower's financials including *log assets*, *book leverage*, *tobin's Q* and *investment grade* dummy. Column 4 and 5 repeat the tests in column 1 and 2 using *close_adj*. In column 3 and column 6, we show that our results are robust by excluding borrowers with (adjusted) RepRisk Indexes (*ESG_Borrower*) equal zero. Panel B repeats the regression analysis using the sample of single-lead loans only. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the lender-borrower pair level. Z statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Panel A	(1) match All	(2) match Public	(3) match ESG_Borrower>0	(4) match All	(5) match Public	(6) match ESG_Borrower>0
Close	-0.00791***	-0.00605***	-0.00711***			
	(-14.10)	(-8.28)	(-7.33)			
Close_adj				-0.00639***	-0.00461***	-0.00413**
				(-11.78)	(-6.30)	(-2.15)
Prior		0.934***	0.851***		0.930***	0.645***
		(19.91)	(15.59)		(19.78)	(6.43)
Log Assets		0.0926***	0.101***		0.0946***	0.118***
		(11.33)	(10.28)		(11.61)	(6.62)
Book Leverage		0.161***	0.210***		0.156***	0.420***
		(3.40)	(3.67)		(3.31)	(3.40)
Tobin's Q		0.0176	0.0109		0.0183	0.0197
		(1.41)	(0.73)		(1.46)	(0.81)
Investment Grade		-0.0698***	-0.0628**		-0.0705***	-0.0746
		(-3.07)	(-2.38)		(-3.10)	(-1.49)
Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	460408	197328	126764	460408	197328	28589
Pseudo R2	0.291	0.334	0.342	0.290	0.333	0.348

Panel B	(1) match All	(2) match Public	(3) match ESG_Borrower>0	(4) match All	(5) match Public	(6) match ESG_Borrower>0
Close	-0.00660***	-0.00845***	-0.0102***			
	(-7.25)	(-5.84)	(-5.01)			
Close_adj				-0.00403***	-0.00549***	-0.00889**
				(-4.40)	(-3.75)	(-1.98)
Prior		1.016***	0.809***		1.017***	0.783***
		(13.16)	(8.07)		(13.21)	(3.86)
Log Assets		-0.00962	-0.00152		-0.00773	-0.0000221
		(-0.71)	(-0.08)		(-0.57)	(-0.00)
Book Leverage		-0.00901	-0.0189		-0.0162	-0.0543
		(-0.10)	(-0.17)		(-0.19)	(-0.18)
Tobin's Q		-0.0230	-0.0242		-0.0207	-0.0248
		(-0.95)	(-0.79)		(-0.86)	(-0.42)
Investment Grade		-0.0170	-0.0238		-0.0175	-0.0388
		(-0.41)	(-0.43)		(-0.43)	(-0.30)
Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Lender FE	Yes	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	147715	48724	24284	147715	48724	3512
Pseudo R2	0.207	0.250	0.250	0.206	0.248	0.243

Table 3

Evolution in Corporate ESG Profile and Bank Lending

The following table reports the OLS regression of the change in the borrower's ESG profile on the ex-ante difference between the bank and borrower's ESG ratings. The change in the borrower's ESG profile (ESG_Chg) is defined as the difference between the borrower's RepRisk Indexes over a two-year window, from one year before to one year after the loan initiation date. The ex-ante difference between the bank and borrower's ESG ratings (ESG_Diff) is defined as the difference between the bank and borrower's RepRisk Indexes measured one year before the loan initiation date. $Lender\_Chg$ controls for the evolution in the lender's ESG indexes over the same two-year window. $ESG\_Borrower$ controls for the potential self-selection problem and is defined as the borrower's RepRisk Index one year before the loan initiation date. In column 1, we report the basic regression without fixed effects and clustering of standard errors. Column 2 clusters the standard errors on the borrower level. Column 3 adds industry and year fixed effects. In column 4, we also include the *log loan amount*, *country of syndication USA*, the *borrower's public status*, and the *number of covenants* in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. In column 5, we should that our results are robust in the sub sample of public firms only, and we further control for borrower's financials including *log assets*, *book leverage*, *ROA*, and *Tobin's Q*. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) ESG_Chg All Loans	(2) ESG_Chg All Loans	(3) ESG_Chg All Loans	(4) ESG_Chg All Loans	(5) ESG_Chg Public
ESG_Diff	0.0746***	0.0746***	0.0623***	0.0376***	0.0379***
	(12.71)	(8.93)	(6.40)	(4.04)	(2.91)
Lender_Chg	0.0558***	0.0558***	0.0431***	0.0199	0.0320*
	(6.35)	(4.16)	(3.11)	(1.49)	(1.78)
ESG_Borrower	-0.393***	-0.393***	-0.413***	-0.506***	-0.583***
	(-45.22)	(-12.99)	(-12.13)	(-16.02)	(-17.38)
Log Loan Amt				1.411***	0.350*
				(11.32)	(1.94)
USA				-0.407	2.519
				(-0.32)	(1.64)
Public				2.068***	
				(6.19)	
# of Covenants				-0.596***	-0.195
				(-4.14)	(-1.23)
Log Assets					2.295***
					(10.94)
Book Leverage					-2.073**
					(-2.39)
ROA					-0.146
					(-0.09)
Tobin's Q					0.751***
					(3.64)
Cluster	No	Yes	Yes	Yes	Yes
Ind FE	No	No	Yes	Yes	Yes
Year FE	No	No	Yes	Yes	Yes
N	12495	12495	12448	12448	7659
adj. R2	0.217	0.217	0.228	0.263	0.310

Table 4

Bank Dependency, Corporate ESG Profile, and Bank Lending

The following table reports the OLS regression of the change in the borrower's ESG profile on the ex-ante difference between the bank and borrower's ESG ratings. The change in the borrower's ESG profile (*ESG_Chg*) is defined as the difference between the borrower's RepRisk Indexes one year after and one year before the loan initiation date. The ex-ante difference between the bank and borrower's ESG ratings (*ESG_Diff*) is defined as the difference between the bank and borrower's RepRisk Indexes measured one year before the loan initiation date. Interaction terms of *ESG_Diff* and proxies of bank dependency are included. Proxies of bank dependency include the *rating* dummy, *number of covenants*, *secure* dummy, and *investment grade* dummy. *ESG_borrower* controls for the potential self-selection problem and is defined as the borrower's RepRisk Index one year before the loan initiation date. We also include the *log loan amount* (in millions), and *country of syndication USA* to control for the heterogeneities in size and regulatory environment. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T statistics for the regressions are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) ESG_Chg	(2) ESG_Chg	(3) ESG_Chg	(4) ESG_Chg
ESG_Diff	0.066***	0.028***	0.029***	0.059***
	(5.28)	(2.79)	(2.67)	(5.53)
Rated	2.678***			
	(6.29)			
ESG_Diff × Rated	-0.051***			
	(-3.38)			
# of Covenants		-0.537***		
		(-2.74)		
ESG_Diff × # of Covenants		0.016**		
		(2.00)		
Secure			-2.306***	
			(-5.34)	
ESG_Diff × Secure			0.027*	
			(1.76)	
Investment Grade				4.653***
				(7.94)
ESG_Diff × Investment grade				-0.085***
				(-4.01)
Lender_Chg	0.017	0.019	0.018	0.019
	(1.29)	(1.38)	(1.39)	(1.40)
ESG_Borrower	-0.512***	-0.493***	-0.503***	-0.534***
	(-16.08)	(-15.92)	(-16.54)	(-18.37)
Log Loan Amt	1.308***	1.537***	1.486***	1.227***
	(10.10)	(12.22)	(12.62)	(10.67)
USA	-0.513	-0.342	-0.401	-0.464
	(-0.43)	(-0.28)	(-0.34)	(-0.41)
Cluster	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	12448	12448	12448	12448
adj. R2	0.262	0.257	0.262	0.272

Table 5

Asymmetric Bank Impact

The following table reports the OLS regression of the change in the borrower's ESG profile on the ex-ante difference between the bank and borrower's ESG ratings. The change in the borrower's ESG profile (*ESG_Chg*) is defined as the difference between the borrower's RepRisk Indexes one year after and one year before the loan initiation date. The ex-ante difference between the bank and borrower's ESG ratings (*ESG_Diff*) is defined as the difference between the bank and borrower's RepRisk Indexes measured one year before the loan initiation date. *ESG_borrower* controls for the potential self-selection problem and is defined as the borrower's RepRisk Index one year before the loan initiation date. Panel A presents the results for the subsamples where lender has a better or worse ESG rating than the borrower. Samples in column (1) and (2) include only loans where the bank's RepRisk Index is smaller or equal to the borrower's RepRisk Index; samples in column (3) and (4) include those where the bank's RepRisk Index is larger than the borrower's. Panel B presents the results for the subsamples partitioned with respect to the difference between lender and borrower ESG ratings. A positive difference indicates that lender has a worse ESG score than the borrower. Column (1) and (2) focus only on loans where the *ESG_Diff* falls in the bottom quintile (*ESG_Diff*<0), while in column (3) and (4) we focus on loans where *ESG_Diff* falls in the top quintile (*ESG_Diff*>32). We also include the *log loan amount* (in millions), *country of syndication USA*, the borrower's *public* status, and the *number of covenants* in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Panel A	(1)	(2)	(3)	(4)
	Better bank = 1		Better bank = 0	
	ESG_Chg	ESG_Chg	ESG_Chg	ESG_Chg
ESG_Diff	0.133***	0.111***	0.017	0.018
	(4.64)	(3.25)	(1.33)	(1.06)
Lender_Chg	0.038	0.062	-0.014	0.012
	(1.01)	(1.40)	(-0.77)	(0.46)
ESG_Borrower	-0.407***	-0.581***	-0.591***	-0.632***
	(-6.34)	(-9.54)	(-20.87)	(-17.31)
Log Loan Amt	2.759***	0.712**	1.082***	0.059
	(9.23)	(2.09)	(6.77)	(0.24)
USA	-6.224**	-0.347	1.426	4.005**
	(-2.40)	(-0.12)	(1.07)	(2.57)
Public	2.719***		2.100***	
	(3.01)		(4.81)	
# of Covenants	-1.493***	-0.733*	-0.501**	-0.120
	(-3.58)	(-1.66)	(-2.43)	(-0.52)
Log Assets		3.698***		2.164***
		(8.96)		(7.90)
Book Leverage		-2.592		-2.363**
		(-1.14)		(-2.21)
Return on Assets		9.131		-2.689
		(1.58)		(-1.52)
Tobin's Q		0.560		0.821***
		(1.08)		(3.19)
Cluster	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	2380	1755	6260	3634
adj. R2	0.303	0.396	0.244	0.275

Panel B	(1)	(2)	(3)	(4)
	Last quintile		First quintile	
	ESG_Chg	ESG_Chg	ESG_Chg	ESG_Chg
ESG_Diff	0.268***	0.275***	0.026	-0.001
	(5.41)	(5.14)	(0.79)	(-0.03)
Lender_Chg	0.084***	0.078**	0.046	0.071
	(3.13)	(2.19)	(1.44)	(1.41)
ESG_Borrower	-0.378***	-0.470***	-0.590***	-0.666***
	(-6.92)	(-7.85)	(-13.02)	(-11.55)
Log Loan Amt	1.754***	0.656*	0.827***	0.034
	(8.09)	(1.85)	(3.86)	(0.10)
USA	-1.804	2.049	3.086	7.295***
	(-0.61)	(0.74)	(1.51)	(2.67)
Public	1.948***		2.392***	
	(2.88)		(3.72)	
# of Covenants	-1.038***	-0.782**	-0.637*	-0.673*
	(-3.36)	(-2.21)	(-1.89)	(-1.65)
Log Assets		2.198***		1.523***
		(5.41)		(3.57)
Book Leverage		-1.885		-1.056
		(-0.99)		(-0.62)
Return on Assets		7.727**		0.557
		(2.05)		(0.28)
Tobin's Q		0.376		0.859**
		(0.84)		(2.11)
Cluster	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
N	2619	1655	2493	1359
adj. R2	0.349	0.400	0.178	0.223

Table 6

Evolutions along Environmental (E), Social (S) and Governance (G) Dimensions

The following table reports the OLS regression of the change in the borrower's environmental (E), social (S), and governance (G) profiles on the ex-ante difference between the bank and borrower's ESG ratings. The change in the borrower's environmental profile (*Chg_E*) is defined as the difference between the borrower's environmental component of RepRisk Indexes over a two-year window, from one year before to one year after the loan initiation date. The changes in the borrower's social (*Chg_S*) and governance (*Chg_G*) profiles are constructed in the similar fashions. The ex-ante difference between the bank and borrower's ESG ratings (*ESG_Diff*) is defined as the difference between the bank and borrower's RepRisk Indexes measured one year before the loan initiation date. *Lender_Chg* controls for the evolution in the lender's ESG indexes over the same two-year window. *ESG_Borrower* controls for the potential self-selection problem and is defined as the borrower's RepRisk Index one year before the loan initiation date. In column 1 to 3, we include the *log loan amount*, *country of syndication USA*, the borrower's *public* status, and the *number of covenants* in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. In column 4-6, we show that our results are robust in the public space only, and we control for borrower's financials including *log assets*, *book leverage*, *ROA*, and *Tobin's Q*. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) Chg_E	(2) Chg_S	(3) Chg_G	(4) Chg_E	(5) Chg_S	(6) Chg_G
ESG_Diff	0.0143***	0.0161**	0.00767	0.0150***	0.0215**	0.00145
	(3.57)	(2.44)	(1.45)	(2.60)	(2.36)	(0.18)
Lender_Chg	0.00190	0.00922	0.00895	0.00827	0.00695	0.0168
	(0.29)	(1.02)	(1.08)	(0.96)	(0.55)	(1.39)
ESG_Borrower	-0.124***	-0.251***	-0.130***	-0.138***	-0.272***	-0.174***
	(-12.00)	(-14.08)	(-8.30)	(-10.38)	(-14.50)	(-8.94)
Log Loan Amt	0.417***	0.583***	0.413***	0.255***	0.102	-0.00748
	(9.20)	(5.75)	(6.76)	(3.47)	(0.60)	(-0.06)
USA	-0.0845	-0.458	0.140	0.251	0.830	1.438**
	(-0.19)	(-0.55)	(0.24)	(0.37)	(0.75)	(2.22)
Public	0.347***	0.712***	0.999***			
	(2.64)	(3.04)	(5.12)			
# of Covenants	-0.124**	-0.190*	-0.281***	-0.119*	0.0756	-0.152
	(-2.02)	(-1.85)	(-3.41)	(-1.80)	(0.63)	(-1.49)
Log Assets				0.389***	0.974***	0.932***
				(4.81)	(6.68)	(6.66)
Book Leverage				-0.758**	-1.129*	-0.186
				(-2.57)	(-1.85)	(-0.31)
ROA				0.358	-0.770	0.267
				(0.29)	(-0.70)	(0.28)
Tobin's Q				0.287***	0.0327	0.431***
				(2.81)	(0.22)	(2.72)
Cluster	Yes	Yes	Yes	Yes	Yes	Yes
Ind FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
N	12448	12448	12448	7659	7659	7659
adj. R2	0.104	0.133	0.064	0.118	0.153	0.091

Table 7

Exposure to Negative Reputational News Incidents and Switch in Lending Relationship

The following table reports the Probit regression of the number of the borrower's negative reputational news on the likelihood of initiating new loan(s) with the same lead lender within 12 months of the original loan's end date. *Num Rep Event* is the number of months with negative news coverage from the start to the end dates of the original loan. *Same* is the dummy variable that turns on if the borrower initiates new loan(s) with at least one of the same lead lenders within 12 months of the original loan end date. *Same Res* is defined more restrictively, as the dummy variable that turns on if the borrower initiates new loan(s) with exactly same group of lead lenders within 12 months of the original loan end date. *Same Sgl (Single)* is defined most restrictively, as the dummy variable that turns on if the original loan has a single lead lender, and the borrower initiates new loan(s) with the same lender within 12 months of the original loan end date. Note that we construct the sample to include only borrowers who need new financing to minimize the demand side heterogeneity. *ESG_borrower_start* is the borrower's adjusted RepRisk Index measured at the start date of the original loan. *Original loan length* refers to the number of years between the start and end dates of the original loan. Controls include the borrower's ex ante level and change (during the original loan window) in *log assets*, *book leverage*, *ROA*, and *Tobin's Q*. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Year FE is based on the end year of the original loan. Standard errors are clustered on the borrower level. Z statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) All same	(2) Public same	(3) All same res	(4) Public same res	(5) All same sgl	(6) Public same sgl
Num Rep Event	-0.0339**	-0.0378**	-0.0307**	-0.0362**	-0.0248	-0.0463
	(-2.56)	(-2.04)	(-2.34)	(-2.05)	(-0.92)	(-1.17)
ESG_Borrower_Start	0.00827***	0.000211	0.00637**	0.00168	0.00651	0.00200
	(2.64)	(0.06)	(2.14)	(0.43)	(0.97)	(0.22)
Book Leverage		0.383		0.223		-0.0956
		(1.44)		(0.85)		(-0.19)
Tobin's Q		-0.149*		-0.201**		-0.261*
		(-1.88)		(-2.47)		(-1.73)
ROA		1.070		0.619		1.847
		(1.51)		(0.96)		(1.30)
Log Assets		0.0787*		0.00415		0.220***
		(1.93)		(0.10)		(2.95)
Chg in Book Leverage		0.414		0.337		-0.439
		(1.09)		(0.93)		(-0.68)
Chg in Tobin's Q		0.0158		-0.0320		-0.173
		(0.19)		(-0.38)		(-1.20)
Chg in ROA		0.561*		0.395		2.555**
		(1.86)		(1.46)		(2.57)
Chg in Log Assets		0.0372		-0.0170		0.00447
		(0.35)		(-0.16)		(0.02)
Original Loan Length		-0.0369		0.00454		0.0559
		(-1.09)		(0.15)		(0.83)
Investment Grade		0.136		0.357***		-0.309
		(1.08)		(2.85)		(-1.37)
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes	Yes	Yes
N	2318	1495	2318	1495	702	388
pseudo R2	0.025	0.043	0.019	0.035	0.043	0.096

Table 8

Balancing Table

The following table reports the balancing test between the ex-ante profiles of borrowers in the treatment and control groups. *Facility date* is the loan initiation date. We construct the control group by selecting loans initiated in the same year-month as the treated loans. *RepRisk Index* is measured *ex-ante* at the facility start date, rather than at the merger and acquisition date. *Public* refers to the public status of the borrowers. *Log (assets)* (if publicly available) compares the size of the borrowers between the treatment and control group. Detailed variable definitions are available in the Appendix A. The T statistics of two-side difference tests are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

Name	Treatment				Control					
	mean	std	min	max	mean	std	min	max	diff-in-mean	t-statistic
Facility Date (*initiation year-month*)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.00	(0.00)
RepRisk Index (*ex-ante*)	4.81	10.81	0	62	4.06	8.48	0	39	0.74	(1.10)
Public (Y/N)	0.42	0.49	0	1	0.44	0.50	0	1	-0.03	(-0.76)
Log (assets)	8.38	2.14	3.15	13.86	8.31	1.62	5.45	13.37	0.07	(0.42)

Table 9
Diff-in-Diff Analysis using Bank Mergers

The following table reports the OLS regression of the borrower's monthly RepRisk Indexes (ESG) over a 48-month window around the M&A event. The sample consists of all loans where the lender is involved in a M&A event within a five-year window after the loan initiation date. We obtain the monthly RepRisk Indexes (if available) from 24 months before to 24 months after the M&A date. *ESG_Shock* is the exogenous variation to the lender's ESG profile in the merger and acquisition. *Post* dummy equals one if the date of the monthly RepRisk Index is after the M&A event date. We also include the *log loan amount*, *country of syndication USA*, the borrower's *public* status, and the *number of covenants* in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) ESG	(2) ESG	(3) ESG	(4) ESG
ESG_Shock×Post	0.170***	0.168***	0.195***	0.172***
	(3.15)	(3.11)	(3.37)	(2.87)
ESG_Shock	-0.0805	-0.0667	-0.193*	-0.0404
	(-1.19)	(-0.94)	(-1.84)	(-0.37)
Post	1.112**	1.117**	1.209**	0.923**
	(2.39)	(2.38)	(2.50)	(1.97)
Log Loan Amt	2.703***	2.496***	2.243***	0.0210
	(8.46)	(8.32)	(6.95)	(0.51)
USA		-9.536***	-13.31***	-0.0745
		(-5.22)	(-4.61)	(-0.49)
Public		4.580***	4.691***	1.736***
		(5.38)	(6.15)	(4.69)
# of Covenants		-0.839**	-0.831**	-0.510*
		(-2.17)	(-2.23)	(-1.66)
Borrower FE	No	No	No	Yes
Lender FE	No	No	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes
N	35805	35805	35805	35805
adj. R2	0.176	0.224	0.267	0.636

Table 10
Robustness Tests

This table reports four robustness tests for the baseline result of ESG evolution. Columns (1) presents the results if *ESG_diff* variable is calculated without country-industry-month adjustments. Sample in columns (2) include only loans with a unique lead arranger in the syndicate. Sample in column (3) repeat the baseline estimation based on the ESG rating of the lead lender with the strongest relationship with the borrower, instead of averaging the ESG ratings of the lead arrangers in the syndicate. Finally, column (4) presents the results under alternative sample selection criteria: USD-denominated loans of non-financial and non-utility US firms. The change in the borrower's ESG profile (*ESG_Chg*) is defined as the difference between the borrower's RepRisk Indexes one year after and one year before the loan initiation date. The ex-ante difference between the bank and borrower's ESG ratings (*ESG_diff*) is defined as the difference between the bank and borrower's RepRisk Indexes measured one year before the loan initiation date. *ESG_borrower* controls for the potential self-selection problem and is defined as the borrower's RepRisk Index one year before the loan initiation date. We also include the *log of loan amount* (in millions), *country of syndication USA*, the borrower's *public* status, and the *number of covenants* in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) ESG_Chg	(2) ESG_Chg	(3) ESG_Chg	(4) ESG_Chg
ESG_Diff	0.022***	0.029**	0.026***	0.039***
	(2.82)	(2.11)	(3.10)	(3.75)
Lender_Chg	0.011	0.004	0.021	0.015
	(0.84)	(0.23)	(1.49)	(1.05)
ESG_Borrower	-0.516***	-0.567***	-0.519***	-0.497***
	(-17.15)	(-12.68)	(-15.12)	(-13.46)
Log Loan Amt	1.442***	0.895***	1.372***	1.577***
	(11.37)	(4.87)	(10.19)	(12.31)
USA	-0.364	-1.721	0.064	-0.534
	(-0.29)	(-0.80)	(0.04)	(-0.27)
Public	2.063***	2.237***	1.994***	1.737***
	(6.16)	(3.99)	(5.36)	(4.59)
# of Covenants	-0.627***	-0.643***	-0.676***	-0.486***
	(-4.32)	(-2.78)	(-4.19)	(-3.10)
Year FE	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes
N	12448	2907	10496	9603
adj. R2	0.261	0.246	0.267	0.259

Appendix A. Variable Definition

Variable name	Description	Source
ESG_Chg	The change in the borrower's RepRisk Index from one year before, to one year after the loan initiation date	RepRisk
ESG_Borrower	The RepRisk Index of the borrower measured one year before the loan initiation date.	RepRisk
Lender_Chg	The change in the lender's RepRisk Index from one year before, to one year after the loan initiation date	RepRisk
ESG_Diff	The difference between the lender and borrower's country-industry-month adjusted RepRisk Indexes measured one year before the loan initiation date.	RepRisk
Rated	An indicator that equals one if the borrower is rated, and zero otherwise.	Compustat
Investment Grade	An indicator that equals one if the borrower is investment grade, and zero otherwise.	Compustat
# of Covenants	Number of financial covenants on the loan.	Dealscan
Secure	An indicator that equals one if the loan is secured, and zero otherwise.	Dealscan
Log Loan Amt	The natural logarithm of the size of the syndicated loan (in millions).	Dealscan
Public	An indicator that equals one if the borrower firm's equity is publicly traded, and zero otherwise.	CRSP
Total Assets	Borrower's total assets at the latest fiscal period that ended prior to loan start date.	Compustat
Book Leverage	The ratio of total book debt to total assets.	Compustat
Return on Assets	The ratio of net income to total assets.	Compustat
Tobin's Q	The ratio of market value of total assets to book value of total assets.	Compustat
Size of Target	The M&A transaction value divided by the percentage of target acquired (in millions).	SDC Platinum
Size of Acquirer	Value of the acquirer's asset LTM (in millions).	SDC Platinum
ESG_Diff_MA	The difference between the acquirer and target's RepRisk Indexes at the time of the M&A.	RepRisk and SDC Platinum
ESG_Shock	The shock to the ESG standard of the lender introduced by the M&A transaction, adjusted by the relative sizes of both parties involved in the transaction.	RepRisk and SDC Platinum

Appendix B. Alternative Specification

This Appendix replicates the analysis in Table 3 using a different specification. $RRI\_borrower_{t+1}$ is defined as the level of borrower's RepRisk Indexes one year after the loan initiation date. RRI_lender_{t-1} is defined as the level of lender's RepRisk Indexes one year before the loan initiation date. $RRI\_borrower_{t-1}$ is defined as the level of borrower's RepRisk Indexes one year before the loan initiation date. We also include the *log of loan amount* (in millions), *country of syndication USA*, the borrower's *public* status, and the *number of covenants* in the loan to control for the heterogeneities in size, regulatory environment, managerial myopia, and the credit risk, respectively. In column 4, we perform a subsample analysis in the public space only, and control for borrower's financials including *log assets*, *book leverage*, *ROA*, and *Tobin's Q*. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Standard errors are clustered on the borrower level. T-statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) RRI_borrower$_{t+1}$	(2) RRI_borrower$_{t+1}$	(3) RRI_borrower$_{t+1}$	(4) RRI_borrower$_{t+1}$
RRI_lender$_{t-1}$	0.0719***	0.0452***	0.0232***	0.0172
	(6.17)	(5.63)	(2.97)	(1.53)
RRI_borrower$_{t-1}$		0.533***	0.461***	0.383***
		(17.08)	(15.92)	(12.99)
Lender_Chg		0.0329**	0.0115	0.0212
		(2.45)	(0.89)	(1.22)
Log Loan Amt			1.441***	0.381**
			(11.38)	(2.08)
USA			-0.363	2.529*
			(-0.29)	(1.65)
Public			2.050***	
			(6.12)	
# of Covenants			-0.625***	-0.227
			(-4.31)	(-1.43)
Log Assets				2.314***
				(10.94)
Book Leverage				-2.045**
				(-2.36)
ROA				-0.144
				(-0.09)
Tobin's Q				0.755***
				(3.66)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes
N	12448	12448	12448	7659
adj. R2	0.065	0.307	0.338	0.393

Appendix C. The ESG Ratings of New Lead Lenders

The following table reports the OLS regression of the number of the borrower's negative reputational news on the changes in the lead lenders' ESG ratings (average ESG ratings of the new lenders minus the average ESG ratings of the lenders of the original loan). The new group of lenders are the banks that lend money to the borrower within 12 months of the original loan's expiration date. *Num Rep Event* is the number of firm-months with negative news coverage from the start to the end dates of the original loan. Note that we construct the sample to include only borrowers who successfully find new financing to minimize the demand side heterogeneity. *ESG_borrower_start* is the borrower's adjusted RepRisk Index measured at the start date of the original loan. Original loan length refers to the number of years between the start and end dates of the original loan. Controls include the borrower's ex ante level and change (during the original loan window) in *log assets*, *book leverage*, *ROA*, and *Tobin's Q*. Detailed variable definitions are available in the Appendix A. Industry FE is based on the Fama-French 12 industry classification. Year FE is based on the end year of the original loan. Standard errors are clustered on the borrower level. T statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) All Lender_Diff	(2) Public Lender_Diff	(3) Public Lender_Diff	(4) Public Lender_Diff
Num Rep Event	0.858***	1.084***	1.086***	0.386*
	(5.74)	(5.92)	(5.78)	(1.72)
ESG_Borrower_Start	-0.260***	-0.219***	-0.200***	-0.0762
	(-6.67)	(-4.14)	(-4.09)	(-1.54)
Book Leverage		-1.473	-2.386	-2.769
		(-0.57)	(-0.73)	(-0.83)
Tobin's Q		-2.174***	-2.488***	-1.875**
		(-2.73)	(-2.76)	(-2.02)
ROA		15.12**	13.29	10.13
		(2.28)	(1.55)	(1.21)
Log Assets		-1.223**	-1.160**	-0.330
		(-2.54)	(-2.38)	(-0.60)
Chg in Book Leverage			-3.233	-4.944
			(-0.62)	(-0.97)
Chg in Tobin's Q			-0.369	-0.446
			(-0.32)	(-0.38)
Chg in ROA			-1.760	-1.823
			(-0.28)	(-0.30)
Chg in Log Assets			1.636	1.070
			(1.15)	(0.79)
Original Loan Length				2.819***
				(6.29)
Investment Grade				0.351
				(0.25)
Industry FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Cluster	Yes	Yes	Yes	Yes
N	2155	1510	1407	1407
adj. R2	0.103	0.114	0.113	0.147

Appendix D. Reputation Risk Exposure and Risk-adjusted Capital Ratios

The following table is from Houston, Shan and Tian (2019). It reports the OLS regression of the bank's ESG and business conduct risk on the level of risk-adjusted Tier 1 capital ratio. The level of observation is on the bank-quarter level. *ESG_Lag* is the RepRisk Index of the bank at t-1 (lagged quarter). *Num_News* is the number of negative news coverage from t-5 to t-1 (in quarters). *Num_News_H* and *Num_News_VH* count the number of high impact and very high impact negative news coverage during the same window. *Num_News_Env*, *Num_News_Soc* and *Num_News_Emp* count the number of negative news coverage related to environmental, social and employee issues during the same window. Bank and Month fixed effects are included to focus on within-bank variations and to preclude the impact of common time trends. Standard errors are clustered on the bank level. T statistics are reported in parentheses. ***, **, and * denote significance at the 1%, 5%, and 10% level, respectively.

	(1) Tier 1	(2) Tier 1	(3) Tier 1	(4) Tier 1	(5) Tier 1	(6) Tier 1	(7) Tier 1
ESG_Lag	0.0194** (2.61)						
Num_News		0.00911*** (2.90)					
Num_News_H			0.0132*** (3.14)				
Num_News_VH				0.0241** (2.46)			
Num_News_Env					0.0429 (1.37)		
Num_News_Soc						0.0267*** (2.84)	
Num_News_Emp							0.0806** (2.41)
Bank FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Month FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Cluster by Bank	Yes	Yes	Yes	Yes	Yes	Yes	Yes
N	1340	1340	1340	1340	1340	1340	1340
adj. R2	0.607	0.616	0.616	0.609	0.606	0.606	0.609

Sea Level Rise and Municipal Bond Yields*

Paul Goldsmith-Pinkham† Matthew T. Gustafson‡ Ryan C. Lewis §
Michael Schwert¶

May 2019

Abstract

This paper examines the effects of climate change on the financing costs of state and local governments in the U.S. Using a sample of bonds issued by school districts in coastal states, we show that greater exposure to sea level rise (SLR) is associated with higher bond yields after controlling for time-varying local economic conditions. We find that SLR is only priced after the sharp upward revision in SLR projections in 2013. The effect is concentrated on the East and Gulf coasts and stronger at long maturities. While the pricing effects of sea level rise are statistically significant, they are economically small and indicate that financial markets do not currently anticipate a high probability of default induced by catastrophe in coastal communities over the next two decades.

*We thank Elaine Yao for excellent research assistance.
†Yale University; paul.goldsmith-pinkham@yale.edu.
‡Penn State University; mtg15@psu.edu.
§University of Colorado; Ryan.C.Lewis@colorado.edu.
¶The Wharton School, University of Pennslyvania; schwert@wharton.upenn.edu.

1 Introduction

Since the 2007 Intergovernmental Panel on Climate Change (IPCC) report, end-of-century sea level rise (SLR) projections have increased more than fourfold, with current upper bound SLR projections of 2.5 meters by 2100 (see e.g., Stocker et al. (2013); Sweet et al. (2017); DeConto and Pollard (2016)). Should these projections manifest, coastal communities will be greatly impacted. Hauer, Evans, and Mishra (2016) finds that a 1.8 meter SLR would inundate areas currently home to six million Americans, while Rao (2017) estimates that nearly one trillion dollars of coastal residential real estate is at risk. Moreover, the systemic and location-specific nature of SLR exposure makes it difficult for coastal communities to diversify the risks associated with SLR exposure.

In this paper, we use the municipal bonds to study how news of SLR risk is priced in financial markets. Three features of the municipal bond market make it an ideal laboratory to examine this question. First, as in other financial markets, the market prices of municipal bonds reflect investors' expectations of future outcomes. Second, the payoff profile of bonds makes the likelihood of large negative shocks a key driver of yields. Third, the sources of repayment for municipal bonds are local in nature, especially so for the school district bonds that comprise our sample and are backed by taxes on local real estate.

Our empirical strategy uses local variation in municipalities' exposure to SLR to measure how municipal bond prices changed as expectations for sea-level rise were revised upwards. During our sample, two key reports were released by the IPCC: in 2007, when the IPCC projected that sea level would rise by only 0.18 to 0.59 meters by the end of the century; and in 2013, when the IPCC doubled its previous projections, and the U.S.-based National Oceanic and Atmospheric Administration (NOAA) supplying an upper bound SLR projection of two meters. A key difficulty in identifying the effect of these reports is that geographic areas with similar SLR exposure may face time-varying shocks that are correlated with SLR exposure. This is particularly true in 2007, as the financial crisis hit immediately afterwards. To address this concern, we exploit variation in SLR exposure across municipalities within the same county, using county-by-year-month fixed effects, and estimate how the relative pricing of bonds issued by SLR exposed municipalities shifted following the release of the different IPCC reports.

We examine the impact of SLR exposure on municipal bond pricing in three periods. First, we find no significant relation between a municipality's SLR exposure and its municipal bond yields prior to

2007. This suggests that our county-year-month fixed effects absorb any differences in the economic conditions in exposed and unexposed municipalities. Second, from 2007 to 2013, we *also* find no significant differences between exposed and unexposed municipalities, suggesting no shift in prices in response to the 2007 IPCC report. In contrast, we find a significant positive relation between SLR exposure and municipal bond yields between 2014 and 2017. The emergence of SLR exposure as a determinant of bond yields suggests that the flurry of SLR-related scientific and news articles in 2013 and 2014 drove the rise in municipal bond yields, rather than other omitted variables.

Our estimates imply that a 10 percentage point (approximately 1.1 standard deviations) increase in the number of SLR exposed properties is accompanied by a 2.4 basis point increase in municipal bond yields, equivalent to two percent of the standard deviation of the yield. This indicates that municipal bond investors do accoutn for SLR expectations. However, the relatively small magnitude of the effect suggests that bond investors believe the probability of a catastrophic event occurring over the life of their bond is relatively small. Indeed, in additional unreported results, we find no significant relation between SLR exposure and the quantity of municipal debt issued.

We also examine heterogeneity in the pricing of municipal bonds. We identify geographic differences as a function of projected SLR (i.e. regions much more exposed to aggregate shifts), and find that the the entire post-2013 increase in the sensitivity of bond yields to SLR exposure was concentrated on the East Coast, where oceans are projected to rise more than twice as fast as on the West Coast. We also examine an area's beliefs regarding climate change. Within East Coast markets, the post-2013 increase in the price effect of SLR exposure is substantially higher in areas that are more worried about the expected impact of climate change. As municipal bonds tend to be held more locally, this likely reflects those beliefs magnified in the financial prices.

Taken together, our findings provide the evidence on how the financial markets view the sharp rise in expected sea-level rise that has occurred over the past decade. The SLR exposure risk has emerged as an increasingly important determinant of bond pricing over this period, but is still economically meaningful. Nevertheless, intuitive cross-sectional and time-series variation in the relation between SLR exposure and bond yields lend credence to the idea that our empirical strategy is precisely identifying SLR exposure as a new consideration among municipal bond investors.

Our paper is closely related to Painter (2018), who also studies the relation between climate change-induced flood risk and the cost of municipal financing. Specifically, Painter (2018) compares the cost of issuing municipal bonds in metropolitan areas shown to be exposed to future flood risk (according to Hallegatte et al. (2013)) to other bonds issued during the same state-year and finds a positive relation between flood exposure and municipal bond issuance fees and yields. Our paper differs in two important ways: first, consistent with evidence in Bernstein, Gustafson, and Lewis (2018) that long-run SLR risk already impacts the coastal property values that underly municipal debt, we find that the yields of both the short- and long-maturity bonds are affected by SLR exposure by the end of our sample period whereas Painter (2018) finds no effect of SLR exposure on short-term bonds. Second, we find no evidence of SLR exposure being priced prior to 2011. On the East Coast, where yields are most related to SLR exposure, the differential pricing of exposed area bonds does not emerge until 2011 for long-maturity bonds and 2014 for shorter maturity bonds.

Differences in our empirical approach may explain the incongruity of our findings. Crucially, we measure SLR exposure at the school district level, which allows us to disentangle the effect of underlying economic conditions from the effect of SLR exposure on municipal bond yields by comparing the yield of two school districts' bonds within the same county and month. In contrast, Painter (2018) compares the costs of initial bond offerings within a state-year, using a measure of climate risk that is only observed for coastal metropolitan counties. As a result, the comparison group for these counties is both non-metro coastal counties and non-coastal counties. These estimates of the effect of climate risks on municipal bonds may include non-climate-induced factors if the treatment and control areas are differentially impacted by economic shocks. This is particularly crucial during the onset of the Great Recession in 2007.

Overall, our findings suggest that financial markets are aware of SLR exposure when pricing municipal bonds. Around 2013 when the IPCC doubled their SLR projections, SLR exposure became a statistically significant factor in pricing municipal bond markets. SLR exposure still is not a first-order determinant of municipal bond prices by the 2017 end of our sample, suggesting that municipal bond investors estimate a low probability of catastrophic loss due to SLR in the coming decades.

2 Sample and Empirical Methods

Our empirical analysis studies the effect of SLR exposure on school district bond yields. We focus on bonds issued by school districts for two key reasons. First, much of the funding for public schools in the U.S. comes from taxes on local real estate, so there is a direct economic link between school districts' ability to repay debts and the anticipated effects of SLR on property values (Bernstein, Gustafson, and Lewis (2018)). Second, public education is an important use of municipal bond proceeds, amounting to 27% of new bond issues and 14% of the dollar amount issued from 1999 to 2017, so we are able to construct a large sample of school district bonds.

Municipal bond yields are drawn from the intersection of the Mergent Municipal Bond Terms and Conditions database and historical transaction price data from the Municipal Securities Rulemaking Board (MSRB). We select school district bonds from these data by screening on primary and secondary education as the use of proceeds. We restrict the sample to general obligation bonds that are backed by the taxing power of the issuer, excluding bonds that are backed by revenues from specific projects. After applying these criteria, our sample consists of 637,328 bonds from 10,871 issuers.

We construct a monthly panel of volume-weighted yields at the bond level. Following past literature (Schwert (2017)), we restrict attention to fixed-coupon tax-exempt bonds that trade at least ten times, to ensure uniformity and a minimum level of liquidity. Additionally, we exclude the first three months after issuance and the last year before maturity because these are times when yields are especially volatile. After applying these filters, the monthly panel consists of 2,220,994 observations for 231,735 bonds.

To link SLR exposures for these issuers, we identify the underlying geography for each school district. Using this geography, we can identify and link it to the expected sea-level rise measures from Bernstein, Gustafson, and Lewis (2018). In order to identify the underlying school districts for each bond issuer, we link the bond issuers names to school district names.[1]

After merging with the data on SLR exposure, our final sample consists of 544,071 observations of 68,241 bonds from 1,532 school district issuers. There are 18 states in our sample, but the observation count is skewed towards more populous states, with California, Texas, New Jersey, and New York accounting for 88% of the bond-month observations and 75% of the school districts. After winsorizing at the 1% level, the bond yields range from 0.74% to 5.73%, with an average yield of 3.38%. The dispersion in municipal bond yields is narrow relative to other bond markets (e.g. corporates) because of the extremely low historical default rate.

To identify the relation between an area's SLR exposure and its municipal bond yields, we regress the yields implied by secondary market municipal bond transactions on an indicator for the percentage of properties within a school district that would be inundated by a 6-foot SLR, which is approximately the upper bound projection for end of century SLR. We mitigate the possibility that SLR

[1] This name matching proceeds in multiple steps. First, we clean and make consistent state names and common abbreviations. We then remove all exact matches, and with the remaining issuers, we remove stop words (e.g. "vocational", "technical" and "elementary") and match issuers and names using these stripped down names. Finally.... Code for linking the districts and issuers is available upon request.

exposure relates to unobserved aspects of the area's economy in two primary ways. First, we include county-year-month fixed effects throughout our analyses, such that we identify the effect of SLR exposure on bond yields by comparing the yields on bonds from different school districts within the same county, traded in the same month. Second, we exploit the fact that SLR projections and awareness have significantly increased over the 1999 to 2017 sample period by focusing on intertemporal variation in the relation between SLR exposure and municipal bond yields. To the extent that a relation between SLR exposure and municipal bond yields emerges or increases as SLR projections become more dire and salient, it is unlikely that the relation we observe is driven non-SLR related factors.

In our first set of analyses, we pool the years following 2013 into a single indicator to capture the average change in the relationship between SLR exposure and yields following the 2014 IPCC report. We expect this to be the largest period of growth for multiple reasons: first, in early 2014, the IPCC released their 2013 climate assessment where they nearly doubled the projection for SLR over the next century; second, between 2013 and 2015, Rohling et al. (2013), Hinkel et al. (2015), and Grinsted et al. (2015) all validated the upper bound SLR projections established by Parris et al. (2012) and dramatically increased the lower bound. Finally, in addition, the potential for glacial collapse in Antarctica became a topic of conversation in the popular press in May of 2014. Bernstein, Gustafson, and Lewis (2018) show that this sequence of events was accompanied by a spike Google trends search intensity, peaking in May of 2014. For these tests, we estimate the following specification:

$$\text{Bond Yield}_{ijt} = c_{jt} + \alpha_1 \text{Pct. Exposed}_{ij} + \alpha_2 (\text{Post} \times \text{Pct. Exposed}_{ijt}) + \beta X_{ijt} + \epsilon_{ijt}, \tag{1}$$

for a bond issued by school district i, located in county j and trading in year-month t. The explanatory variable of interest is the interaction between Post and Pct. Exposed, where Post is an indicator for observations after 2013 and Pct. Exposed is the percentage of properties within school district i that would be inundated with a 6-foot rise in sea levels.

To construct the percent exposed measure we obtain property-level data from the real estate assessor and transaction datasets in the Zillow Transaction and Assessment Dataset (ZTRAX). We then determine each properties SLR exposure using the NOAA SLR viewer (Marcy et al. (2011)). Importantly, we use the NOAA's SLR calculator, which accounts for the fact that tidal variation and other coastal geographic factors affect the impact of global oceanic volume increases on local SLR.[2] The

[2]See Bernstein, Gustafson, and Lewis (2018) for more details regarding this SLR exposure definition.

measure we use is the percentage of homes in an area that would be inundated by a 6-foot SLR, since that is approximately the upper bound projections for year 2100 SLR and the largest SLR exposure reported on the NOAA website [3] SLR exposure is highly skewed, even in our sample, which is restricted to counties near the coast. Most school districts in our sample do not have any SLR exposed properties. The 75th, 90th, and 95th percentiles of Pct. Exposed are approximately 1%, 10%, and 20%, respectively.

To better understand how the relation between SLR exposure and municipal bond prices evolves around the two IPCC reports, released in 2007 and 2014, we then conduct separate regressions for each year of our sample period of the form:

$$\text{Bond Yield}_{ijt} = c_{jt} + \alpha_{1,t}\text{Pct. Exposed}_{ij} + \beta_t X_{ijt} + \epsilon_{ijt}, \tag{2}$$

for a bond issued by school district i, located in county j and trading in month t.

Table 1 summarizes the primary variables utilized in our analysis. In total we have 544,071 district by month observations in coastal counties, 238,360 of which have properties that will experience chronic inundation after 6 feet of global average sea level rise. On average, 7% of properties are exposed at the the 6 foot level in these districts. Municipal bonds issued by school districts in our sample average 3.36% for all districts and 3.38% for exposed districts. We find little difference in maturity, age and bond price variance between the exposed and full sample. Housing prices are approximately 10% higher in exposed districts and dollar bond volumes are approximately 20% higher than in the full sample.

3 Empirical Results

In Columns 1 and 2 of Table 2, we estimate Equation 1 over our full sample from 1998 to 2017, and a shorter 10-year sample concentrated around the post-2013 rise in the projections and publicity of SLR. In both cases, the Pct. Exposed main effect is statistically insignificant and small compared to the Post $\times$ Pct. Exposed interaction, which is positive and statistically significant. The small main effect suggests that there are limited differences across municipal bond yields in exposed and unexposed counties prior to the 2014 IPCC release. This bolsters our preferred interpretation of the positive and significant Post $\times$ Pct. Exposed interaction as the increased importance of SLR exposure risk in determining municipal bond yields since the 2014 IPCC release.

[3] We find qualitatively similar results using the percentage of properties exposed to a 3 foot SLR.

The estimated coefficient on the Post $\times$ Pct. Exposed interaction is 0.24 in both columns. This suggests (1) that our choice regarding the pre-period length does not affect our inferences, and (2) that the 2007 IPCC release had a limited effect on the relation between SLR exposure and municipal bond yields.

The estimates in Columns 1 and 2 imply that since the beginning of 2014, a 10 percentage point (or approximately 1.1 standard deviation) increase in the number of SLR exposed properties within a school district is accompanied by a 2.4 basis point or 0.5% increase in municipal bond yields. The above result indicates that municipal bond investors do factor SLR expectations into their investment decisions. However, the relatively small magnitude of the effect suggests that bond investors believe the probability of a catastrophic event occurring over the life of their bond is relatively small. Consistent with SLR exposure not being a first-order determinant of the cost of municipal debt, we find no significant relation between SLR exposure and the quantity of municipal debt a school district issues.

In Columns 3 to 5 of Table 2, we investigate the extent to which pricing varies according to geographic regions. Such differences may arise from multiple sources. First, there is substantial heterogeneity in historical SLR across the United States, which Piecuch et al. (2018) argue is primarily due to geological processes that will persist for centuries. East coast sea levels have been rising by between 1.70 mm/year (in Massachusetts and Maine) and 3.89 mm/year (in New Jersey, New York, Connecticut, Rhode Island, Maryland, and Virginia), which is approximately twice as fast as on the west coast. Second, state perceptions about the dangers posed by climate change vary dramatically across the United States which may impact the beliefs of the marginal Municipal bond investor. Consistent with both sources of heterogeneity, results from Bernstein, Gustafson, and Lewis (2018) suggest that the home price discount from SLR exposure is concentrated in the East Coast where SLR projections are more pessimistic. School district bond yields follow a similar pattern with pricing effects evident on the eastern seaboard but not on the Pacific coast. We find that the estimated coefficient on Post $\times$ Pct. Exposed interaction is 0.36 in the East Coast sample and close to zero in the West Coast and Gulf implying that our full sample results are driven entirely by the East Coast.

In Figure 1, we decompose the pre- and post-periods to provide evidence on the timing with which East Coast municipal bond markets began pricing SLR exposure. Each point on the solid line represents the estimated coefficient on Pct. Exposed obtained from estimating Equation 2 over the year indicated on the x-axis.[4] The dashed lines present the 95% confidence interval on this coefficient

[4]By estimating separate regression each year we allow the coefficients on control variables to vary by year, meaning that the estimated coefficients in the figure need not average to those reported in Table 2.

estimate.

The figure shows that the coefficient is only significant during two of the seventeen years prior to the 2014 IPCC release. Moreover, the coefficients between 1998 and 2013 exhibit no particular pattern, oscilating between 0.05 and -0.15. Since the increased SLR projections and media attention in 2014, however, the Pct. Exposed coefficient has been consistently positive and significant. In 2014, the coefficient rose to approximately 0.18 (from approximately 0.01 in 2013), and since 2014 the coefficient has been stable at over 0.3. This pattern is consistent with the East Coast municipal bond markets incorporating news about SLR exposure risk into bond prices around the time that such news was released.

3.1 Additional Heterogeneity

In Table 3, we conduct two sets of tests to better understand what geographical characteristics predict the extent to which the municipal bond market prices SLR exposure. First, we more directly examine the role of past SLR, which also proxies for expected future SLR (see e.g., Piecuch et al. (2018)). To do this, we interact Pct. Exposed with a continuous measure of a state's historical SLR. We use a standardized measure of state-level SLR exposure, which we denote State SLR. Specifically, we obtain the historical state-level SLR reported by the NOAA, subtract the sample-wide average, and divide by the standard deviation.[5]

In Columns 1 and 2 of Table 3, the primary coefficient of interest is that on the triple interaction between Post × Pct. Exposed × State SLR. The positively significant coefficient on this triple {interaction} indicates that SLR exposure becomes increasingly important in pricing municipal bonds following the 2014 IPCC release in areas with faster rising sea levels. The statistically insignificant Post × Pct. Exposed interaction indicates that there is no increased role of SLR exposure in municipal bond pricing following the 2014 IPCC release in the U.S. regions with average past SLR.

This result is interesting because in theory, our measure of exposure reflects the effect of a 6-foot global sea level rise, accounting for the accompanying regional differences in SLR. Nevertheless, this could be a rational response if bond prices are responding to the low probability of catastrophic loss, since such losses are measured with a high degree of uncertainty due to the unpredictability of SLR

[5] According to the NOAA's Regional Sea Levels and Future Scenarios maps, the extent of recent SLR varies considerably across states, with SLR on the East and Gulf coasts ranging from 1.70 mm/year (in Massachusetts and Maine) to 3.89 mm/year (in New Jersey, New York, Connecticut, Rhode Island, Maryland, and Virginia) and SLR on the West Coast ranging from 1.28 mm/year in Oregon and Washington to 1.73 mm/year in California.

exposure in more inland regions, which will disproportionately impact areas with faster rising seas.[6]

In Columns 3 and 4, we investigate whether regional beliefs regarding the impact of climate change affect the manner in which municipal bond markets price SLR exposure. Bernstein, Gustafson, and Lewis (2018) and Baldauf, Garlappi, and Yannelis (2018) both find evidence of climate changes beliefs affecting how real estate markets price SLR exposure. It is reasonable to expect that local beliefs will also matter for municipal bond pricing because buyers are often local retail investors, in part because the tax advantages to municipal bond purchases are often largest for local clientele. To measure an area's beliefs about climate change we merge our data with the Yale Climate Opinions map data (Howe et al. (2015)). Specifically, we aggregate 2014 county level survey data on the respondents' answer' to the question "worried about global warming" to the state-level (on an equal-weighted basis by school district). To form our State Worry measure, we then subtract the average state's level of worry and divide by the standard deviation, resulting in a standardized measure that ranges from -2.68 to 0.79.

These columns augment the standard specification from equation 1 by adding the Post $\times$ Pct. Exposed $\times$ State Worry triple interaction (along with interactions between State Worry and both Post and Pct. Exposed). The positive and statistically significant triple interaction in Column 4 indicates that on the East Coast and Gulf there is a positive relation between how municipal bond markets price SLR exposure and the reported level of concern over global warming in the state. However, we do not find the same relation in the West Coast, which makes the triple interaction in the full sample statistically insignificant.

We next examine the empirical question of whether the pricing of SLR exposure risk depends on the bond's maturity. If market participants view SLR exposure risk as a long-term risk, the long maturity bonds may be more affected. On the other hand, since municipal bonds are largely supported by taxes on the value of local property, which should incorporate expectations of future outcomes, even short maturity bonds may be affected. For example, evidence in Bernstein, Gustafson, and Lewis (2018) that SLR exposed coastal real estate already trades at a 7% discount relative to observably similar unexposed properties suggests that it may become increasingly difficult to role over short-term debt that is based on the value of coastal economies.

Figure 2 illustrates the relation between bond maturity and the extent to which the municipal bond market appears to price SLR exposure. The figure is constructed in the same manner as Figure 1, except that each regression is run only on the subsample of bonds with remaining maturity of less

[6] See the discussion in Bernstein, Gustafson, and Lewis (2018) regarding the imprecision with with SLR exposure is measured, especially in areas farther from the coast.

than 10 years (dashed line) or greater than 10 years (solid line). For legibility, we do not report standard error bands in the figure, but as in Figure 1, few of the observations prior to 2012 are statistically significant. Thus, we find no evidence that the municipal bond markets were pricing SLR exposure risk for either short or long maturity bonds prior to 2012.

The solid line, which plots the Pct. Exposed by year for the sample of bonds with over ten years in remaining maturity spikes in 2012 to a statistically significant estimate of approximately 0.15. It then increases in the following year to approximately 0.25 and continues to rise over the remainder of our sample to over 0.3. The dashed line indicates that the market did not start pricing SLR exposure risk in the short-maturity municipal bond market until several years later, in 2015. Since 2015, the short maturity bond market has priced SLR exposure risk similarly to the long-run bond market.

Finally, in Table 4 we both corroborate the evidence from Figure 2 in a regression framework and examine the potential channels for our effect. In particular, Bernstein, Gustafson, and Lewis (2018) shows a negative house price associated with SLR risk within a similar set of geographic regions and following a similar time horizon as the effect we document here. Perhaps the higher yields on municipal bonds simply reflect this the degradation of the underlying assets that form an exposed communities taxable base (e.g. houses). Column 1 specifically addresses this concern by including a local house price index. We do find a negative coefficient on district house prices (e.g. an increase in house prices results in a decrease in bond yields) but this effect is completely orthogonal to the relation between SLR risk and yields. The coefficient on Post $\times$ Pct. Exposed remains at 0.24 in the specifications with or without house prices. This finding is consistent with a rational channel where bond market participants are responding to a shock to expected volatility of the underlying housing stock associated with SLR risk. It is also consistent with a model of segmented markets where municipal bond holders price SLR risk without considering the immediate house price movements.

While our regressions control for a suite of potential drivers, additional concerns remain that unobservable district level effects are responsible for the pricing results. However, given the long term nature of SLR risk, we should also expect a differential effect between long and short maturity bonds in response thus allowing us to more directly control for district level unobservables In Column 2, we perform the same setup as in Table 1 but include a triple interaction between Post $\times$ Pct. Exposed $\times$ Time to Maturity. We show that the increased yields for SLR exposed districts is larger for securities with longer maturities. However, this specification allows us to further mitigate concerns about unobservable drivers of municipal yields by including district by month fixed effects (this absorbs Post $\times$ Pct. Exposed). In Column 3, we find a significant and positive coefficient on the triple interaction term Post $\times$ Pct. Exposed $\times$ Time to Maturity indicating that longer term bonds are more sensitive to the increasingly pessimistic SLR forecasts for exposed districts.

4 Conclusion

Many argue that climate change is one of the biggest threats facing the world today. Yet, it is often challenging to quantify climate change risks, in part because they will manifest over centuries. Financial markets offer a unique opportunity to overcome this challenge for cases in which climate change risks affect current asset prices, which incorporate investors' future expectations.

We use the municipal bond market to examine the extent to which financial markets view increasing sea level rise projections as a material risk to U.S. coastal communities. Sea level rise projections have risen substantially over the past decade, culminating with a 2013 IPCC report that doubled end-of-century estimates. Currently over $1 trillion of coastal real estate is at risk over the coming century.

Prior to 2013, we find no significant relation between sea level rise and municipal bond yields. After the 2013 IPCC doubling of sea level rise projections, we find consistent that sea level rise exposure has become a statistically significant predictor of municipal bond yields in coastal communities. The importance of sea level rise exposure in determining bond yields is largest on the east coast (where seas have been are expected to continue rising faster) and in states that report high levels of belief in climate change.

Despite consistent evidence that municipal bond markets now price sea level rise exposure, the economic importance of sea level rise exposure as an input into bond prices remains small. A 10-percentage point (or approximately 1.1 standard deviation) increase in the number of SLR exposed properties within a school district is accompanied by a 2.4 basis point increase in municipal bond yields, equivalent to 0.5% of the average yield. Thus, municipal bond investors appear to believe that there is a small probability of a sea level rise related catastrophe in the coming decades. Whether this is due to the maturity of municipal bonds being measured in decades instead of centuries or because the market expects coastal communities to take on successful remediation efforts is an important question for future research.

References

Baldauf, Markus, Lorenzo Garlappi, and Constantine Yannelis. 2018. “Does Climate Change Affect Real Estate Prices? Only If You Believe in it.” *Review of Financial Studies Forthcoming* .

Bernstein, Asaf, Matthew Gustafson, and Ryan Lewis. 2018. “Disaster on the horizon: The price effect of sea level rise.” *Journal of Financal Economics Forthcoming* .

DeConto, Robert M and David Pollard. 2016. “Contribution of Antarctica to past and future sea-level rise.” *Nature* 531 (7596):591.

Grinsted, Aslak, Svetlana Jevrejeva, Riccardo EM Riva, and Dorthe Dahl-Jensen. 2015. “Sea level rise projections for northern Europe under RCP8. 5.” *Climate Research* 64 (1):15–23.

Hallegatte, Stephane, Colin Green, Robert J Nicholls, and Jan Corfee-Morlot. 2013. “Future flood losses in major coastal cities.” *Nature climate change* 3 (9):802.

Hauer, Mathew E, Jason M Evans, and Deepak R Mishra. 2016. “Millions projected to be at risk from sea-level rise in the continental United States.” *Nature Climate Change* 6 (7):691.

Hinkel, Jochen, Carlo Jaeger, Robert J Nicholls, Jason Lowe, Ortwin Renn, and Shi Peijun. 2015. “Sea-level rise scenarios and coastal risk management.” *Nature Climate Change* 5 (3):188.

Howe, Peter D, Matto Mildenberger, Jennifer R Marlon, and Anthony Leiserowitz. 2015. “Geographic variation in opinions on climate change at state and local scales in the USA.” *Nature Climate Change* 5 (6):596–603.

Marcy, Doug, Nate Herold, Kirk Waters, William Brooks, Brian Hadley, Matt Pendleton, Keil Schmid, Mike Sutherland, Kyle Dragonov, John McCombs, and Sean Ryan. 2011. “New Mapping Tool and Techniques for Visualizing Sea Level Rise and Coastal Flooding Impacts.” *In Proceedings of the 2011 Solutions to Coastal Disasters Conference, Anchorage, Alaska* :474–490.

Painter, Marcus. 2018. “An Inconvenient Cost: The Effects of Climate Change on Municipal Bonds.” *Journal of Financial Economics Forthcoming* .

Parris, Adam S, Peter Bromirski, Virginia Burkett, Daniel R Cayan, Mary Evans Culver, John Hall, Radley M Horton, Kevin Knuuti, Richard H Moss, Jayantha Obeysekera et al. 2012. “Global sea level rise scenarios for the United States National Climate Assessment.” .

Piecuch, Christopher G, Peter Huybers, Carling C Hay, Andrew C Kemp, Christopher M Little, Jerry X Mitrovica, Rui M Ponte, and Martin P Tingley. 2018. “Origin of spatial variation in US East Coast sea-level trends during 1900–2017.” *Nature* 564 (7736):400.

Rao, Krishna. 2017. *Climate Change and Housing: Will a Rising Tide Sink all Homes?* URL https://www.zillow.com/research/climate-change-underwater-homes-12890/.

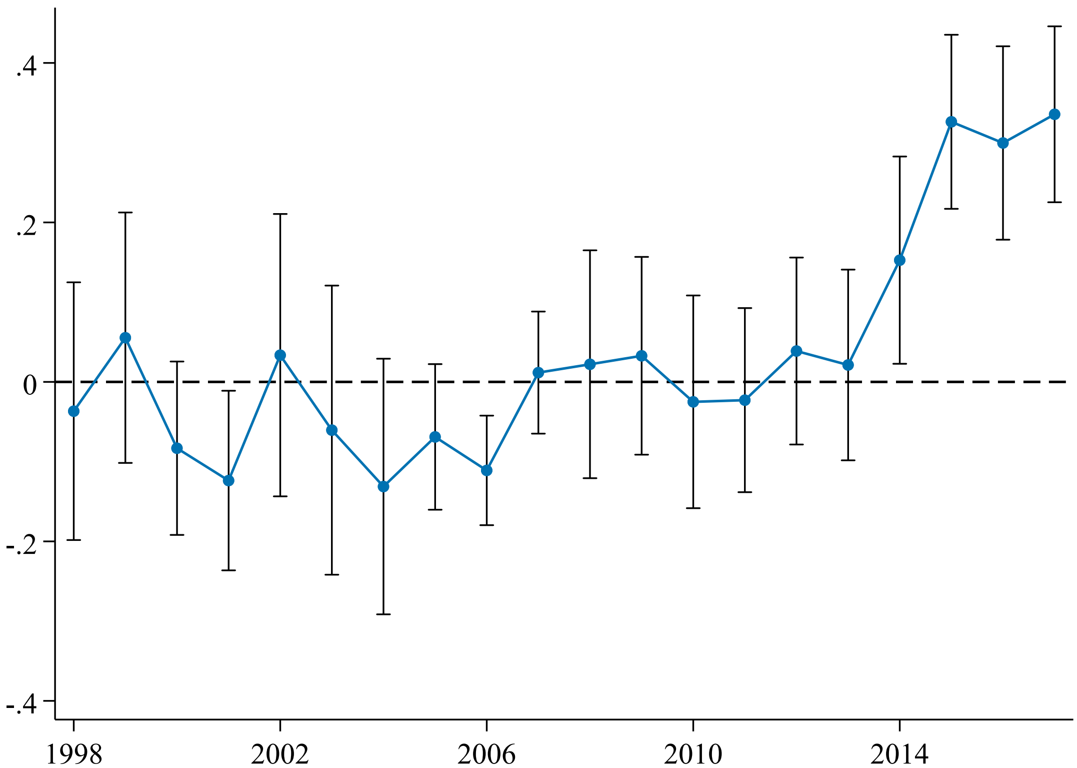
Rohling, Eelco J, Ivan D Haigh, Gavin L Foster, Andrew P Roberts, and Katharine M Grant. 2013. “A geological perspective on potential future sea-level rise.” *Scientific reports* 3:3461.

Schwert, Michael. 2017. “Municipal bond liquidity and default risk.” *The Journal of Finance* 72 (4):1683–1722.

Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P.M. Midgley, editors. 2013. *Summary for Policymakers*, book section SPM. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press, 1–30. URL www.climatechange2013.org.

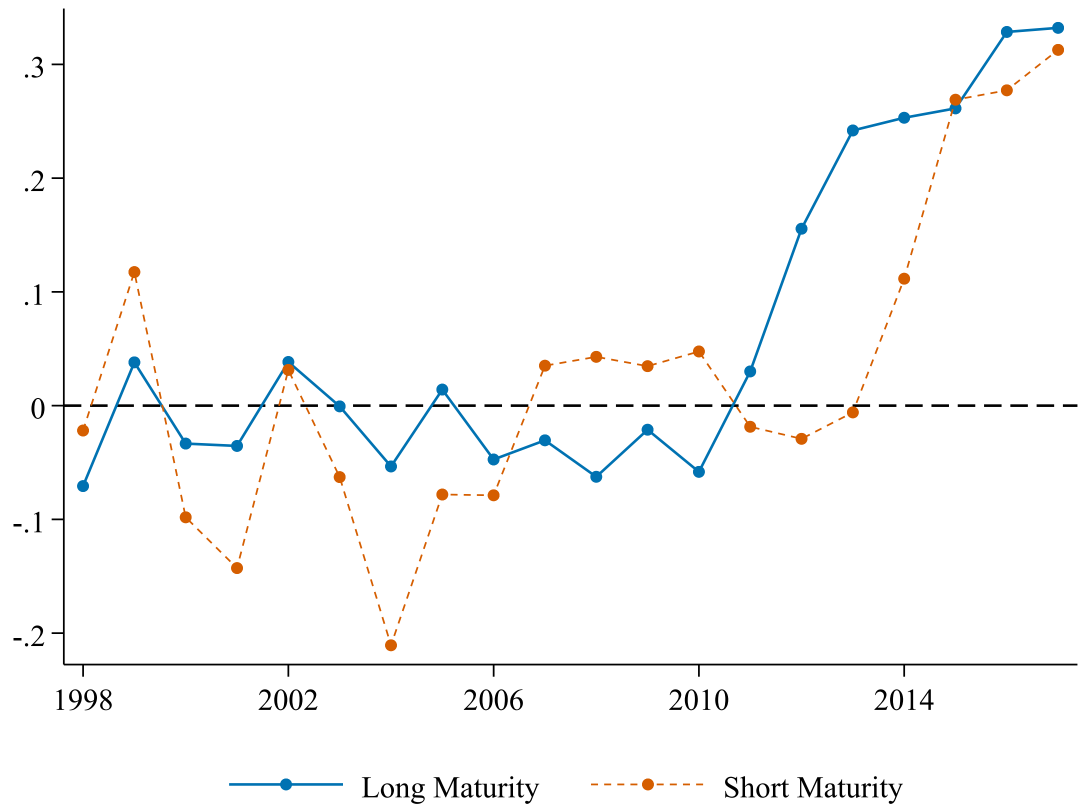
Sweet, William V, Robert E Kopp, Christopher P Weaver, Jayantha Obeysekera, Radley M Horton, E Robert Thieler, and Chris Zervas. 2017. “Global and regional sea level rise scenarios for the United States.” .

Figure 1: Effect of Sea Level Rise on Bond Yields



Note: This figure plots the year-by-year effect of Pct. Exposed on municipal bond yields. The coefficients come from a regression identical to that in Equation 1, except instead of including a Post $\times$ Pct. Exposed interaction, we restrict the sample to only bonds traded during the year indicated on the x-axis. The vertical bars reflect the 95% confidence intervals, where standard errors are clustered at the county-level.

Figure 2: Heterogeneous Effects of Sea Level Rise on Bond Yields by Maturity



Note: This figure plots the year-by-year effect of Pct. Exposed on municipal bond yields for long and short-maturity bonds. The coefficients come from a regression identical to that in Equation 1, except instead of including a Post × Pct. Exposed interaction, we restrict the sample to only bonds traded during the year indicated on the x-axis. Long maturity is defined as greater than 10 years, and is denoted by the solid line. Short maturity is denoted by the dashed line. The vertical bars reflect the 95% confidence intervals, where standard errors are clustered at the county-level.

Table 1: Summary Statistics

	Full Coastal Sample			*SLR Exposed Districts*		
	Mean	Std.Dev.	Obs	Mean	Std.Dev.	Obs
Percent of Properties Exposed (3ft GASLR)	1.08	3.26	544,071	2.47	4.56	238,360
Percent of Properties Exposed (6ft GASLR)	3.18	9.12	544,071	7.26	12.65	238,360
Average Price, Single Family Home ($000s)	404.91	447.44	544,071	442.60	564.11	238,360
Bond Yield to Worst1	3.38	1.25	544,071	3.36	1.25	238,360
Time to Maturity	10.10	6.35	544,071	9.92	6.21	238,360
Bond Age	3.90	2.69	544,071	3.89	2.66	238,360
Monthly Trading Volume ($mm)	509.26	3138.31	544,071	624.75	4157.25	238,360
Monthly SD of Price	0.91	0.72	480,004	0.92	0.72	211,061

Note: This table reports the summary statistics for the main variables used in this paper. Observations are at the district by month frequency. Sea Level Rise exposures are measured within school districts and represent the percent of residential properties that would be inundated after a X feet of global average sea level rise.

Table 2: Effect of Sea Level Rise on Bond Yields

	Full Sample		East	Gulf	West
	(1)	(2)	(3)	(4)	(5)
Pct. Exposed	-0.015	0.029	0.022	0.123	-0.156
	(-0.31)	(0.40)	(0.46)	(0.60)	(-1.51)
Post X Pct. Exposed	0.240**	0.240**	0.438***	0.037	-0.039
	(2.34)	(2.41)	(4.15)	(0.20)	(-0.26)
Time to Maturity	0.129***	0.146***	0.135***	0.121***	0.129***
	(41.64)	(48.67)	(34.11)	(35.91)	(31.95)
Log(Average Value)	-0.022**	-0.028**	-0.015*	0.036	-0.074**
	(-2.29)	(-2.24)	(-1.93)	(1.01)	(-2.55)
Coupon	0.066***	0.110***	-0.031***	0.215***	0.078***
	(6.82)	(10.56)	(-3.31)	(9.82)	(6.96)
Bond Age	0.091***	0.104***	0.088***	0.112***	0.089***
	(18.03)	(18.13)	(17.88)	(19.20)	(10.82)
Log(Monthly Volume)	-0.035***	-0.045***	-0.037***	-0.038***	-0.030***
	(-10.87)	(-11.96)	(-13.93)	(-11.76)	(-6.29)
Monthly SD of Price	0.094***	0.114***	0.094***	0.075***	0.097***
	(21.65)	(22.74)	(17.30)	(8.84)	(18.24)
Constant	2.020***	1.481***	2.372***	0.603	2.628***
	(14.48)	(8.01)	(23.58)	(1.34)	(6.40)
Adj. R-squared	0.802	0.800	0.832	0.791	0.789
Adj. Within R-squared	0.683	0.729	0.683	0.692	0.693
Outcome Mean	3.391	3.080	3.282	3.447	3.466
Outcome SD	1.247	1.268	1.220	1.268	1.254
Outcome Within-SD	0.960	1.068	0.853	1.016	1.024
Observations	474,363	336,241	182,154	93,683	198,437

Note: This table reports the effect of Pct. Exposed on municipal bond yields in the pre- and post-2013 period. In Column 1, we estimate Equation 1 over our full sample from 1997 to 2017. In Column 2, we estimate the same equation over a shorter 10-year sample concentrated around 2013. In Column 3, we limit the sample to school districts on the East Coast, Column 4, the Gulf Coast and Column 5 the West Coast. Pct. Exposed measures the share of homes that would be inundated with a six foot SLR.

Table 3: Heterogeneity in Effect of Sea Level Rise on Bond Yields by Sentiment

	(1)	(2)	(3)	(4)
Pct. Exposed	-0.077	-0.001	-0.024	0.010
	(-1.25)	(-0.01)	(-0.49)	(0.20)
Post X Pct. Exposed	0.106	-0.185	0.226**	0.428***
	(1.19)	(-0.87)	(2.16)	(3.86)
Post X Pct. Exposed X State SLR	0.236***	0.488***		
	(3.01)	(2.86)		
Post X Pct. Exposed X State Worry			0.138	0.363***
			(1.35)	(2.90)
Time to Maturity	0.129***	0.135***	0.129***	0.135***
	(41.60)	(34.14)	(41.62)	(34.10)
Coupon	0.067***	-0.031***	0.066***	-0.031***
	(6.84)	(-3.34)	(6.81)	(-3.29)
Bond Age	0.091***	0.088***	0.091***	0.088***
	(18.05)	(17.95)	(18.05)	(17.92)
Log(Monthly Volume)	-0.035***	-0.037***	-0.035***	-0.037***
	(-10.96)	(-13.98)	(-10.93)	(-13.99)
Monthly SD of Price	0.094***	0.094***	0.094***	0.094***
	(21.66)	(17.31)	(21.65)	(17.30)
Constant	1.760***	2.263***	1.756***	2.198***
	(25.13)	(31.50)	(24.98)	(35.87)
Sample	All	East Coast	All	East Coast
Adj. R-squared	0.802	0.832	0.802	0.832
Adj. Within R-squared	0.683	0.683	0.683	0.683
Outcome Mean	3.391	3.282	3.391	3.282
Outcome SD	1.247	1.220	1.247	1.220
Outcome Within-SD	0.960	0.853	0.960	0.853
Observations	474,363	182,154	474,363	182,154

Note: This table explores heterogeneity in the effect of Pct. Exposed on municipal bond yields in the pre- and post-2013 period. In Column 1 and 3, we use the full sample. In Column 2 and 4, e limit the sample to school districts on the East Coast. Pct. Exposed measures the share of homes that would be inundated with a six foot SLR.

Table 4: The Effect of Sea Level Rise on Bond Yields: Identification and Channel

	(1)	(2)	(3)
Post X Pct. Exposed	0.240**	-0.175	
	(2.34)	(-1.13)	
Pct. Exposed X Time to Maturity		0.004	0.016
		(0.52)	(0.65)
Post X Time to Maturity		0.007	0.007
		(1.59)	(1.53)
Post X Pct. Exposed X Time to Maturity		0.043***	0.034**
		(2.76)	(2.01)
Log(Average Value)	-0.022**		
	(-2.29)		
Constant	2.020***	1.742***	1.660***
	(14.48)	(25.81)	(24.51)
Controls	Y	Y	Y
FIPS X Month FE	Y	Y	N
District X Month FE	N	N	Y
Adj. R-squared	0.802	0.802	0.824
Adj. Within R-squared	0.683	0.684	0.696
Outcome Mean	3.391	3.391	3.396
Outcome SD	1.247	1.247	1.241
Outcome Within-SD	0.960	0.960	0.846
Observations	474,363	474,363	430,517

Note: This table explores the role of bond maturity on the effect of Pct. Exposed on municipal bond yields in the pre- and post-2013 period. In Column 1 and 2, we use the same set of county-by-month fixed effects and controls as in Table 2 and 3. In Column 3, we use school-district-by-month fixed effects. Pct. Exposed measures the share of homes that would be inundated with a six foot SLR.

Pricing Poseidon: Extreme Weather Uncertainty and Firm Return Dynamics*

Mathias S. Kruttli, Brigitte Roth Tran, and Sumudu W. Watugala†

May 2019

Abstract

We investigate the uncertainty dynamics surrounding extreme weather events, particularly hurricanes, through the lens of stock and option markets. We combine firm establishment-level data with novel hurricane forecast and damage data to identify market responses to both the uncertainty regarding potential landfall and subsequent economic impact. In the days following landfall, stock options on firms exposed to the landfall region exhibit increases in implied volatility of up to 8 percent, reflecting the impact uncertainty. Using hurricane forecasts, we show that landfall uncertainty and potential impact uncertainty are reflected in prices before landfall, consistent with investors paying attention to the forecasts. We find no evidence that markets incorporate better hurricane forecasts than those from the National Oceanic and Atmospheric Administration. Improvements to hurricane forecasts could have economically significant effects in financial markets.

JEL classification: G12, G14, Q54.

Keywords: extreme weather events, uncertainty, implied volatility, stock returns, hurricanes, climate change.

*We thank Kerry Emanuel, Kristine Watson Hankins, Andrew Karolyi, Fang Liu, David Ng, Neil Pearson, Brian Seok, Aurelio Vasquez, Scott Yonker, Youngsuk Yook, and seminar participants at the Federal Reserve Board, NOAA, Cornell University, and the Risk Management and Financial Innovation Conference in Memory of Peter Christoffersen for helpful comments. Keely Adjorlolo, David Rubio, and Alan Yan provided outstanding research assistance. The analysis and conclusions set forth are those of the authors and do not indicate concurrence by the Board of Governors of the Federal Reserve System or its research staff.

†Kruttli: The Board of Governors of the Federal Reserve System. Email: mathias.s.kruttli@frb.gov. Roth Tran: The Board of Governors of the Federal Reserve System. Email: brigitte.rothtran@frb.gov. Watugala: Cornell University. Email: sumudu@cornell.edu.

1 Introduction

Extreme weather events can lead to substantial damages and devastate economic regions. In 2017, the total damages of extreme weather events in the United States were over $300 billion.[1] Although significant work has been done to understand how extreme weather events affect real economic activity and decision making by households, firms, and financial institutions, little is known about the uncertainty surrounding extreme weather events.[2] Given that uncertainty affects real economic activity and is important in the context of economic agents' decision making (see, for example, Bernanke (1983); Bloom, Bond, and van Reenen (2007); Bloom (2009)), a comprehensive assessment of the economic effects of extreme weather events requires understanding the uncertainty dynamics surrounding them.

This paper examines extreme weather uncertainty through the lens of financial markets. Asset prices are an ideal instrument to assess the dynamics and magnitude of extreme weather uncertainty because of the frequency and scale at which financial data are available and the fact that the investor behavior underlying the asset prices is guided by financial incentives. We focus our analysis on hurricanes, arguably one of the most devastating types of extreme weather events.[3] In our framework, we distinguish between two components of extreme weather uncertainty: (a) the "landfall uncertainty" that exists about where and when a hurricane will make landfall, if at all, and (b) the "impact uncertainty" about what effect a hurricane will have on the local economy and firms conditional on it making landfall.

Using implied volatility measures of single stock options as a proxy for uncertainty, we find that immediately after a hurricane has made landfall, option prices reflect substantial impact uncertainty through higher implied volatilities for firms in the affected region. Further, our results based on novel data on hurricane forecasts issued by the National Oceanographic and Atmospheric Administration (NOAA) show that the impact and landfall uncertainty are priced before the hurricane makes landfall, in line with investors paying attention to the forecasted path of the hurricane.

[1] The damage estimate is from the National Oceanic and Atmospheric Administration and can be found here: https://www.climate.gov/news-features/blogs/beyond-data/2017-us-billion-dollar-weather-and-climate-disasters-historic-year.

[2] For research on the real effects and decision making of extreme weather events, see, for example, Belasen and Polachek (2008); Imberman, Kugler, and Sacerdote (2012); Barrot and Sauvagnat (2016); Bernile, Bhagwat, and Rau (2017); Dessaint and Matray (2017); Brown, Gustafson, and Ivanov (2017).

[3] Of the $300 billion in total damages in the US from extreme weather events in 2017, approximately $265 billion were caused by hurricanes.

The combination of impact and landfall uncertainty can cause higher levels of uncertainty before landfall than immediately after landfall.

While our framework can also be applied to other extreme weather events,[4] we focus on hurricanes because they develop and resolve over fairly short but well-defined time frames, thereby isolating the associated uncertainty in time and location to allow for better identification, NOAA publishes a range of data that allow us to precisely estimate the uncertainty surrounding hurricanes, and finally they are likely to garner investor attention due to significant impacts and extensive media coverage.

We combine firm establishment and sales data at the county level with hurricane forecast and damage data in order to identify firms that operate within regions (potentially) exposed to a particular hurricane. We use these data to test two hypotheses regarding how financial markets price a hurricane's landfall and impact uncertainty.

Our first hypothesis is that immediately after a hurricane has made landfall, implied volatilities of options of firms in the disaster region are elevated due to impact uncertainty. Implied volatility is a proxy for uncertainty (see, for example, Bloom (2009) and Kelly, Pastor, and Veronesi (2016)), and heightened implied volatility is consistent with investors hedging their exposure to weather event impacts by buying options on the firms with a physical presence in the affected region. Our results support this hypothesis. Indicative of substantial impact uncertainty, we find that immediately after hurricane landfall the implied volatility of options of firms in the disaster region are 8 percent higher than before the hurricane's inception. Importantly, the average firm that experiences damage from hurricanes in our sample is similar in size to the average firm that does not, indicating that our results are based on a representative sample of firms.[5]

Further, we examine the post-landfall stock price reactions of firms with a physical presence in a hurricane disaster region. While both the short- and long-term mean abnormal return relative to the Fama-French three factor model is generally not significantly different between firms in disaster regions and control firms, the abnormal return differences are substantial for the weakest performing

[4]Most extreme weather events like snow storms and severe rain involve uncertainty about when and where they will occur ex ante as well as uncertainty in the immediate period after they have occurred about their full eventual impact.

[5]The market capitalization of firms with at least 20% of its establishments in the hurricane disaster region is \$6.0 billion on average, while the market capitalization of firms with less than 20% of its establishments in the disaster region is \$6.6 billion on average.

stocks in the long-term. Over the 120 trading days after hurricane landfall, the 10^{th} percentile of the abnormal return distribution for firms in disaster regions is 12-15 percentage points lower than the 10^{th} percentile of the abnormal return distribution of the control firms. Instead of disaster region stocks being affected uniformly, a subset of affected firms underperform substantially. Our results are consistent with investors learning which firms are in this subset as impact uncertainty resolves over time.

Our second hypothesis is that investors pay attention to hurricane forecasts before landfall and demand compensation for the landfall uncertainty and the potential impact uncertainty. This hypothesis implies that hurricane forecasts contain valuable information for investors and, if financial markets are efficient, this information should be reflected in asset prices. Using NOAA forecasts issued in the days or weeks leading up to a hurricane's landfall or dissipation (in the case of a hurricane that "missed") to measure landfall uncertainty, we find that implied volatility and stock returns react as predicted even at low landfall probabilities of 10 percent and below, indicating that investors pay attention to hurricane forecasts. Further, the combined landfall uncertainty and potential impact uncertainty cause the implied volatilities to be higher before landfall than shortly after landfall.

As a natural extension to the second hypothesis, we analyze whether financial markets can forecast hurricanes better than NOAA and estimate the effects of improving NOAA's forecasts, which arguably are the most prominent public forecasts heavily used by media outlets. Financial markets might reflect better (private) predictions about which firms will be hit by a hurricane if large institutional investors, like hedge funds that often act as the marginal investor in asset markets and affect asset prices, are able to outperform weather agencies when forecasting extreme weather events. There is reason to believe that sophisticated institutional investors can forecast hurricanes better than NOAA. First, the annual budget of the National Weather Service, which is the division of NOAA responsible for hurricane forecasts, is orders of magnitude smaller than the value of assets managed by large institutional investors.[6] Second, there is anecdotal evidence that hedge funds obtain information on hurricane forecasts from sources other than NOAA.[7] We test this hypothesis

[6]The total budget of the National Weather Service, a subdivision of NOAA, was around $1 billion in 2017. Only a fraction of this budget was used for hurricane forecasts as the National Weather Service is also tasked with providing weather forecasts besides hurricane forecasts. The budget of the National Weather Service for 2017 can be found here: https://www.corporateservices.noaa.gov/nbo/.

[7]See, for example, the discussion of the hedge fund with the name Nephila by Michael Lewis here: https://www.

by estimating if firms that are *not* in the ex ante forecasted path of a hurricane but end up in the disaster region also see increases in the implied volatility of their options. Despite reasonable statistical power, our estimates fail to reject the null hypothesis that markets do not reflect superior information to NOAA forecasts on hurricanes even when accounting for the institutional ownership of the underlying stock.

We compute the potential effects on option markets of improving NOAA forecast accuracy. We estimate the additional change in implied volatility due to forecast errors for firms for which the forecasted exposure was larger (smaller) than the eventual exposure to the disaster region. We find large average effects of up to 200 basis points for thousands of firms over the sample period from 2007 to 2017. This result speaks to the outsized importance of NOAA's hurricane forecasts for financial markets and is valuable information for legislators who make budgetary decisions.

While this paper focuses on how extreme weather uncertainty affects the broad universe of US public firms, we also conduct a separate analysis on insurance firms. The sample of public property and casualty insurance firms with liquid options and stocks is fairly small, and we only have data on their exposure at the state level and not at the county level. However, we do find that single stock options of property and casualty insurance firms price in substantial impact uncertainty immediately following a hurricane landfall.

Our findings on the dynamics and magnitude of extreme weather uncertainty as measured through financial markets have several important implications. First, given that research has shown that other types of uncertainty can affect a firm's decision making[8], the large economic magnitudes of our estimated responses suggest that extreme weather uncertainty is an important factor for real economic activity and the decision making of economic agents. Second, our uncertainty estimates imply that extreme weather events impose significant financial costs that have generally not been accounted for. Uncertainty regarding the impact of an extreme weather event on a firm increases investor costs of trading and hedging investments related to that firm. These costs are an important component of accounting for the toll these events take when policymakers determine how to respond to threats of more frequent and severe extreme weather events (see, for example, Melillo, Richmond, and Yohe (2014)). Third, our results show that investors are attentive to firm exposure to hurricanes

nytimes.com/2007/08/26/magazine/26neworleans-t.html?pagewanted=all.

[8]For example, Bernanke (1983), Bloom, Bond, and van Reenen (2007), and Julio and Yook (2012) show how uncertainty can reduce firm investments.

even before landfall. Investor attention to extreme weather risk is important for correctly pricing assets with extreme weather and climate change exposure and reduces the risks of sudden large price corrections that could disrupt financial stability (see, for example, Carney (2015)). Fourth, knowing how investors perceive extreme weather uncertainty empirically can help the modeling of uncertainty for economic agents in integrated assessment models that try to assess the social costs of climate change (see, for example, Nordhaus and Yang (1996); Tol (1997); Nordhaus (2014)).[9]

Related literature. Our paper ties in to several diverse bodies of literature. By analyzing extreme weather uncertainty, our paper contributes to the uncertainty literature, in which several papers have focused on economic policy uncertainty and its effects on firms (see, for example, Bloom, Bond, and van Reenen (2007) and Bloom (2009)). Other researchers have focused on political uncertainty proxied by elections and how they affect firm investments and financial markets (see, for example, Julio and Yook (2012); Kelly, Pastor, and Veronesi (2016); Jens (2017)). Our paper complements this body of work by showing that extreme weather uncertainty is a different but important source of uncertainty that affects prices in financial markets. Moreover, in the case of elections, there is uncertainty about outcomes, but generally not about when and whether the elections themselves will occur because they are scheduled in advance.[10] Our analysis introduces an additional layer of complexity as we separately examine the effects of the uncertainty about the landfall of a hurricane and the uncertainty about the impact of the event itself. Our paper differs from the research on macroeconomic uncertainty and economic growth (see, for example, Jurado, Ludvigson, and Ng (2015); Baker, Bloom, and Davis (2016); Baker, Bloom, and Terry (2018)) in that our analysis is at the firm level and more granular than the macroeconomy as a whole, which is important as extreme weather events are generally local phenomena. Also, the uncertainty shock in our case, the hurricane, is exactly determined.

Further, by showing that extreme weather events cause substantial uncertainty in economic regions before and after landfall, our work proposes an additional factor that should be considered by the literature that examines extreme weather events' real effects and their impact on economic

[9]See Gillingham, Nordhaus, Anthoff, Blanford, Bosetti, Christensen, McJeon, and Reilly (2018) for a detailed summary of uncertainty in integrated assessment models for climate change.

[10]Empirical work on political uncertainty focuses on scheduled elections in order to isolate political uncertainty from economic uncertainty. Unscheduled elections and regime changes can be precipitated by economic conditions. In contrast, hurricanes are exogenous to economic uncertainty (economic conditions do not make hurricanes more likely), so we do not face this identification issue.

agents' decision making. This growing literature includes work that examines the effects of extreme weather on labor markets and schooling (see Belasen and Polachek (2008) and Imberman, Kugler, and Sacerdote (2012)). Barrot and Sauvagnat (2016) find that shocks of extreme weather events propagate in customer-supplier firm networks. Bernile, Bhagwat, and Rau (2017) analyze the relationship between risk taking behavior and the early-life disaster experiences of CEOs. Dessaint and Matray (2017) show that managers overreact to hurricane risks after experiencing a hurricane. Brown, Gustafson, and Ivanov (2017) report that firms experience decreased cash flows after extreme snowfall events and that they respond by increasing their use of credit lines. Looking at storm-level total damages, Martinez (2018) finds that damages increase with forecast error of landfall location 12 hours before landfall. Roth Tran and Wilson (2018) find that hurricanes have a wide range of impacts on local economic activity, including on employment, population, and home prices.

Finally, this paper introduces a novel topic to an emerging literature on climate finance that includes early empirical work on how Florida temperature fluctuations affect orange juice futures prices (see Roll (1984) and Boudoukh, Richardson, Shen, and Whitelaw (2007)) and how the use of a time series forecasting approach is useful for pricing weather derivatives (see Campbell and Diebold (2005)). Our research contributes to three branches of the climate finance literature.

First, by examining hurricane effects, this paper builds on recent papers in the finance literature focused on extreme weather events and investor attention. Hong, Li, and Xu (2016) show that drought indices are predictive of food company stock returns, indicating that investors are inattentive to droughts' impacts on food companies. Choi, Gao, and Jiang (2018) examine how investors' climate change beliefs when temperatures are warmer than usual and find evidence of a positive relationship. Addoum, Ng, and Ortiz-Bobea (2019) show that firms' earnings are affected by high temperatures and that analysts and investors do not immediately react to temperature shocks.

Second, our paper adds to climate finance papers that develop hedging strategies. While Baker, Hollifield, and Osambela (2018) and Roth Tran (2018) present theoretical models in which green or emission-oriented investors can hedge risks by investing in polluters, Andersson, Bolton, and Samama (2016) show empirically that investors can hedge against potential future prices on carbon emissions by investing in a decarbonized index. Engle, Giglio, Kelly, Lee, and Stroebel (2018) develop a climate change news index and assess strategies that can hedge an investor against such news. In contrast to these papers, we focus on market dynamics that reflect investor behavior around specific disaster events that occur at a local level.

Third, by using daily hurricane forecasts from NOAA, this paper adds to recent climate finance research that analyzes how NOAA forecasts are reflected in asset prices. Drawing mixed conclusions, several papers (see Bernstein, Gustafson, and Lewis (2018); Giglio, Maggiori, Rao, Stroebel, and Weber (2018); Murfin and Spiegel (2018)) use NOAA sea level rise predictions to examine whether residential real estate prices reflect sea level rise risks. Our use of NOAA forecasts is substantially different, because for hurricanes we can observe multiple isolated events from inception to resolution, whereas NOAA's forecasts for sea level rise are long-term and cannot yet be compared to realizations. We are thus able to show not only that price reactions in options and stocks are consistent with investors paying attention to NOAA's hurricane forecasts, but also assess whether this attention is in line with the realized outcomes.

The remainder of this paper is structured as follows. In Section 2, we describe the empirical design. Section 3 describes the datasets we use. We present our main results in Section 4 and conclude in Section 5.

2 Empirical design

2.1 Landfall uncertainty and impact uncertainty

Our framework distinguishes between two types of uncertainty that surround a hurricane: impact uncertainty and landfall uncertainty. While this paper focuses on hurricanes as an example for an extreme weather event due to the availability of high quality data, the framework of landfall and impact uncertainty can also be applied to other types of extreme weather events. The impact uncertainty is the uncertainty about a hurricane's impact on firms with exposure to the landfall area. More formally, if hurricane h is expected to make landfall at time $t+1$ and an all-equity firm i's stock return at $t+1$ is given by

$$r_{i,t+1} = \epsilon_{i,t+1} + \theta_{h,t+1} g_{i,h,t+1}, \tag{1}$$

where $\epsilon \sim N(0, \sigma^2)$ represents a random shock to the firm's return at time $t+1$ with a mean of zero and variance of σ^2. The random variable $g_{i,h,t+1} \sim N(\mu_g, \sigma_g^2)$ is independent of ϵ and captures the impact of the hurricane on the value of firm i, conditional on hurricane landfall in the firm's geographic region. The random variable θ captures whether or not the firm is hit by the hurricane and has a Bernoulli distribution or, equivalently, a binomial distribution with one draw, $\theta \sim B(1, \phi)$, where $Pr(\theta = 1) = 1 - Pr(\theta = 0) = \phi$ and $0 \leq \phi \leq 1$. The product of the two random variables, $\theta_{h,t+1} g_{i,h,t+1}$, is the component of the return attributable to the hurricane.

Conditional on hurricane landfall at time $t+1$, σ_g^2 represents the *impact uncertainty*. Defining uncertainty as the variance of an unpredictable disturbance is in line with Pastor and Veronesi (2012 and 2013) and Jurado, Ludvigson, and Ng (2015). In this framework, a hurricane landfall introduces uncertainty for the local economy and firms. Predicting at the time of landfall which firms will be most affected could be challenging for several reasons. First, the number of hurricane landfalls for a given local economy are not sufficient to predict the exact economic effect. For example, Houston, TX, had not experienced a hurricane for more than two decades before Hurricane Harvey hit in 2017. Second, a hurricane's impact on individual firms operating within a disaster region is to a large extent unpredictable. Knowing ex-ante exactly which areas will actually flood in a particular storm, the extent of power outages, or whether a levy will break, is challenging if not impossible.

Prior to the potential hurricane landfall, there is a second source of uncertainty which we call *landfall uncertainty* about whether the hurricane will make landfall. More generally, in other contexts, this corresponds to the uncertainty of the incidence or occurrence of an event or the uncertainty on the extensive margin. Similarly, impact uncertainty can be thought of as uncertainty on the intensive margin. At time t, we can decompose the uncertainty generated for the firm from the hurricane into *expected* impact uncertainty and landfall uncertainty as follows.

The expected return conditional on whether or not landfall occurs is, intuitively, $E_t[r_{i,t+1}|\theta = 1] = \mu_g$ and $E_t[r_{i,t+1}|\theta = 0] = 0$. The conditional variance of firm i's return is,

$$Var_t(r_{i,t+1}|\theta = 0) = \sigma^2, \tag{2}$$

$$Var_t(r_{i,t+1}|\theta = 1) = \sigma^2 + \sigma_g^2. \tag{3}$$

Then, we can find the expected conditional variance[11] and the variance of the conditional expectation,[12]

$$E[Var_t(r_{i,t+1}|\theta)] = \sigma^2 + \phi\sigma_g^2, \tag{4}$$

$$Var(E_t[r_{i,t+1}|\theta]) = \phi(1-\phi)\mu_g^2. \tag{5}$$

Applying the law of total variance, we can derive $Var_t(r_{i,t+1})$ using (4) and (5),

$$\begin{aligned} Var_t(r_{i,t+1}) &= E[Var_t(r_{i,t+1}|\theta)] + Var(E_t[r_{i,t+1}|\theta]), \\ &= \sigma^2 + \phi\sigma_g^2 + \phi(1-\phi)\mu_g^2. \end{aligned} \tag{6}$$

Landfall uncertainty is captured in the total variance by the third term in equation (6), $\phi(1-\phi)\mu_g^2$. For a given $\mu_g \neq 0$, landfall uncertainty is highest when the probability of landfall, $\phi = 0.5$. When $\mu_g = 0$, there is no contribution from landfall uncertainty to total variance at time t. In this case, $Var_t(r_{i,t+1})$ varies with ϕ purely due to the expected impact uncertainty, $\phi\sigma_g^2$.

Figure 1 depicts how the total variance prior to landfall ($Var_t(r_{i,t+1})$) varies with the probability of hurricane landfall (ϕ). The figure has parameters $\sigma = 0.4$ and $\sigma_g = 0.05$. The four dashed lines have μ_g values of 0.1, 0.07, 0.05, and 0, respectively. The solid line shows the level of variance following hurricane landfall, $Var_t(r_{i,t+1}|\theta = 1) = \sigma^2 + \sigma_g^2$.

Depending on the the parameter values of μ_g and σ_g^2, as ϕ varies from 0 to 1, prior to landfall, the relative contribution to total variance from the landfall uncertainty and the expected impact uncertainty will vary. All else equal, as μ_g increases, the contribution of landfall uncertainty to total variance increases. In Figure 1, landfall uncertainty at a given ϕ is the vertical distance between a curve and the dot-dash (red) straight line depicting $Var_t(r_{i,t+1})$ when $\mu_g = 0$. $Var_t(r_{i,t+1})$ will in fact be greater than $Var_t(r_{i,t+1}|\theta = 1)$ when $|\mu_g| > \frac{1}{\sqrt{\phi}}\sigma_g$. In the figure, this is the case where the dashed lines are above the solid black line. When $\phi > 0$ and at least one of μ_g or σ_g is non-zero, $Var_t(r_{i,t+1})$ is greater than $Var_t(r_{i,t+1}|\theta = 0) = \sigma^2$.

[11] $E[Var_t(r_{i,t+1}|\theta)] = (1-\phi)\sigma^2 + \phi(\sigma^2 + \sigma_g^2) = \sigma^2 + \phi\sigma_g^2$

[12] $E[E_t[r_{i,t+1}|\theta]] = \phi\mu_g$,
$Var(E_t[r_{i,t+1}|\theta]) = E[(E_t[r_{i,t+1}|\theta] - \phi\mu_g)^2] = \phi(\mu_g - \phi\mu_g)^2 + (1-\phi)(0-\phi\mu_g)^2 = \phi(1-\phi)\mu_g^2$.

2.2 Identification strategy

Changes to the expected volatility of stock returns due to a hurricane event can be measured by the changes to the implied volatility of the stock's options. In our analysis, we use single stock options of firms that are in the damage region or forecasted path of a hurricane to estimate treatment effects, while using single stock options of firms unaffected by a hurricane as controls.

Because an extreme weather event like a hurricane is generally a local phenomenon, our identification strategy is based on selecting counties where a hurricane has made (or is predicted to make) landfall. For each hurricane, we have data on which counties were damaged and which counties lay in the forecasted path of the hurricane on a given day. A firm's exposure to these counties are then measured through the share of establishments or sales located in such counties. For firm i on day T_h, the landfall day of hurricane h, the exposure to counties damaged by hurricane h is given by

$$HurricaneDamageExposure_{i,T_h} = \sum_c (FirmCountyExposure_{i,T_h,c} \times I_{c \in D_{T_h}}), \tag{7}$$

where $FirmCountyExposure_{i,T_h,c}$ is the share of firm i's establishments (sales) located in county c, and D_{T_h} is the set of counties damaged by hurricane h. Therefore, a firm's exposure to a hurricane is a continuous variable that can range from 0 percent to 100 percent. To measure a firm's exposure to a forecasted hurricane path, we can use the set of counties in the forecasted path of a hurricane Γ days before the landfall or dissipation of the hurricane, denoted $F_{T_h-\Gamma}$, in place of D_{T_h}:

$$HurricaneForecastExposure_{i,T_h-\Gamma} = \sum_c (FirmCountyExposure_{i,T_h-\Gamma,c} \times I_{c \in F_{T_h-\Gamma}}). \tag{8}$$

For each hurricane, there are two groups of firms, those with and without exposure, with the degree of exposure being heterogeneous. Therefore, our analysis can be thought of as a differences-in-differences setting, where each hurricane represents a treatment, and we jointly estimate the treatment effect across all the hurricanes.

Because a hurricane has an identified inception date, we can isolate and estimate a hurricane's impact uncertainty, described in Section 2.1 by comparing the implied volatility of firms exposed to the landfall area shortly after landfall to the implied volatility before the inception of the hurricane. The implied volatility dynamics of the options of firms with zero exposure over the same time window comprise the control set. To measure landfall uncertainty, we rely on hurricane forecasts from NOAA. NOAA releases forecasts for the path of a hurricane starting from the hurricane's inception. For each hurricane, these forecasts provide a landfall probability for each county and each day since the inception of the hurricane. The landfall uncertainty defined in Section 2.1, can be computed based on the probabilities issued by NOAA.

3 Data and summary statistics

We combine data from a range of sources. We use data both from the Federal Emergency Management Agency (FEMA) and the Spatial Hazard Events and Losses Database for the United States (SHELDUS) when determining which counties are affected by hurricanes. We identify county level pre-landfall hurricane risk levels using archived forecasts from the NOAA. We combine these data sources with National Establishment Time-Series (NETS) data on locations of firm establishments and sales to identify firm exposure to hurricanes. And finally, our stock and option outcome data come from CRSP-Compustat and OptionMetrics, respectively. We describe each of these data sources below.

3.1 Hurricane damages

We use FEMA Disaster Declarations Summary data in combination with Spatial Hazard Events and Losses Database for the United States (SHELDUS) data to identify counties that experienced significant hurricane damages. We create an indicator that equals 1 if a county received a FEMA disaster declaration qualifying residents for individual and household program (IHP) assistance due to a hurricane. Counties are only eligible for IHP aid if they sustain significant damage on a per capita basis.

We build on the basic disaster indicator based solely on FEMA declarations by combining it with SHELDUS data in order to allow for the possibility that there are areas with significant damages which do not receive IHP aid.[13] SHELDUS data draw upon National Centers for Environmental Information (formerly National Climatic Data Center) Storm Data and Unusual Weather Phenom-

[13] A reason for why we may see SHELDUS damages in excess of reported FEMA declaration thresholds without seeing FEMA declarations could be because of measurement errors and disagreements or if FEMA chooses to diverge from its reported guidelines.

ena for hurricanes to provide county level per capita estimates of damages for named hurricane events. We set our second disaster indicator variable equal to 1 if there was a FEMA declaration for IHP aid or if the SHELDUS-reported per capita damages exceeded the published threshold for FEMA to provide IHP aid. Summary statistics are provided in Table 1. Figure 2 shows the number of times each county received an IHP declaration from FEMA for a hurricane between 2007 and 2017, while Figure 3 shows which counties received IHP aid or met the threshold according to SHELDUS in the case of hurricane Sandy.

3.2 Hurricane forecasts

We use NOAA's National Hurricane Center (NHC) wind speed probability forecasts to develop our measure of a hurricane's landfall uncertainty prior to landfall. In particular, we use text files containing probabilities that particular locations will experience winds in excess of 64 knots (KT), which is the lower bound windspeed for hurricanes. Because NOAA does not issue forecasts for the damage that counties could experience, the hurricane windspeed forecasted for a county acts as a proxy for the amount of damage, with higher windspeed implying larger damage.

The windspeed probabilities are presented cumulatively for 12-hour windows up to five days out from the time of each forecast. The NHC reports these wind speeds for cities, towns, and military bases along the coast as well as some major cities that are more inland (including Birmingham, AL, Savannah, GA, and Washington, DC.) There are three wind speed thresholds included in these reports, the lowest and highest of which are the cutoffs for tropical storm-force and hurricane-force winds, respectively. These windspeed data are derived from the same underlying data of the hurricane forecast charts published by the NHC in real time and used by news outlets in the run-up to hurricanes. Figure 4 shows an example of the forecast chart of cumulative probability bands for hurricane force winds, as presented by the NHC, over a five day period in the case of Hurricane Sandy in 2012.

Table 2 reports that the number of storms for which we observe forecasts decreases as the windspeed or probability thereof increases. We have several times the number of storm events to examine when we look not only at the hurricanes that make landfall (shown in the final two rows) but also at the storms that had a real possibility of making landfall (shown in the second and third rows.)

We have taken two steps to deal with the fact that the wind speed probabilities in these text files are only reported for particular locations, most of which are coastal. First we define a set of criteria that counties with data must meet in order to be considered at risk. For example, we might define a location as being at risk if it has at least a 20 percent chance of experiencing hurricane-force winds, that is 64 KT and above, within the next five days. Second, we define as at-risk any county within a 75-mile radius of a county with data that has been defined as being at-risk in the first step. For each day, we only use the last available forecast before close of trading, as forecasted hurricane paths can change meaningfully over the course of a day. Figure 5 illustrates a sample of processed wind speed data for two hurricanes.

One benefit of using the wind speed probability data is that it provides us information not only on where the eye of the storm is expected to be but also on how intense the winds will likely be and how wide the impact will be geographically. A hurricane that technically never makes landfall because the eye of the storm never passes over land can still pass close enough to a coastline to generate significant damage through strong winds, heavy rainfall, and storm surge. The wind speed forecast would show strong winds in the coastal areas closest to this hurricane. We will be referring to locations with forecasted windspeeds of 64 KT or more as the locations where the hurricane is predicted to make landfall. These windspeed forecasts are available from 2007 to 2017.

3.3 Firm's geographic footprint

We use data on locations of firm establishments and sales in order to precisely estimate firm exposure to specific hurricanes. In particular, we use NETS data (see, for example, Neumark, Wall, and Zhang (2011) and Barnatchez, Crane, and Decker (2017)) to compute the geographic footprint of a firm. The NETS data contain establishment locations of public and private firms at the county level at an annual frequency. The data also include sales data for each establishment. The time series starts in 1990 and ends in 2014. For each hurricane season, we use the firms' geographic footprints from the previous year. For hurricanes from 2015-2017, we use the geographic footprint from 2014.

3.4 Financial data

To obtain financial data for the firms in NETS, we map them to OptionMetrics and CRSP-Compustat, both data sources are described below. The mapping is conducted based on the name of the firm and the address of the headquarters. We exclude all financial firms with SIC numbers from 6000 to 6799 from our analysis.[14] Summary statistics for the NETS data are reported in Table 3. From 1995 to 2014, we have 4,999 (4,990) firms in our sample with establishment (sales) data. On average, a county has 61 establishments and 478 million in sales. For counties that have experienced hurricane damage the number is higher with an average of 82 establishments and 634 million in sales. Figure 6 shows counties sorted into deciles based on the number of establishments for the years 2010 and 2014. The economic activity as measured by firm establishments is high in the hurricane prone areas along the Atlantic and the Gulf Coast.

We obtain daily frequency stock data from CRSP-Compustat and single-name stock options from OptionMetrics. Similar to previous studies we use the data from out-of-the-money traded options with valid pricing information,[15] and we restrict the set of options to slightly out-of-the-money options. These are more liquid and have a relatively smaller effect from any potential early-exercise premium for American options over European options for the time horizons that we consider.

Accordingly, we include single-stock options with: (i) standard settlement, (ii) a positive open interest, (iii) a positive bid price and bid-ask spread, (iv) a valid implied volatility estimate, (v) greater than 7 days and at most 200 days to expiry, and (vi) an option delta, δ, that satisfies $0.2 \geq |\delta| \leq 0.5$. The estimate for the average implied volatility of firm i at date t is, $IV_{i,t} = \frac{1}{N}\sum_{j=1}^{N} IV_{i,j,t,M}$, where M is the nearest-to-maturity expiration at time t with valid options which satisfy the above criteria and N is the number of valid stock options for firm i with that expiry. Using this methodology for the period from 2007 to 2017, we obtain 7,551,261 firm-date observations of implied volatility from OptionMetrics. We merge these data with CRSP-Compustat on firm CUSIP, which yields 7,089,684 observations covering 5,175 firms and 2,769 dates. Of these, 2,937,420 observations are from 1,995 firms that appear at least once in the firm establishment data from NETS. The summary statistics for these data are in Table 4.

[14] We provide a separate analysis on insurance firms in Section 4.3.5.

[15] See, among others, Carr and Wu (2009); Kelly, Pastor, and Veronesi (2016); Martin and Wagner (2018).

4 Results

In this section, we describe the regression specifications that we employ to test our hypotheses and the corresponding results. The three hypotheses are: (i) stock options of firms in hurricane disaster regions exhibit a higher implied volatility right after a hurricane has hit, in line with investors trying to hedge their exposure to the impact uncertainty surrounding the firms that are in the disaster zone of the hurricane; (ii) abnormal stock returns of firms with a physical presence in a hurricane disaster region show a large dispersion that is negatively skewed in the period after a hurricane has hit; and (iii) pre-landfall, stock and option prices react to hurricane forecasts, with implied volatility increasing and stock prices decreasing for firms located in the forecasted hurricane paths, consistent with investors paying attention to hurricane forecasts.

4.1 Impact uncertainty estimation

We begin with testing the hypothesis that stock options of firms in disaster regions have higher implied volatilities. The implied volatility is an estimate of expected future volatility and is commonly used as a measure of uncertainty. If a hurricane landfall leads to impact uncertainty for firms in the disaster region, the implied volatilities of these firms should increase. The impact uncertainty can be isolated and estimated by looking at the implied volatilities shortly after landfall, when investors know where the hurricane hit, that is all the landfall uncertainty has disappeared, but do not know what the eventual impact on the firms located in the damage region will be. To test this first hypothesis, we estimate the following panel regression model for each τ,

$$log\left(\frac{IV_{i,T_h+\tau}}{IV_{i,T_h^*}}\right) = \lambda HurricaneDamageExposure_{i,T_h} + \theta_h + \psi_{Ind} + \epsilon_{i,h,\tau}, \tag{9}$$

where τ is the number of trading days since the hurricane made landfall on day T_h.[16] The last trading day before the *inception* of the hurricane is T_h^* and IV_i is the implied volatility of firm i. $HurricaneDamageExposure_{i,T_h}$ is a measure of firm i's exposure to the counties with hurricane damage, as defined in equation (7). This measure can vary from 0 percent, for firms with no exposure to the hurricane disaster region, to 100 percent, for firms with all of their establishments

[16] If a hurricane makes multiple landfalls, the first landfall date is used as T_h in the analysis.

(or sales) located within the disaster region. The NETS data allow us to use different variables to measure the county exposure of a firm, namely, the amount of establishments or sales in a specific county as a percent of the firm's total establishments and sales throughout the U.S. We include hurricane fixed effects (θ_h), which is equivalent to including time fixed effects because there is one time period per hurricane. We include industry fixed effects (ψ_{Ind}) based on firm SIC codes. As our treatment selection is at a geographic level, we cluster the standard errors based on the county where the headquarters of a firm is located (see, for example, Dessaint and Matray (2017)).

The regression model in equation (9) can also be seen as a differences-in-differences estimation where each hurricane acts as a treatment, that is firms with exposure to the disaster zone are considered treated and firms with no exposure to the disaster zone act as controls. Following the recommendation of Bertrand, Duflo, and Mullainathan (2004), for each hurricane, we collapse the time series information into a pre- and post-period, where the pre-period is T_h^*, the day before the inception, and the post-period is $T_h + \tau$, τ days after the landfall.

The coefficient estimate of λ is expected to be significant and positive if investors perceive that hurricane landfall leads to impact uncertainty surrounding the local firms. A hurricane making landfall is thought to introduce severe uncertainty for the local economy and firms. Knowing ex-ante which firms will be most affected is likely impossible because of several factors. First, the number of hurricane landfalls for a given local economy are mostly insufficient to predict the exact economic effect. Second, the hurricane's impact on individual firms in the disaster zone is to a large extent random, as described in Section 2.1. Therefore, investors are expected to hedge themselves in the immediate aftermath of a hurricane against the impact uncertainty.

The estimation results of the model given in equation (9) are reported in Table 5. Panel A shows the results when the exposure of a firm to the hurricane disaster zone is based on establishments. We consider selecting counties in the disaster region solely based on FEMA damage data and FEMA damage data enhanced with SHELDUS data. The value τ, trading days after landfall, is set to 5 days, but our results are robust to choosing a different τ close to landfall. We have high quality option data available from 2002. We also show the results for the time period from 2007 to 2017, as the hurricane forecast data used in the subsequent analysis starts in 2007.

The estimate of λ is significant and positive for all specifications, which is in line with our first hypothesis that hurricane landfall causes impact uncertainty for local firms. In particular, we find that a firm with 100 percent of its establishments located in the disaster region will experience a 8 percent increase in its implied volatility relative to before inception of the hurricane. This economic magnitude is considerable and comparable to Kelly, Pastor, and Veronesi (2016), who show that political uncertainty leads to implied volatilities of index options increasing on average around 5 percent compared to non-election periods. These results are robust to including industry times time fixed effects, which implies that the effects are preset within industry.

These results are robust to measuring the geographic footprint of a firm by sales at the county level instead of establishments, as shown in Panel B. The estimates of λ in Panel B are also strongly significant for all the specifications. The largest coefficient estimates for five days after the landfall are 0.06, implying that a firm with a 100 percent of its sales in the disaster region has an implied volatility that is 6 percent higher than before the inception of the hurricane. The magnitudes for effects based on exposure of sales (Panel B) are somewhat smaller than those based on exposure of establishments (Panel A). This could be explained by investors being more concerned about damages to production facilities or investors having better information about locations of firm establishments than sales.

Importantly, our results are not driven by small firms. The average market capitalization of a firm with exposure to disaster region counties of at least 25% is \$4.8 billion and \$8.1 billion when measuring the exposure by establishments and sales, respectively. The average market capitalization of a firm with less than 25% exposure to the disaster region is similar in magnitude with \$6.6 billion and \$6.4 billion, respectively.

We also test if these results are driven by a particular industry but find that the impact uncertainty of hurricanes affects firms of all industries similarly. Table A.1. in the appendix adds an industry dummy interacted with $HurricaneDamageExposure_{i,T_h}$ to the model in equation (9). The coefficient estimate of the interaction term is insignificant for almost all specifications, which suggests that the reported effect is not driven by one particular industry. Only the transport industry shows a stronger relationship between hurricane exposure and implied volatility compared to other industries when measuring the geographic footprint of firms by sales. However, when measuring the geographic footprint by establishments, the relationship between hurricane exposure and implied volatility is not significantly different from other industries.

4.2 Impact uncertainty resolution

The large impact uncertainty measured in the previous section suggests that firms in the disaster region face uncertain outcomes. The resolution of this impact uncertainty should be reflected in the firms' stock prices in the months following a hurricane landfall. In this section, we test if the abnormal stock returns of firms with exposure to a hurricane disaster region show a large dispersion that is negatively skewed in the post-landfall period. A negatively skewed dispersion would be in line with a slow resolution of impact uncertainty, that is investors learning over time which firms were most adversely affected, and could explain the increases in implied volatility as a measure of uncertainty right after landfall.

To estimate how a hurricane affects firm stock returns after landfall, we first estimate daily abnormal returns relative to the Fama-French three-factor model (see Fama and French (1993)). For each firm and each hurricane in our sample, we estimate the following model:

$$r_{i,d} = \alpha_i + \beta_{1,i} r_{m,d} + \beta_{2,i} r_{smb,d} + \beta_{3,i} r_{hml,d} + \epsilon_{i,d}, \tag{10}$$

where $r_{m,d}$ is the daily market return on day d, $r_{smb,d}$ and $r_{hml,d}$ are the daily returns of the small-minus-big and high-minus-low portfolios, respectively. We estimate this model using 200 trading days before the day of hurricane landfall. We then use the coefficient estimates from this first stage regression to compute abnormal returns for each firm and hurricane as follows:

$$r^a_{i,d} = r_{i,d} - (\hat{\alpha}_i + \hat{\beta}_{1,i} r_{m,d} + \hat{\beta}_{2,i} r_{smb,d} + \hat{\beta}_{3,i} r_{hml,d}). \tag{11}$$

We next aggregate the abnormal returns to a cumulative abnormal return, denoted $r^{ac}_{i,T_h:T_h+\tau}$, for each firm and hurricane over the time period T_h to $T_h + \tau$, where T_h is again the day of the landfall and τ is a number of trading days.[17] The time period starts in 1995 and ends in 2017.[18]

We compute the differences in cumulative abnormal returns between firms hit by the hurricane and control firms for the mean and nine quantiles of returns, which are reported in Table 6 along with the corresponding t-stats. We estimate the standard errors using a bootstrap that clusters by

[17] If a hurricane makes landfall on a non-trading day, then we take the next trading day as T_h.

[18] This time periods is longer than for the sample of implied volatilities with which we estimate the regression model in equation (9), as we have a longer time series of high quality stock return data available.

county based on firm headquarters.[19] The cumulative abnormal returns are computed for up to 5 and 120 days after landfall. For Panel A, we consider firms to be in the disaster zone if at least 50 percent of the establishments are in the disaster region. For Panel B, the threshold is 50 percent of the sales.

As shown in Panel A, we find that the cumulative abnormal returns from the landfall day to five days after yield a negative difference for the mean and all quantiles except the top one. These differences are generally between -10 and -50 basis points, but they are not significant. When looking at the cumulative abnormal returns from landfall day to 120 days after the landfall, the differences in cumulative abnormal returns are strongly negatively skewed. For the bottom quantile, the difference in cumulative abnormal returns is -12 percent, but for the top quantile the difference is 3 percent. While the mean difference is not statistically significant, the median and bottom half of the quantiles differences are statistically significant. In Panel B, firm exposure to hurricane disaster regions is measured based on a firm's sales in a county. The differences in cumulative abnormal returns are again negatively skewed and comparable to Panel A in magnitude and statistical significance.

These results show that considering just the average negative impact on stock returns post hurricane landfall gives an incomplete picture. There are some firms that are adversely affected and these firms have strongly negative abnormal stock returns. This negative skewness in the abnormal return distribution for the firms in the landfall area is in line with investor concern about potentially large drops in stock prices once the impact uncertainty is slowly resolved. Unable to discern at the time of landfall the implications of hurricanes for specific firms, investors hedge themselves through options, thereby increasing the implied volatilities.

4.3 Uncertainty before landfall

In Table 5 we show that days after landfall, options price in substantial impact uncertainty in hurricane disaster regions. However, over the course of the days or weeks while a hurricane makes its approach toward the Atlantic or the Gulf Coast, NOAA issues hurricane forecasts that contain the probabilities of the hurricane making landfall in a particular region. Such forecasts are often highly publicized through news outlets. For example, the shape of the forecasted path of Hurricane

[19] To ensure clusters of a sufficient size, we exclude hurricanes with fewer than 25 affected firms.

Sandy in 2012 shown in Figure 4 likely looks familiar to people who tend to follow the news during hurricane season. Based on the efficient market hypothesis, investors should pay attention to these forecasts, and the forecasts should be priced in. If investors pay attention to hurricane forecasts before landfall, then the impact uncertainty will increasingly be priced into options as the likelihood of a hurricane making landfall in a specific region increases, which is represented by the term $\phi\sigma^2$ in equation (6). In addition, investor attention to hurricane forecasts will also lead to landfall uncertainty, given by the term $\phi(1-\phi)\mu_g$ in equation (6), being reflected in option prices through higher implied volatilities.

We use the NOAA forecasts described in Section 3.2 to examine how hurricane forecasts affect implied volatilities of firms located in the path of a hurricane. We estimate the following panel regression model

$$log\left(\frac{IV_{i,T_h-\Gamma}}{IV_{i,T_h^*}}\right) = \lambda HurricaneForecastExposure_{i,T_h-\Gamma} + \theta_h + \psi_{Ind} + \epsilon_{i,h,\Gamma}, \tag{12}$$

where Γ represents the number of calendar days before the landfall or dissipation of the hurricane and we estimate the regression separately for each $\Gamma \in \{1,2,3,4,5\}$, as NOAA forecasts hurricane paths up to five days out.[20] Firm i's exposure to hurricane h's forecasted path, $HurricaneForecastExposure_{T_h-\Gamma}$, is as defined in equation (8). The remaining parameters are as described for regression equation (9). Only hurricanes for which the day $T_h - \Gamma$ is a trading day are included a regression. The time series starts in 2007, because we have high quality hurricane forecast data from 2007 onwards, and ends in 2017. As described in detail in Section 3.2, the hurricane forecasts provide a probability that a county will experience windspeeds of at least 64 KT within five days. A storm's windspeed has to be at least 64 KT to be classified as a hurricane by NOAA.

If investors pay attention to hurricane forecasts, the estimate of λ is expected to be positive and significant. Particularly, the change in a firms' implied volatilities should depend on the probability that a hurricane will make landfall in counties in which the firm operates. In our framework presented in Section 2.1, we show in Figure 1 that for any probability of landfall greater than zero, given by the term ϕ, the implied volatility will be higher than before the inception of the hurricane. Further, the total uncertainty given in equation (6) can be higher before landfall, when

[20]If a hurricane makes landfall on multiple days, we only consider the first landfall day.

landfall and impact uncertainty are present, then right after landfall when there is no uncertainty about landfall but only uncertainty about the impact of a hurricane. Figure 1 shows that depending on the parametrization, the total variance (uncertainty) can be higher before landfall, when ϕ is smaller than 1, than at landfall, that is when ϕ equals 1. Whether total uncertainty is higher before landfall than right after landfall is ultimately an empirical question.

We report the estimation results of the model in equation (12) in Table 7. The parameter Γ is between 1 and 5 days, and the probabilities of hurricane-level windspeeds that we require to designate a county as at-risk ranges from 1 to 50 percent. For each Γ and probability, we require that at least three hurricanes and 25 firms that have an exposure of 20 percent or more of their establishments or sales in counties in the path of a hurricane. Because the days before the landfall or dissipation of a hurricane can fall on non-trading days and different hurricanes reach probability thresholds of making landfall on different days, the hurricanes included in the estimation can vary across the table's columns.

The results are in line with investors paying attention to hurricane forecasts and the uncertainty surrounding a hurricane being reflected in the implied volatilities of firms located in the forecasted path of a hurricane. The estimates of λ are always positive, regardless of whether a firm's exposure to a hurricane is based on establishments (Panel A) or sales (Panel B). The λ estimates are also significant with the exception of the estimates five days before landfall/dissipation. For each day, the magnitude of λ increases with higher landfall probabilities. Particularly, for high probabilities, the increase in implied volatilities is in some cases larger than the increase in implied volatilities right after landfall reported in Table 5.[21] This result suggests that the landfall uncertainty is strongly reflected in option prices and can push the total uncertainty before landfall above the impact uncertainty measured at landfall, as suggest by our framework in Section 2.1. Overall, these results are consistent with hurricane forecasts containing valuable information and investors paying attention to them.

One interesting observation is that the estimated magnitude of λ for the same probability is sometimes lower for days closer to the landfall or dissipation of the hurricane, although not

[21] A caveat is that the sample of hurricanes in the two tables can differ. In particular, while in Table 5 we include only the hurricanes that make landfall, in Table 7 we also consider hurricanes that dissipate without making landfall. Also, for some hurricanes included in Table 5, a specific day before landfall can be a non-trading day, and thus, the hurricane would not be included for that day in Table 7. However, the result of higher total uncertainty before the landfall than right after landfall holds when comparing the same hurricanes.

significantly so. A possible explanation for this is that hurricanes that reach a specific probability of making landfall when they are still far off the coast are simply stronger hurricanes that can lead to more devastating effects.

4.3.1 Alternative specification for forecasts and implied volatilities

The estimates of the regression model shown in equation (12) support the hypothesis that investors pay attention to hurricane forecasts and the uncertainty surrounding a hurricane is reflected in option prices before landfall. To further test the robustness of this result, we use an alternative estimation where we allow for the fact that firms can reach an exposure threshold to a specific hurricane, for example, 10 percent of establishments are located in the forecasted path of a hurricane, on different days. In the regression model in equation (12), it is not possible to jointly estimate the change in implied volatilities for these firms. The model specification below allows for a joint estimation, but the hurricane exposure variable is an indicator variable instead of a continuous variable as in equation (12). We compute the measure given by

$$IVD_{i,h} = log\left(\frac{IV_{i,t_{i,h}}}{IV_{i,T_h^*}}\right) - \frac{1}{|J_h|}\sum_{J_h} log\left(\frac{IV_{j,t_{i,h}}}{IV_{j,T_h^*}}\right), \tag{13}$$

where $t_{i,h}$ is the first trading day when the number of establishments (sales) of firm i in the path of hurricane h exceed a certain threshold, and T_h^* is again the last trading day before the inception of hurricane h. The set of control firms, J_h, for hurricane h are the firms with zero exposure to the forecasted path of the hurricane. We exclude from this analysis the days on which a hurricane makes landfall. We compute $IVD_{i,h}$ for all hurricanes and firms and estimate the mean, $\overline{IVD}$, for the sample from 2007 to 2017. A positive and significant $\overline{IVD}$ would be consistent with the results shown in Table 7.

The results for this specification are presented in Table 8. Here we use five probability thresholds ranging from 1 to 50 percent to designate which counties lie within the forecasted path of a hurricane. A 1 percent probability threshold implies that a county has at least a 1 percent chance of experiencing hurricane-force winds in the next 5 days. We consider three thresholds for a firm's exposure to a hurricane wherein 10, 25, and 50 percent of a firm's establishments (sales) are located in counties that we have designated as being in the hurricane's forecasted path. For a 10 percent threshold, we compute the difference in the implied volatility of firm i on the first trading day that 10 percent of firm i establishments (sales) are located in the hurricane's forecasted path and the implied volatility on the last trading day before the inception of the hurricane.

Panel A reports the estimates of $\overline{IVD}$ when the firm geographic footprints are computed based on the share of establishments in a county. The estimates of $\overline{IVD}$ are positive and significant for the great majority of the specifications. The only two exceptions are for the probability threshold of 1 percent. As in Table 7, the magnitude of the estimates is monotonically increasing with the probability of the firms being hit by a hurricane. Further, when selecting only firms with at least 25 or 50 percent of their establishments or sales in at-risk counties, the estimates are substantially larger than for the firms with an exposure of at least 10 percent. These results further support the hypotheses that investors pay attention to hurricane forecasts which leads to uncertainty being reflected in the implied volatilities of firms located in the forecasted path of a hurricane.

4.3.2 Forecasts and stock returns

While uncertainty is generally measured with implied volatilities computed from option prices, uncertainty is also expected to affect stocks prices. Investors are thought to require a premium to hold stocks during a time period of high uncertainty, as, for example, discussed by Pastor and Veronesi (2012) and Pastor and Veronesi (2013) in the context of political uncertainty. Therefore, the previously documented impact uncertainty that is reflected in option prices before and right after a hurricane makes landfall should also be reflected in stock returns. However, while a hurricane's impact uncertainty leads to higher implied volatilities, its effect on stock prices should be the opposite, that is, stock returns should be negative.

To test this hypothesis, we estimate the regression model in equation (12), but with log stock returns as the dependent variable:

$$r_{i,T_h^*:T_h-\Gamma} = \lambda HurricaneForecastExposure_{i,T_h-\Gamma} + \theta_h + \psi_{Ind} + \epsilon_{i,h,\Gamma}, \tag{14}$$

where $r_{i,T_h^*:T_h-\Gamma}$ is the log return of firm i from the inception of hurricane h to Γ calendar days before the landfall or dissipation of the hurricane. An estimate of λ that is significant and negative would support the hypothesis that the uncertainty surrounding firms in a hurricane's path leads to negative stocks returns.

The results are reported in Table 9, which is structured as Table 7. The estimates of λ are negative in all cases except for the last column that examines effects five days before landfall/dissipation with a probability of a hurricane hit of 10 percent. The estimates are strongly significant for the majority of the specifications. The estimates are also economically significant. The smallest estimate is -0.09, which implies that a firm with a 100 percent exposure to the forecasted path of a hurricane experiences a negative return of 9 percent from the inception of the hurricane to a few days before the landfall/dissipation. The estimates are similar when the geographic footprint of a company is based on establishments, as in Panel A, or on sales, as in Panel B. These results are consistent with our analysis on implied volatilities and support the hypothesis that the uncertainty associated with a hurricane leads to negative stock returns.

4.3.3 Can the market forecast better than NOAA?

The previous results show that market prices react to hurricane forecasts by pricing in the impact uncertainty caused by a potential hurricane strike. The hurricane forecasts in our analysis are taken from NOAA. NOAA's hurricane forecasts are arguably the most prominent as they are widely publicized through the media. However, it is possible that large institutional investors like hedge funds, which often act as marginal investors in asset markets and move asset prices, could preempt NOAA hurricane forecasts by trading on superior proprietary hurricane forecasts. In this case, markets would predict hurricane damages more precisely than NOAA forecasts. There are a couple of reasons to believe that markets could predict damages more precisely. First, the budget of NOAA's subdivision responsible for hurricane forecasts, the National Weather Service, is minuscule compared to the value of assets managed by large institutional investors.[22] Second, there is anecdotal evidence that hedge funds buy information on hurricane forecasts from private companies.[23]

We test this hypothesis by estimating the panel regression model in equation (12) with an additional term that measures whether option markets can predict which firms end up more exposed

[22]The total budget of the National Weather Service, a subdivision of NOAA, was around $1 billion in 2017. However, this budget also includes funds appropriated for weather forecasts other than hurricane forecasts. The budget of the National Weather Service for 2017 can be found here: https://www.corporateservices.noaa.gov/nbo/.

[23]See, for example, the discussion of the hedge fund with the name Nephila by Michael Lewis here: https://www.nytimes.com/2007/08/26/magazine/26neworleans-t.html?pagewanted=all.

to the hurricane than predicted by the NOAA forecasts:

$$
\begin{aligned}
log\left(\frac{IV_{i,T_h-\Gamma}}{IV_{i,T_h^*}}\right) =& \lambda HurricaneForecastExposure_{i,T_h-\Gamma} \\
&+ \gamma UnderPrediction_{i,T_h-\Gamma} + \theta_h + \psi_{Ind} + \epsilon_{i,h,\Gamma}.
\end{aligned}
\tag{15}
$$

Here $UnderPrediction_{i,T_h-\Gamma}$ is defined as the difference between a firm's exposure to counties that eventually experience hurricane related damages and the exposure to counties in a hurricane's forecasted path:

$$
\begin{aligned}
UnderPrediction_{i,T_h-\Gamma} =& (HurricaneDamageExposure_{i,T_h} - HurricaneForecastExposure_{i,T_h-\Gamma}) \\
&\times I_{(HurricaneDamageExposure_{i,T_h} - HurricaneForecastExposure_{i,T_h-\Gamma})>0}.
\end{aligned}
\tag{16}
$$

Firm i will have a positive value for $UnderPrediction_{i,T_h-\Gamma}$ if the share of its establishments or sales in counties that experience hurricane damages is greater than the share of its establishments or sales predicted to be affected based on NOAA forecasts made Γ days before landfall. Otherwise, $UnderPrediction_{i,T_h-\Gamma}$ will assume a value of 0. If the market can forecast which counties will experience hurricane damage more accurately than NOAA, the estimate of γ in equation (15) would be significant and positive.[24]

We look at underpredicted firms rather than overpredicted firms, because a smaller than average increase in implied volatility for overpredicted firms (firms that end up with less exposure to the damage region than forecasted) could be explained by markets being less attentive to low probability forecasts, which include many firms in the forecasted path that end up with no or little exposure to the damage region. Therefore, analyzing underpredicted firms allows us to isolate and estimate the forecast ability of financial markets.

The results are shown in Table 10, which has the same structure as Table 7. While the estimates of γ are positive, they are insignificant for all but one case, and there the coefficient is only weakly significant. Therefore, we do not find support for the hypothesis that markets can forecast hurri-

[24]Suppose, for example, that the NOAA forecast implies zero exposure for a firm four days before a hurricane's actual landfall. If the firm ends up with significant exposure to counties affected by the hurricane four days later, $UnderPrediction_{i,T_h-4}$ would equal the exposure of the firm to the actual hurricane. If the markets are able to predict the final exposure four days ahead when the NOAA forecast did not, γ would reflect this by being significant and positive.

canes better than NOAA. Given that the number of firms with a non-zero $UnderPrediction_{i,T_h-\Gamma}$ measure is large with an average of 1,267 across the specifications, we should have sufficient power to detect the market's ability, if any, to beat NOAA forecasts. In contrast, the coefficient estimates on $HurricaneForecastExposure_{i,T_h-\Gamma}$ are positive and strongly significant for most specifications as in Table 7.

We also test if our results change when we focus on options for which the underlying stocks have a large institutional ownership, because institutional investors are more likely than retail investors to have the means to obtain hurricane forecasts that are not generated by NOAA. For this purpose, the variable $UnderPrediction_{i,T_h-\Gamma}$ in equation (16) is interacted with an indicator variable that takes the value one if the institutional ownership share of the underlying stock is above or equal to the sample mean at the quarter end before the inception of the hurricane.[25] Table 11 presents the results and they are largely in line with the results in Table 7. While the coefficient on $HurricaneForecastExposure_{i,T_h-\Gamma}$ remains strongly significant up to four days before hurricane landfall, the coefficient estimates on the $UnderPrediction_{i,T_h-\Gamma}$ variable with or without interaction with the institutional ownership indicator variable are almost always insignificant.

4.3.4 The economic effect of improved forecasts

The previous findings show that markets price in NOAA's hurricane forecasts and furthermore do not appear to be able to outperform NOAA forecasts. These findings lead to the question: what economic effect would improved hurricane forecasts have? In other words, how much of the price variation in options around hurricanes were caused by mispredictions, that is over- and underpredictions.[26] This price variation could potentially be reduced by more accurate forecasts.

To answer this question, we compute the average change in implied volatilities resulting from an overprediction (underprediction) of a firm's exposure to a hurricane.[27] To estimate by how much implied volatilities are too high due to an overprediction, we compute the average overprediction for each combination of number of days (before landfall or dissipation) and probability threshold

[25] The data on institutional ownerships of stocks are from Thomson Reuters.

[26] While the measure of underprediction is described in equation 16, the overprediction measure is defined as $OverPrediction_{i,T_h-\Gamma} = (HurricaneDamageExposure_{i,T_h} - HurricaneForecastExposure_{i,T_h-\Gamma}) \times I_{(HurricaneDamageExposure_{i,T_h} - HurricaneForecastExposure_{i,T_h-\Gamma})<0}$.

[27] Alley, Emanuel, and Zhang (2019) show that hurricane forecasts have indeed improved dramatically in recent decades. In particular, they find that "modern 72-hour predictions of hurricane tracks are more accurate than 24-hour forecasts were 40 years ago."

shown in Table 7 (e.g. 2 days at 10 percent.) We then multiply this average overprediction by the coefficient estimate on $HurricaneForecastExposure_{i,T_h-\Gamma}$ given in Panel A of Table 7. To estimate by how much implied volatilities are too low due to an underprediction, we multiply the average underprediction by the equation (9) coefficient on $HurricaneDamageExposure_{i,T_h}$ estimated for the respective hurricanes in the sample using five days post-landfall. The damage exposure of a firm is based on the combined FEMA and SHELDUS dataset, and number of establishments act as the geographic footprint measure.

Figure 7 presents the results of this analysis. Panel A shows the changes in implied volatilities resulting from over- and underprediction. The magnitudes of the changes are large, reaching 100 basis points for the overpredictions and 200 basis points for the underpredictions. Panel B presents the number of firms with non-zero over- or underprediction summed across all the hurricanes in our sample by forecast horizon and probability threshold. The number of firms with overprediction is particularly large at the 1 percent probability thresholds, reaching up to 6,000 firms three days prior to landfall. For underprediction, the probability thresholds show a more even distribution with average number of firms at 1,273. These results imply that improvements to hurricane forecasts could have large economic effects on pricing of hurricane related uncertainty in option markets as the price variations in options caused by under- and overpredictions are considerable.

4.3.5 Insurance firms

The analysis and discussion so far in this paper has been focused on the universe of firms excluding financial firms as it is common in the asset pricing literature. One contribution of this paper is to show that the uncertainty around extreme weather events affects a wide range of firms and not only insurance firms which are often thought of in the context of disaster events. However, we also want to investigate if extreme weather uncertainty is reflected in the asset prices of insurance firms. The challenge that we face is that the number of publicly traded insurance firms with liquid options is relatively limited and we only have the exposure of an insurance firm at the state but not at the county level.

We use data on insurance statutory financials from S&P Global Market Intelligence, which provides us with the share of total premiums written by state for property and casualty insurance firms in the US. We estimate the regression in equation (9) for these property and casualty insurance firms, with $HurricaneDamageExposure_{i,T_h}$ being replaced by a variable that measures the share of total premiums, lagged by one year, written in states that experienced damage by hurricane h. The results are reported in Table 12. The table is structured similarly as Table 5. Panel A (B) considers a state to have experienced hurricane damage if at least 10% (25%) of the counties experienced hurricane damage as measured by FEMA data and FEMA data enhanced with SHELDUS.

The coefficient estimate is positive and significant for all specifications implying that the impact uncertainty for property and casualty insurance firms is substantial in the aftermath of a hurricane. The magnitude of the coefficient estimates are economically significant, with the implied volatility being up to 40% higher for insurance firms with a 100% exposure to the damage region of the hurricane. The statistical significance is slightly weaker than for the universe of firms in Table 5 as the number of insurance firms in our sample is relatively small, we have on average 20 to 30 insurance firms per hurricane depending on the specification.

5 Conclusion

This paper studies how extreme weather uncertainty affects prices in option and stock markets by analyzing uncertainty surrounding hurricanes. Hurricanes are well suited for this purpose, as they can have severe effects on local economies, but the effect is isolated in time and place and can therefore be precisely estimated. Our framework distinguishes between landfall uncertainty (on when and where the hurricane will hit, if at all) and impact uncertainty (on the consequences to the local firms and economy following landfall). Knowing if and how extreme weather uncertainty affects option and stock prices is important for financial stability, assessing the cost of extreme weather events, and climate change regulation.

We find that options of firms operating in regions affected by hurricanes have considerably higher implied volatilities, of up to 8 percent, in the immediate aftermaths of those hurricanes. The higher implied volatilities are in line with investors being concerned about substantial impact uncertainty right after a hurricane has hit and hedging their exposure to this uncertainty. We use daily hurricane forecasts from NOAA to account for the uncertainty before landfall, that is the landfall uncertainty combined with the potential impact uncertainty. Our results show that investors pay attention to hurricane forecasts. Implied volatilities and stock returns increase and decrease, respectively, as the probability of a firm being hit increases. Further, we find no evidence that markets can forecast hurricanes better than NOAA. These results are in line with markets being efficient and investors considering hurricane forecasts to contain valuable information and being attentive to them. Also, our analysis shows that an improvement in NOAA's forecasts can have large economic effects in option markets.

Our novel analysis and framework contribute to a burgeoning climate finance literature. Further, we add to the existing uncertainty literature by showing that extreme weather uncertainty is important and reflected in the prices of options and stock markets.

References

Addoum, Jawad M., David Ng, and Ariel Ortiz-Bobea, 2019, Temperature shocks and earnings news, *Working Paper*.

Alley, Richard B., Kerry A. Emanuel, and Fuqing Zhang, 2019, Advances in weather prediction, *Science* 363, 342–344.

Andersson, Mats, Patrick Bolton, and Frederic Samama, 2016, Hedging climate risk, *Financial Analysts Journal* 72, 13–32.

Baker, Steven D., Burton Hollifield, and Emilio Osambela, 2018, Asset prices and portfolios with externalities, Discussion paper.

Baker, Scott R., Nicholas Bloom, and Steven J. Davis, 2016, Measuring economic policy uncertainty, *Quarterly Journal of Economics* 131, 1593–1636.

Baker, Scott R., Nicholas Bloom, and Stephen J. Terry, 2018, Does uncertainty reduce growth? Using disasters as natural experiments, *Working Paper*.

Barnatchez, Keith, Leland D. Crane, and Ryan A. Decker, 2017, An assessment of the national establishment time series (nets) database, *Working Paper*.

Barrot, Jean-Noel, and Julien Sauvagnat, 2016, Input specificity and the propagation of idiosyncratic shocks in production networks, *Quarterly Journal of Economics* 131, 1543–1592.

Belasen, Ariel R., and Solomon W. Polachek, 2008, How hurricanes affect wages and employment in local labor markets, *American Economic Review* 98, 49–53.

Bernanke, Ben S., 1983, Irreversibility, uncertainty, and cyclical investment, *Quarterly Journal of Economics* 98, 85–106.

Bernile, Gennaro, Vineet Bhagwat, and P Raghavendra Rau, 2017, What doesn't kill you will only make you more risk-loving: Early-life disasters and CEO behavior, *The Journal of Finance* 72, 167–206.

Bernstein, Asaf, Matthew Gustafson, and Ryan Lewis, 2018, Disaster on the horizon: The price effect of sea level rise, *Journal of Financial Economics forthcoming.*

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, 2004, How much should we trust differences-in-differences estimates?, *Quarterly Journal of Economics* 119, 249–275.

Bloom, Nicholas, 2009, The impact of uncertainty shocks, *Econometrica* 77, 623–685.

Bloom, Nick, Stephen Bond, and John van Reenen, 2007, Uncertainty and investment dynamics, *Review of Economic Studies* 74, 391–415.

Boudoukh, Jacob, Matthew Richardson, YuQing Shen, and Robert F. Whitelaw, 2007, Do asset prices reflect fundamentals? freshly squeezed evidence from the OJ market, *Journal of Financial Economics* 83, 397–412.

Brown, James R., Matthew T. Gustafson, and Ivan T. Ivanov, 2017, Weathering cash flows, *Working Paper.*

Campbell, Sean D., and Francis X. Diebold, 2005, Weather forecasting for weather derivatives, *Journal of the American Statistical Association* 100, 6–16.

Carney, Mark, 2015, Breaking the tragedy of the horizon - climate change and financial stability, *Speech given at Lloyd's of London (29 September).*

Carr, P., and L. Wu, 2009, Variance risk premiums, *Review of Financial Studies* 22, 1311–1341.

Choi, Darwin, Zhenyu Gao, and Wenxi Jiang, 2018, Attention to global warming, *Working Paper.*

Dessaint, Olivier, and Adrien Matray, 2017, Do managers overreact to salient risks? evidence from hurricane strikes, *Journal of Financial Economics* 126, 97–121.

Engle, Robert, Stefano Giglio, Bryan Kelly, Heebum Lee, and Johannes Stroebel, 2018, Hedging climate change news, *Working Paper.*

Fama, Eugene F., and Kenneth R. French, 1993, Common risk factors in the returns on stocks and bonds, *Journal of Financial Economics* 33, 3–56.

Giglio, Stefano, Matteo Maggiori, Krishna Rao, Johannes Stroebel, and Andreas Weber, 2018, Climate change and long-run discount rate: Evidence from real estate, *Working Paper.*

Gillingham, Kenneth, William D. Nordhaus, David Anthoff, Geoffrey Blanford, Valentina Bosetti, Peter Christensen, Haewon McJeon, and John Reilly, 2018, Modeling uncertainty in integrated assessment of climate change: A multimodel comparison, *Journal of the Association of Environmental and Resource Economists* 5, 791–826.

Hong, Harrison, Frank W. Li, and Jiangmin Xu, 2016, Climate risks and market efficiency, *Journal of Econometrics forthcoming.*

Imberman, Scott A., Adriana D. Kugler, and Bruce I. Sacerdote, 2012, Katrina's children: Evidence on the structure of peer effects from hurricane evacuees, *American Economic Review* 102, 2048–2082.

Jens, Candace E., 2017, Political uncertainty and investment: Causal evidence from U.S. gubernatorial elections, *Journal of Financial Economics* 124, 563–579.

Julio, Brandon, and Youngsuk Yook, 2012, Political uncertainty and corporate investment cycles, *Journal of Finance* 67, 45–83.

Jurado, Kyle, Sydney C. Ludvigson, and Serena Ng, 2015, Measuring uncertainty, *American Economic Review* 105, 1177–1216.

Kelly, Bryan, Lubos Pastor, and Pietro Veronesi, 2016, The price of political uncertainty: Theory and evidence from the option market, *Journal of Finance* 71, 2417–2480.

Martin, Ian, and Christian Wagner, 2018, What is the expected return on a stock?, *Working Paper.*

Martinez, Andrew, 2018, A false sense of security: The impact of forecast uncertainty on hurricane damages, *Working Paper.*

Melillo, Jerry M., Terese Richmond, and Gary W. Yohe, 2014, Climate change impacts in the United States: The third national climate assessment, U.S. Global Change Research Program.

Murfin, Justin, and Matthew Spiegel, 2018, Is the risk of sea level rise capitalized in residential real estate?, *Review of Financial Studies forthcoming.*

Neumark, David, Brandon Wall, and Junfu Zhang, 2011, Do small businesses create more jobs? New evidence for the united states from the national establishment time series, *Review of Economics and Statistics* 93, 16–29.

Nordhaus, William D., 2014, Estimates of the social cost of carbon: Concepts and results from the DICE-2013R model and alternative approaches, *Journal of the Association of Environmental and Resource Economists* 1, 273–312.

——— , and Zili Yang, 1996, A regional dynamic general-equilibrium model of alternative climate-change strategies, *American Economic Review* 86, 741–765.

Pastor, Lubos, and Pietro Veronesi, 2012, Uncertainty about government policy and stock prices, *Journal of Finance* 67, 1219–1264.

——— , 2013, Political uncertainty and risk premia, *Journal of Financial Economics* 110, 520–545.

Roll, Richard, 1984, Orange juice and weather, *American Economic Review* 74, 861–880.

Roth Tran, Brigitte, 2018, Philanthropic endowment investments in objectionable firms, *American Economic Review: Insights forthcoming.*

——— , and Daniel J. Wilson, 2018, The local economic impact of natural disasters, *Working Paper.*

Tol, Richard S.J., 1997, On the optimal control of carbon dioxide emissions: An application of fund, *Environmental Modeling and Assessment* 2, 151–163.

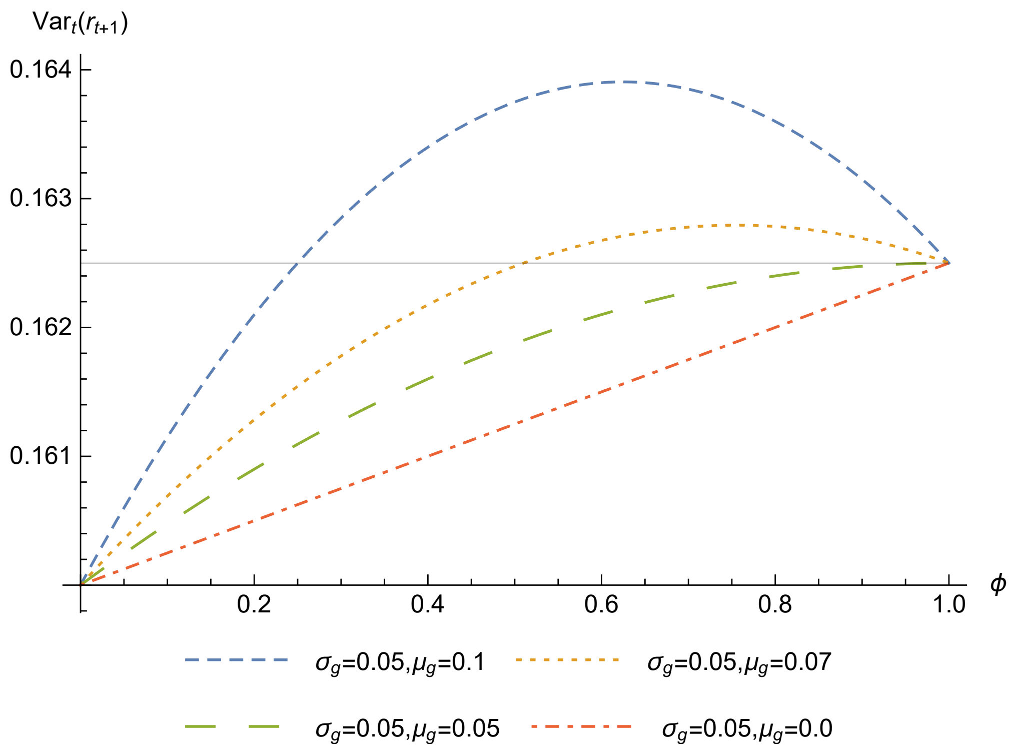


Figure 1: Variance as a function of the probability of hurricane landfall

This figure shows the total variance prior to landfall, $Var_t(r_{i,t+1})$ derived in Eqn. (6), as the probability of landfall, ϕ, varies from 0 to 1. In this figure, $\sigma = 0.4$ and $\sigma_g = 0.05$. The four dashed lines have μ_g values of 0.1, 0.07, 0.05, and 0. The solid line shows the level of variance conditional on hurricane landfall, $Var_t(r_{i,t+1}|\theta = 1) = \sigma^2 + \sigma_g^2$, as defined in Eqn. (3).

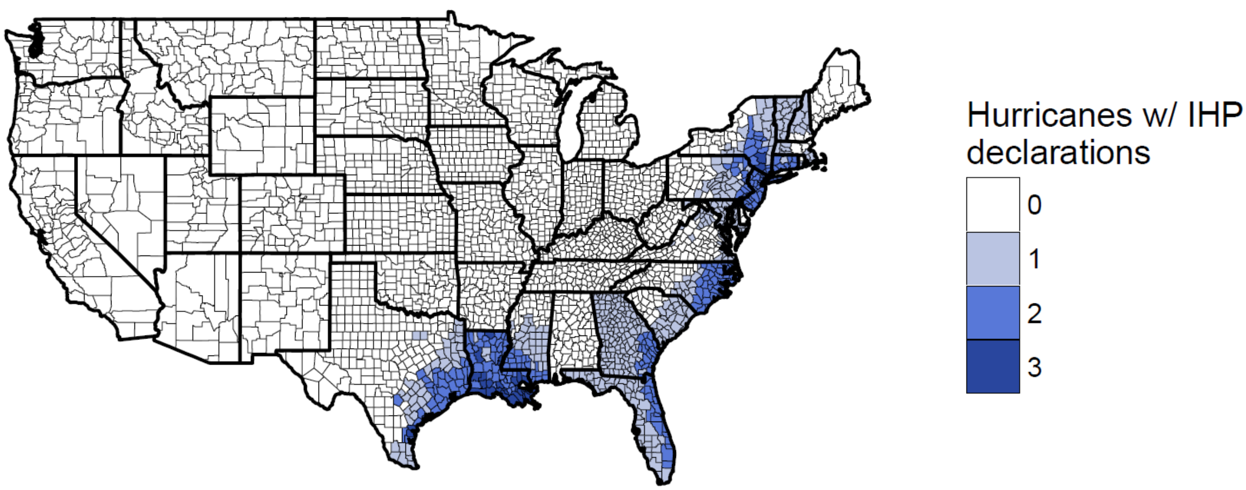


Figure 2: US counties with hurricane damage.

This figure highlights the counties with at least one IHP declarations during the sample period from 2007 to 2017. The map is constructed using data from FEMA.

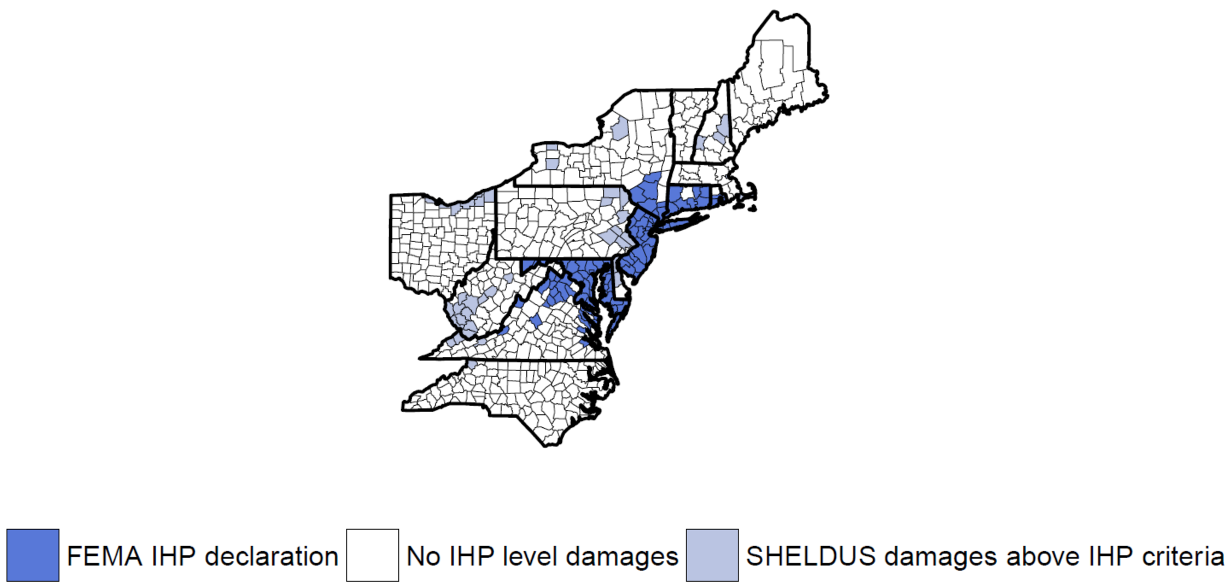


Figure 3: Counties with damage from Hurricane Sandy.

This figure highlights the counties with IHP-level damages from Hurricane Sandy in 2012. The map is constructed using data from FEMA and SHELDUS.

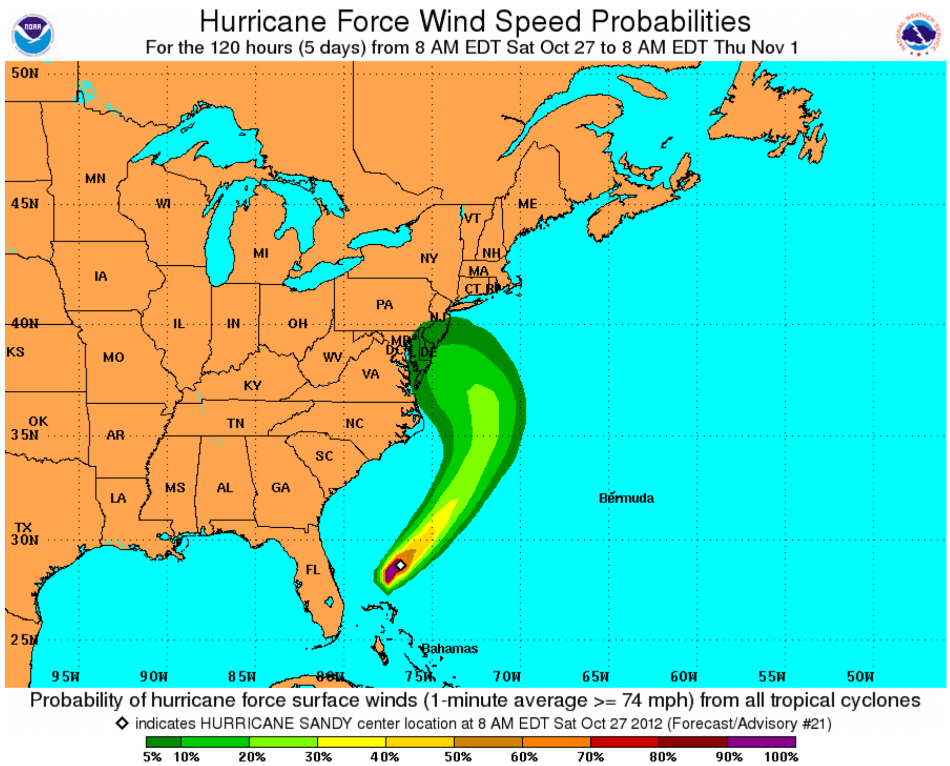


Figure 4: Example of a five-day forecast of a hurricane.

This figure from NOAA illustrates the five-day forecast for Hurricane Sandy on October 27, 2012.

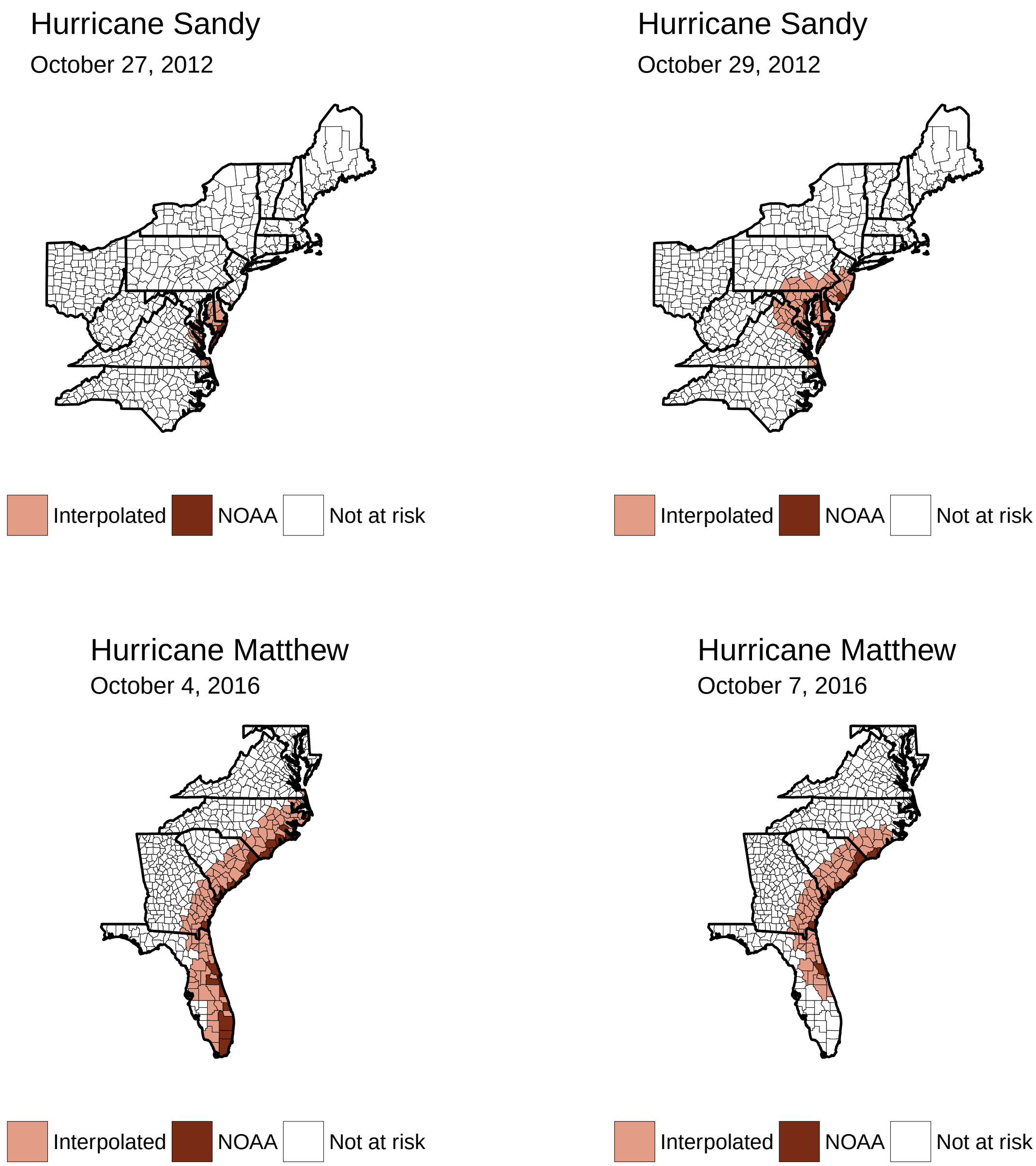
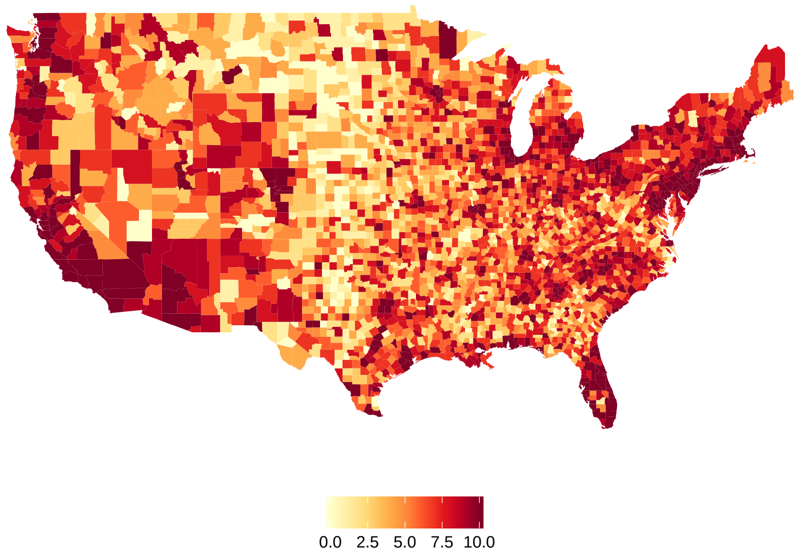
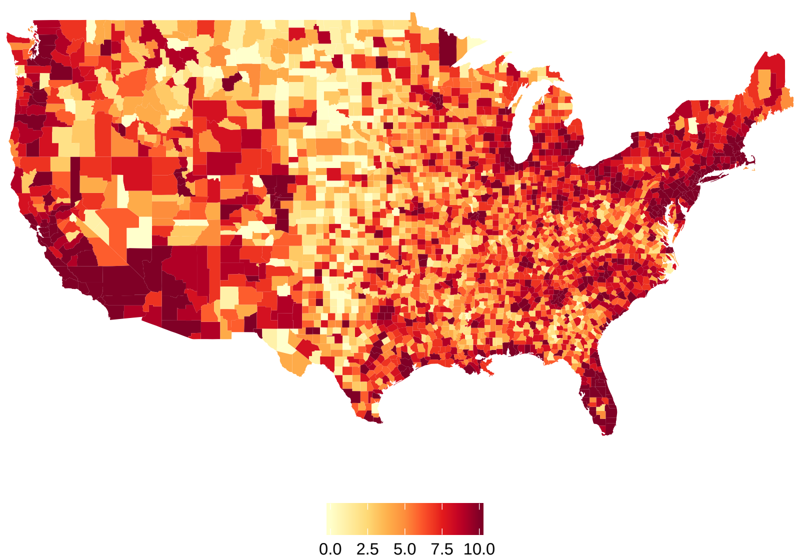


Figure 5: Wind forecasts over a five-day horizon.

This figure highlights counties with a forecasted cumulative probability of at least 10 percent for wind speeds of 64 knots or higher over a five-day horizon. The darker shade denote counties with direct wind speed observations. The lighter shade denotes counties within a 75-mile radius of these direct observations. The counties untreated at these levels are not shaded. The maps are generated after processing raw wind speed forecast data obtained from NOAA.



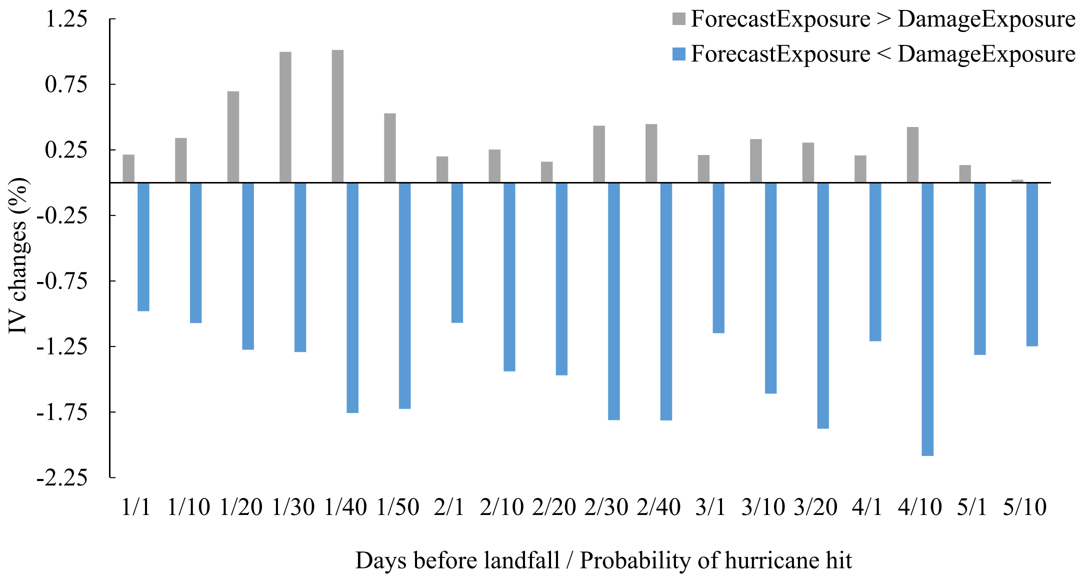
(a) Year 2010



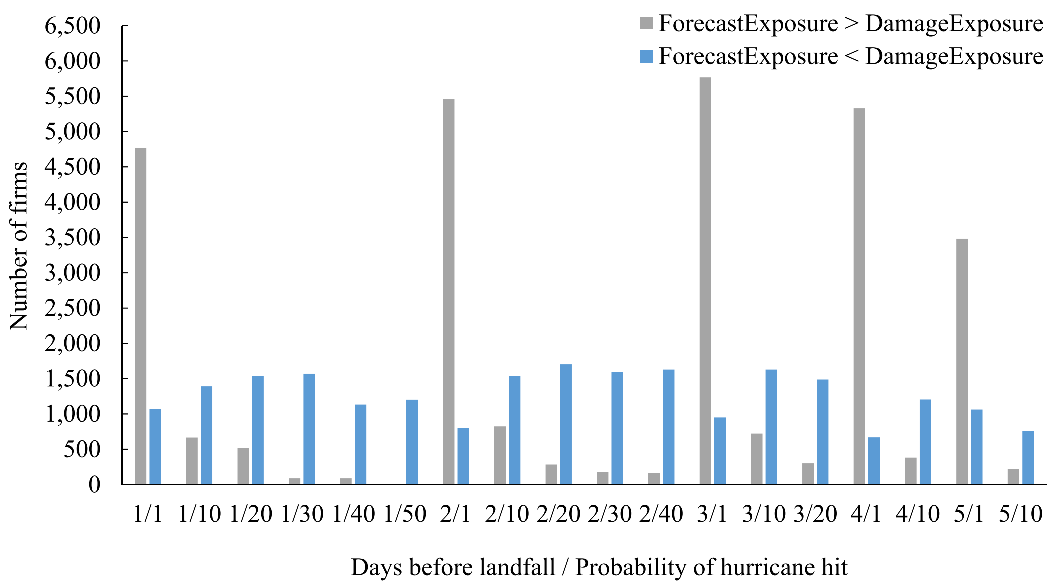
(b) Year 2014

Figure 6: Firm establishments by county

This chart plots counties based on the number of establishments located in that county for the years 2010 (Panel A) and 2014 (Panel B). The data are from NETS. Only firms that could be mapped to CRSP-Compustat are included. The counties are sorted into deciles based on the number of establishments.



(a) Change in implied volatility



(b) Number of firms

Figure 7: Economic effects of differences in forecast and damage exposure

Panel A plots by how much the implied volatility is too high (low) on average due to an overprediction (underprediction), i.e., the firm's forecast exposure to a hurricane is greater (smaller) than the firm's eventual damage exposure. Panel B plots the number of firms for which the forecast exposure to a hurricane is greater (smaller) than the firm's eventual damage exposure summed across all the hurricanes from 2007 to 2017. The analysis uses up to five calendar days before the landfall or dissipation day and six probability thresholds. The geographic footprint of firms is based on the number of establishments in counties. A combination of days before landfall (dissipation) and probability threshold is only included if there exist at least three hurricanes and 25 firms with an exposure of over 20 percent in establishments to the counties in the forecasted path. The damage exposure is measured with FEMA and SHELDUS data.

Table 1: Summary statistics of hurricane damage data

This table reports summary statistics on the damage data for each hurricane from 1995 to 2017 based on FEMA and SHELDUS. Counties are eligible for individual and household program (IHP) assistance if there is a FEMA disaster declaration due to significant damage from a hurricane on a per capita basis.

	Mean	Median	Std Dev	Min	Max	Obs
FEMA Counties with IHP declarations	80	77	48	16	254	26
SHELDUS Counties with property damage	180	151	121	24	466	26
SHELDUS Counties with IHP-worthy damages	103	84	83	5	345	26
SHELDUS Property damages estimate ($millions)	10,042	3,408	21,833	32	89,432	26

Table 2: Counts of storms in forecast data

This table reports the number of storms in NOAA forecasts from 2007 to 2017, which are forecasted to make landfall with windspeeds of at least 64KT with a given minimum probability within a specified time period.

Probability	Time Horizon	Number of Storms
$\geq 1\%$	5 days	59
$\geq 10\%$	5 days	23
$\geq 50\%$	5 days	11
$\geq 50\%$	24 hours	11

Table 3: Firm establishment and sales summary statistics

This table reports the summary statistics on the number of establishments and amount of sales (in USD) in the NETS dataset from 1995 to 2017 for the firms that were matched to equity data from CRSP-Compustat.

Number of firms with establishment/sales data					
Establishments	4,999				
Sales	4,990				
Statistics by firm-year					
	Avg	SD	10% percentile	50% percentile	90% percentile
Establishments	59.744	284.307	1.000	4.000	95.000
Sales (in millions)	464.744	2,277.302	0.447	29.270	816.455
Statistics by county-year					
	Avg	SD	10% percentile	50% percentile	90% percentile
Establishments	61.436	198.227	2.000	12.000	128.000
Sales (in millions)	477.512	1,984.836	2.090	52.356	847.390
Statistics by county-year for hurricane damaged counties					
	Avg	SD	10% percentile	50% percentile	90% percentile
Establishments	82.237	232.929	3.000	19.000	204.000
Sales (in millions)	633.887	2,701.185	2.656	85.089	1,204.906

Table 4: Summary statistics of implied volatility

This table reports the summary statistics on the options data from OptionMetrics. Panel A includes the dataset once merged with CRSP-Compustat. Panel B further restricts the sample to firms appearing at least once in the NETS firm establishment data.

Panel A: Firms matched to CRSP-Compustat

	N	Mean	Median	Stdev	25th	75th	10th	90th
$IV_{i,t}$	7,089,684	0.459	0.384	0.284	0.271	0.557	0.205	0.796
$log\left(\frac{IV_{i,t}}{IV_{i,t-1}}\right)$	6,938,562	0.001	0.000	0.132	-0.045	0.049	-0.113	0.119
Days to expiry$_{i,t}$	7,089,684	39.359	29.000	36.582	17.000	40.000	10.000	95.000
Total open interest$_{i,t}$	7,089,684	2,437.218	270.000	12,107.768	53.000	1,368.000	13.000	5,169.000

Panel B: Firms matched to CRSP-Compustat and NETS

	N	Mean	Median	Stdev	25th	75th	10th	90th
$IV_{i,t}$	2,937,420	0.447	0.373	0.279	0.266	0.539	0.202	0.774
$log\left(\frac{IV_{i,t}}{IV_{i,t-1}}\right)$	2,874,389	0.001	0.000	0.135	-0.046	0.049	-0.115	0.121
Days to expiry$_{i,t}$	2,937,420	39.732	29.000	36.653	17.000	40.000	10.000	95.000
Total open interest$_{i,t}$	2,937,420	2,111.815	234.000	7,933.214	48.000	1,210.000	12.000	4,614.000

Table 5: Hurricane effects on implied volatility

This table reports the coefficients and test statistics when estimating the panel model in equation (9). The dependent variable is the change (in percent) in the implied volatility of firm i from the day before the inception day of the hurricane T_h^* until 5 trading days after the landfall T_h. The independent variable measures how much of the geographic footprint of a firm is exposed to the disaster area. For Panel A, the geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county, and for Panel B, the geographic footprint is based on sales per county. To identify counties that have been damaged by a hurricane we use FEMA data and FEMA data enhanced with SHELDUS data. The data are from 2002 to 2017. Results are also shown for the subsample from 2007 to 2017, which corresponds to the time period for which we have hurricane forecast data used in the subsequent analysis. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). Industry and time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Dependent variable: Change in IV, $log\left(IV_{i,T_h+5}/IV_{i,T_h^*}\right)$

	2002-2017				2007-2017			
	FEMA		FEMA+SHELDUS		FEMA		FEMA+SHELDUS	
$DamageExposure_{i,T_h}$	0.064***	0.045***	0.064***	0.050***	0.073***	0.052***	0.075***	0.060***
	4.032	2.809	4.076	3.123	3.958	2.535	3.904	2.907
Adjusted R^2 (%)	14.203	14.971	14.343	15.117	15.711	16.703	15.735	16.734
Obs. total	12,059	12,059	12,524	12,524	7,107	7,107	7,107	7,107
Obs. firm exposure $> 0\%$	5,347	5,347	5,896	5,896	2,980	2,980	3,219	3,219
Obs. firm exposure $\geq 20\%$	798	798	966	966	472	472	585	585
Obs. firm exposure $\geq 50\%$	284	284	315	315	173	173	197	197
Hurricanes	19	19	20	20	10	10	10	10
Industry FE	Yes	No	Yes	No	Yes	No	Yes	No
Time FE	Yes	No	Yes	No	Yes	No	Yes	No
Industry × Time FE	No	Yes	No	Yes	No	Yes	No	Yes

Panel B: Firms' hurricane exposure based on sales

Dependent variable: Change in IV, $log\left(IV_{i,T_h+5}/IV_{i,T_h^*}\right)$

	2002-2017				2007-2017			
	FEMA		FEMA+SHELDUS		FEMA		FEMA+SHELDUS	
$DamageExposure_{i,T_h}$	0.048***	0.035***	0.047***	0.037***	0.058***	0.043***	0.058***	0.048***
	3.495	2.583	3.675	2.978	3.837	2.693	3.940	3.111
Adjusted R^2 (%)	14.204	14.990	14.334	15.124	15.714	16.718	15.730	16.741
Obs. total	12,029	12,029	12,493	12,493	7,097	7,097	7,097	7,097
Obs. firm exposure $> 0\%$	5,325	5,325	5,874	5,874	2,966	2,966	3,207	3,207
Obs. firm exposure $\geq 20\%$	800	800	945	945	480	480	587	587
Obs. firm exposure $\geq 50\%$	388	388	444	444	232	232	276	276
Hurricanes	19	19	20	20	10	10	10	10
Industry FE	Yes	No	Yes	No	Yes	No	Yes	No
Time FE	Yes	No	Yes	No	Yes	No	Yes	No
Industry × Time FE	No	Yes	No	Yes	No	Yes	No	Yes

Table 6: Abnormal returns post hurricane landfall

This table reports differences in cumulative abnormal returns post landfall between firms with exposure and firms without exposure to the hurricane disaster region and corresponding t-stats. The differences are reported for the mean and nine quantiles. The abnormal returns are estimated based on the Fama-French three factor model. FEMA and SHELDUS data are used to identify counties that have been hit by a hurricane. For Panel A, the hit firms are defined as firms that have 50 percent or more of their establishments in the counties of the disaster region, and for Panel B, 50 percent or more of the sales have to be located in the disaster area counties. We exclude hurricanes that affected less than 25 firms. The data are from 1995 to 2017. The standard errors are bootstrapped and clustered by county (headquarter location). The significance of the difference in abnormal returns is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms selected based on establishments in disaster region $\geq 50\%$

	0 to 5 days post landfall		0 to 120 days post landfall	
	Cumulative r diff.	T-stat	Cumulative r diff	T-stat
Mean	-0.159	-0.298	-3.715	-0.970
Quantiles				
10%	-0.396	-0.585	-12.119^{**}	-2.072
20%	-0.453	-1.324	-10.576^{***}	-2.587
30%	-0.243	-1.023	-8.281^{*}	-1.842
40%	-0.256	-1.466	-5.676^{**}	-2.028
50%	-0.315	-1.388	-5.469^{**}	-2.091
60%	-0.253	-1.181	-3.955	-1.162
70%	-0.135	-0.440	-1.489	-0.606
80%	-0.156	-0.351	0.957	0.285
90%	1.308	1.198	3.044	0.372
Hit firms	566	566	567	567
Control firms	13,334	13,334	13,336	13,336

Panel B: Firms selected based on sales in disaster region $\geq 50\%$

	0 to 5 days post landfall		0 to 120 days post landfall	
	Cumulative r diff.	T-stat	Cumulative r diff	T-stat
Mean	-0.350	-0.661	-6.216*	-1.844
Quantiles				
10%	-0.987	-1.818*	-14.717^{*}	-1.916
20%	-0.430	-1.293	-10.641^{**}	-2.324
30%	-0.269	-1.124	-5.948	-1.545
40%	-0.311	-1.276	-4.505^{**}	-1.979
50%	-0.225	-0.721	-3.861^{*}	-1.682
60%	-0.030	-0.128	-3.694^{*}	-1.714
70%	0.155	0.373	-2.189	-0.880
80%	0.025	0.065	-1.029	-0.320
90%	0.560	0.650	-2.686	-0.511
Hit firms	804	804	805	805
Control firms	13,977	13,977	13,979	13,979

Table 7: Forecasted hurricane path and implied volatility

This table reports the coefficients and test statistics when estimating the panel model in equation (12). The dependent variable is the change (in percent) in the implied volatility of firm i from inception of the hurricane to Γ days before landfall or dissipation, T_h, of the hurricane. The independent variable measures how much (in percent) of the geographic footprint of a firm is exposed to the forecasted path of a hurricane Γ days before the landfall or dissipation of the hurricane. The geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county for Panel A and based on sales per county for Panel B. The forecasted path of the hurricane is from NOAA and gives a probability of being hit by a specific hurricane for each county. The included probability thresholds have at least three hurricanes and 25 firms with an exposure of over 20% in establishments or sales to the counties in the forecasted path. The data are from 2007 to 2017. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). Industry and time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Dependent variable: Change in IV from hurricane inception to Γ days before landfall/dissipation, $log\left(IV_{i,T_h-\Gamma}/IV_{i,T_h^*}\right)$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit $\geq$	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	0.038***	0.064***	0.095***	0.155***	0.159***	0.020*	0.072***	0.075***	0.176***	0.012*	0.106***	0.153***	0.016*	0.138***	0.012	0.029
	(3.230)	(3.711)	(5.675)	(5.717)	(6.168)	(1.960)	(4.493)	(4.108)	(5.608)	(1.911)	(4.026)	(6.473)	(1.840)	(3.701)	(1.310)	(0.592)
Adjusted R^2	13.977	18.456	17.974	22.855	22.816	10.752	13.304	14.528	17.429	9.798	13.534	14.318	14.797	24.432	9.098	12.066
Obs. total	21,949	6,525	5,741	2,587	2,587	16,699	7,538	5,881	4,315	14,795	7,460	5,099	12,313	4,172	10,568	2,440
Obs. ForecastExposure > 0%	5,928	1,832	1,525	763	696	6,475	1,978	1,397	979	6,893	2,058	1,354	6,243	1,143	4,651	680
Obs. ForecastExposure $\geq$ 20%	433	185	132	75	64	1,325	174	123	71	2,036	156	96	1,932	90	1,136	36
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Firms' hurricane exposure based on sales

Dependent variable: Change in IV from hurricane inception to Γ days before landfall/dissipation, $log\left(IV_{i,T_h-\Gamma}/IV_{i,T_h^*}\right)$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit $\geq$	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	0.030**	0.052***	0.075***	0.133***	0.148***	0.020**	0.068***	0.082***	0.155***	0.014*	0.090***	0.142***	0.015*	0.123***	0.009	0.038
	(2.253)	(4.303)	(3.942)	(4.629)	(5.807)	(2.083)	(4.679)	(5.004)	(4.626)	(2.655)	(2.784)	(5.261)	(2.064)	(3.316)	(1.125)	(1.172)
Adjusted R^2	14.009	18.489	18.023	22.958	22.962	10.770	13.329	14.583	17.490	9.823	13.573	14.447	14.863	24.484	9.078	12.043
Obs. total	21,921	6,516	5,733	2,584	2,587	16,679	7,529	5,874	4,310	14,776	7,451	5,093	12,297	4,167	10,555	2,437
Obs. ForecastExposure > 0%	5,856	1,821	1,514	759	690	6,443	1,962	1,386	969	6,867	2,040	1,346	6,214	1,134	4,630	673
Obs. ForecastExposure $\geq$ 20%	483	201	149	82	69	1,254	183	130	80	1,835	152	102	1,713	98	1,003	40
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 8: Alternative specification for forecasted hurricane path and implied volatility

This table reports the sample mean, denoted $\overline{IVD}$, and t-stat for $IVD_{i,h}$ described in equation (13). For each hurricane and firm i, we subtract the implied volatility (IV) on the day before the inception of the hurricane from the IV on the day when the firm's exposure to counties in the hurricane's path exceeds a threshold of X percent. Landfall days are excluded. The log difference in IV is then demeaned by the mean log difference in IV of firms that are not exposed to the hurricane. The geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county for Panel A and based on sales per county for Panel B. The forecasted path of the hurricane is from NOAA and gives a probability of being hit by a specific hurricane for each county. The data are from 2007 to 2017. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Exposure to hurricane path ≥	10%					25%					50%				
Probability of hurricane hit ≥	1%	10%	20%	40%	50%	1%	10%	20%	40%	50%	1%	10%	20%	40%	50%
$\overline{IVD}$	-0.348	1.738**	2.450**	6.519*	8.529**	0.279	3.489***	4.656***	11.122***	12.495***	0.985**	2.178**	3.376***	9.404***	11.311***
	(-1.475)	(2.213)	(2.293)	(1.715)	(2.538)	(0.728)	(3.130)	(4.135)	(3.045)	(3.606)	(2.065)	(2.094)	(2.836)	(3.269)	(5.835)
Number of firms	6,088	649	414	178	145	2,775	244	165	68	61	1,282	118	81	32	27
Number of hurricanes	40	13	9	5	5	39	11	7	5	5	37	11	7	4	4

Panel B: Firms' hurricane exposure based on sales

Exposure to hurricane path ≥	10%					25%					50%				
Probability of hurricane hit ≥	1%	10%	20%	40%	50%	1%	10%	20%	40%	50%	1%	10%	20%	40%	50%
$\overline{IVD}$	-0.135	1.887**	2.877**	7.494*	12.217***	0.199	2.466**	4.783***	10.633***	11.634***	0.440	2.721**	3.882***	9.096**	11.348***
	(-0.499)	(1.996)	(2.282)	(1.827)	(3.740)	(0.610)	(1.961)	(7.899)	(2.892)	(4.276)	(1.204)	(2.545)	(5.498)	(2.418)	(3.499)
Firms	4,788	538	354	160	132	2,792	279	193	84	73	1,706	174	122	57	48
Number of hurricanes	39	14	10	6	6	37	12	9	5	5	37	12	8	5	5

Table 9: Forecasted hurricane path and stock returns

This table reports the coefficients and test statistics when estimating the panel model in equation (14). The dependent variable is the stock return (in percent) of firm i from inception of the hurricane to Γ days before landfall or dissipation of the hurricane. The independent variable measures how much (in percent) of the geographic footprint of a firm is exposed to the forecasted path of a hurricane Γ days before the landfall or disappearance of the hurricane. The geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county for Panel A and based on sales per county for Panel B. The forecasted path of the hurricane is from NOAA and gives a probability of being hit by a specific hurricane for each county. The included probability thresholds have at least three hurricanes and 25 firms with an exposure of over 20 percent in establishments or sales to the counties in the forecasted path. The data are from 2007 to 2017. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). Industry and time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Dependent variable: Cumulative return from hurricane inception to Γ days before landfall/dissipation, $r_{i,T_h^*:T_h-\Gamma}$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit $\geq$	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	-0.014	-0.025***	-0.039***	-0.076***	-0.084***	-0.005	-0.043***	-0.047***	-0.089***	-0.003	-0.048**	-0.058**	-0.007	-0.076***	-0.003	0.017
	(-1.259)	(-4.794)	(-4.201)	(-3.029)	(-3.562)	(-0.986)	(-3.781)	(-5.266)	(-3.519)	(-1.103)	(-2.494)	(-2.423)	(-1.508)	(-2.820)	(-0.769)	(1.559)
Adjusted R^2	21.217	20.658	18.506	15.559	15.630	16.224	19.952	21.044	24.342	9.905	14.981	19.010	13.099	28.808	10.970	32.036
Obs. total	24,783	7,164	6,251	2,421	2,421	17,302	7,631	6,032	4,107	15,934	8,048	5,016	13,923	4,293	11,239	2,469
Obs. ForecastExposure $> 0\%$	6,510	1,823	1,537	740	680	6,586	1,919	1,347	936	7,252	2,071	1,310	6,807	1,126	5,071	663
Obs. ForecastExposure $\geq 20\%$	560	171	129	69	62	1,381	164	114	66	2,124	155	93	1,973	86	1,344	36
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Panel B: Firms' hurricane exposure based on sales

Dependent variable: Cumulative return from hurricane inception to Γ days before landfall/dissipation, $r_{i,T_h^*:T_h-\Gamma}$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit $\geq$	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	-0.012	-0.021***	-0.031***	-0.061***	-0.069***	-0.004	-0.034***	-0.039***	-0.075***	-0.003	-0.036**	-0.048**	-0.008**	-0.062***	-0.004	0.008
	(-1.470)	(-5.401)	(-4.574)	(-3.090)	(-3.814)	(-0.952)	(-3.354)	(-5.374)	(-3.920)	(-1.260)	(-2.285)	(-2.524)	(-2.058)	(-2.792)	(-1.316)	(0.972)
Adjusted R^2	21.220	20.655	18.488	15.502	15.592	16.222	19.923	21.036	24.348	9.907	14.906	18.971	13.127	28.811	10.982	31.982
Obs. total	24,783	7,164	6,251	2,421	2,421	17,302	7,631	6,032	4,107	15,934	8,048	5,016	13,923	4,293	11,239	2,469
Obs. ForecastExposure $> 0\%$	6,441	1,811	1,526	737	675	6,554	1,903	1,336	927	7,229	2,055	1,303	6,780	1,118	5,051	656
Obs. ForecastExposure $\geq 20\%$	632	198	153	86	72	1,352	175	125	78	1,977	164	111	1,841	103	1,247	42
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 10: Market and NOAA forecasts comparison

This table reports the coefficients and test statistics when estimating the panel model in equation (15). The dependent variable is the change (in percent) in the implied volatility of firm i from inception of the hurricane to Γ days before landfall or dissipation, T_h, of the hurricane. The independent variable $HurricaneForecastExposure_{i,T_h-\Gamma}$ measures how much (in percent) of the geographic footprint of a firm is exposed to the forecasted path of a hurricane Γ days before the landfall or dissipation of the hurricane. The independent variable $UnderPrediction_{i,T_h-\Gamma}$ measures the difference between the firm's exposure to the eventual hurricane damage region and the firm's exposure to the forecasted path of a hurricane, as shown in equation 16. The geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county for Panel A and based on sales per county for Panel B. The forecasted path of the hurricane is from NOAA and gives a probability of being hit by a specific hurricane for each county. The included probability thresholds have at least three hurricanes and 25 firms with an exposure of over 20% in establishments or sales to the counties in the forecasted path. The data are from 2007 to 2017. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). Industry and time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Dependent variable: Change in IV from hurricane inception to Γ days before landfall/dissipation, $log\left(IV_{i,T_h-\Gamma}/IV_{i,T_h^*}\right)$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit $\geq$	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	0.040***	0.053***	0.089***	0.146***	0.153***	0.016	0.073***	0.072***	0.174***	0.012*	0.105***	0.142***	0.012	0.131***	0.010	0.005
	(3.507)	(4.325)	(5.722)	(4.658)	(4.834)	(1.528)	(4.738)	(4.591)	(5.728)	(1.906)	(3.957)	(5.851)	(1.345)	(3.388)	(1.008)	(0.116)
$UnderPrediction_{i,T_h-\Gamma}$	0.030	0.055*	0.021	0.016	0.019	-0.005	0.016	0.024	0.009	-0.024	0.017	0.022	0.019	0.038	0.027	0.004
	(0.875)	(1.749)	(0.818)	(0.581)	(0.666)	(-0.094)	(0.579)	(0.813)	(0.352)	(-0.734)	(0.791)	(0.921)	(0.420)	(1.429)	(1.115)	(0.167)
Adjusted R^2	13.679	18.100	17.220	20.527	20.510	10.410	13.245	14.738	17.316	9.471	12.549	13.338	14.443	22.693	9.181	8.389
Obs. total	21,949	6,525	5,741	2,587	2,587	16,699	7,538	5,881	4,315	14,795	7,460	5,099	12,313	4,172	10,568	2,440
Obs. ForecastExposure > 0%	5,928	1,832	1,525	763	696	6,475	1,978	1,397	979	6,893	2,058	1,354	6,243	1,143	4,651	680
Obs. ForecastExposure $\geq$ 20%	433	185	132	75	64	1,325	174	123	71	2,036	156	96	1,932	90	1,136	36
Obs. UnderPrediction	1,068	1,391	1,535	1,132	1,202	798	1,536	1,703	1,629	1,629	876	1,487	669	1,205	1,062	758
Obs. UnderPrediction $\geq$ 20%	84	128	189	182	189	67	222	251	312	101	221	237	79	232	158	156
Obs. UnderPrediction $\geq$ 50%	27	42	65	61	61	23	79	87	108	31	75	79	24	75	69	54
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 10: Market and NOAA forecasts comparison (continued)

Panel B: Firms' hurricane exposure based on sales

Dependent variable: Change in IV from hurricane inception to Γ days before landfall/dissipation, $log\left(IV_{i,T_h-\Gamma}/IV_{i,T_h^*}\right)$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit ≥	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	0.033***	0.046***	0.073***	0.127***	0.142***	0.016*	0.068***	0.079***	0.153***	0.013***	0.093***	0.136***	0.013*	0.123***	0.008	0.025
	(2.828)	(4.487)	(4.036)	(3.851)	(4.609)	(1.787)	(5.009)	(5.289)	(4.507)	(2.581)	(2.940)	(4.932)	(1.693)	(3.417)	(0.999)	(0.807)
$UnderPrediction_{i,T_h-\Gamma}$	0.004	0.024	0.005	0.001	0.003	-0.032	0.003	0.005	-0.001	-0.031	0.013	0.011	0.011	0.029	0.026	0.004
	(0.137)	(0.847)	(0.243)	(0.042)	(0.126)	(-0.722)	(0.116)	(0.208)	(-0.068)	(-1.002)	(0.698)	(0.513)	(0.315)	(1.345)	(1.501)	(0.216)
Adjusted R^2	13.715	18.119	17.274	20.641	20.658	10.431	13.269	14.780	17.376	9.504	12.597	13.463	14.503	22.774	9.175	8.370
Obs. total	21,921	6,516	5,733	2,584	2,584	16,679	7,529	5,874	4,310	14,776	7,451	5,093	12,297	4,167	10,555	2,437
Obs. ForecastExposure > 0%	5,856	1,821	1,514	759	690	6,443	1,962	1,386	969	6,867	2,040	1,346	6,214	1,134	4,630	673
Obs. ForecastExposure ≥ 20%	483	201	149	82	69	1,254	183	130	80	1,835	152	102	1,713	98	1,003	40
Obs. UnderPrediction	1,098	1,387	1,526	1,130	1,199	859	1,531	1,695	1,622	980	1,643	1,481	739	1,203	1,118	757
Obs. UnderPrediction ≥ 20%	117	164	216	213	218	100	246	277	325	139	234	246	100	230	200	141
Obs. UnderPrediction ≥ 50%	51	67	100	96	99	45	115	123	153	56	106	106	43	104	93	67
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 11: Market and NOAA forecasts comparison accounting for institutional ownership

This table reports the coefficients and test statistics when estimating the panel model in equation (15) with an institutional ownership interaction term. The dependent variable is the change (in percent) in the implied volatility of firm i from inception of the hurricane to Γ days before landfall or dissipation, T_h, of the hurricane. The independent variable $HurricaneForecastExposure_{i,T_h-\Gamma}$ measures how much (in percent) of the geographic footprint of a firm is exposed to the forecasted path of a hurricane Γ days before the landfall or dissipation of the hurricane. The independent variable $UnderPrediction_{i,T_h-\Gamma}$ measures the difference between the firm's exposure to the eventual hurricane damage region and the firm's exposure to the forecasted path of a hurricane, as shown in equation 16. The independent variable $InstShareIndicator_{i,T_h-\Gamma}$ takes the value one if the institutional ownership of the underlying stock is above or equal to the sample mean at the quarter end before the inception of the hurricane. The geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county for Panel A and based on sales per county for Panel B. The forecasted path of the hurricane is from NOAA and gives a probability of being hit by a specific hurricane for each county. The included probability thresholds have at least three hurricanes and 25 firms with an exposure of over 20% in establishments or sales to the counties in the forecasted path. The data are from 2007 to 2017. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). Industry and time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Dependent variable: Change in IV from hurricane inception to Γ days before landfall/dissipation, $log\left(IV_{i,T_h-\Gamma}/IV_{i,T_h^*}\right)$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit $\geq$	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	0.040***	0.052***	0.087***	0.148***	0.156***	0.014	0.061***	0.064***	0.166***	0.008	0.090***	0.122***	0.007	0.124***	0.009	-0.003
	(3.154)	(4.263)	(4.935)	(3.758)	(3.870)	(1.293)	(4.182)	(4.337)	(4.745)	(1.180)	(2.922)	(3.876)	(0.911)	(3.229)	(0.850)	(-0.050)
$UnderPrediction_{i,T_h-\Gamma}$	0.052	0.070	0.034	0.026	0.028	0.033	0.036	0.038	0.012	-0.027	0.037	0.033	0.020	0.025	0.063**	0.038
	(0.926)	(1.595)	(0.867)	(0.581)	(0.642)	(0.471)	(1.025)	(1.131)	(0.376)	(-0.761)	(1.282)	(1.035)	(0.345)	(0.802)	(2.122)	(1.181)
$InstShareIndicator_{i,T_h-\Gamma}$	-0.260	0.418	0.770	-0.142	-0.130	-0.486*	-0.390	-0.338	-0.015	-0.556**	-0.399	-0.086	-0.797**	-0.492	-0.295	-0.062
	(-1.079)	(0.691)	(1.236)	(-0.138)	(-0.127)	(-1.777)	(-0.903)	(-0.601)	(-0.019)	(-1.968)	(-0.829)	(-0.141)	(-2.376)	(-0.862)	(-1.052)	(-0.098)
$UnderPrediction_{i,T_h-\Gamma}\times$	-0.025	-0.016	-0.017	-0.005	-0.005	-0.050	-0.037	-0.037	-0.014	0.038	-0.018	-0.004	0.054	0.021	-0.054	-0.066
$InstShareIndicator_{i,T_h-\Gamma}$	(-0.309)	(-0.267)	(-0.356)	(-0.088)	(-0.097)	(-0.734)	(-0.904)	(-0.907)	(-0.355)	(0.765)	(-0.465)	(-0.097)	(0.637)	(0.446)	(-1.241)	(-1.453)
Adjusted R^2	13.769	17.835	16.845	20.188	20.173	10.376	12.991	14.534	16.932	9.260	12.231	12.914	14.296	22.550	9.197	8.255
Obs. total	20,311	6,034	5,320	2,447	2,447	15,552	7,050	5,480	4,062	13,713	6,892	4,773	11,428	3,889	9,835	2,281
Obs. ForecastExposure > 0%	5,928	1,832	1,525	763	696	6,475	1,978	1,397	979	6,893	2,058	1,354	6,243	1,143	4,651	680
Obs. ForecastExposure $\geq$ 20%	433	185	132	75	64	1,325	174	123	71	2,036	156	96	1,932	90	1,136	36
Obs. UnderPrediction	1,068	1,391	1,535	1,132	1,202	798	1,536	1,703	1,629	950	1,629	1,487	669	1,205	1,062	758
Obs. UnderPrediction $\geq$ 20%	84	128	189	182	189	67	222	251	312	101	221	237	79	232	158	156
Obs. UnderPrediction $\geq$ 50%	27	42	65	61	61	23	79	87	108	31	75	79	24	75	69	54
Number of hurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 11: Market and NOAA forecasts comparison accounting for institutional ownership (continued)

Panel B: Firms' hurricane exposure based on sales

Dependent variable: Change in IV from hurricane inception to Γ days before landfall/dissipation, $log\left(IV_{i,T_h-\Gamma}/IV_{i,T_h^*}\right)$

Γ	1 Day					2 Days				3 Days			4 Days		5 Days	
Probability of hurricane hit ≥	1%	10%	20%	40%	50%	1%	10%	20%	40%	1%	10%	20%	1%	10%	1%	10%
$ForecastExposure_{i,T_h-\Gamma}$	0.032**	0.044***	0.070***	0.127***	0.145***	0.014	0.057***	0.071***	0.147***	0.010*	0.082**	0.124***	0.009	0.117***	0.005	0.016
	(2.506)	(3.855)	(3.453)	(3.267)	(3.886)	(1.532)	(3.995)	(5.114)	(3.910)	(1.751)	(2.279)	(3.806)	(1.248)	(3.188)	(0.651)	(0.467)
$UnderPrediction_{i,T_h-\Gamma}$	0.064	0.065**	0.033	0.028	0.030	0.023	0.028	0.022	0.006	-0.016	0.029	0.009	0.047	0.036	0.055**	0.025
	(1.498)	(2.018)	(1.022)	(0.770)	(0.802)	(0.410)	(0.942)	(0.781)	(0.191)	(-0.575)	(1.123)	(0.291)	(0.923)	(1.292)	(2.392)	(0.749)
$InstShareIndicator_{i,T_h-\Gamma}$	-0.236	0.528	0.871	0.141	0.157	-0.468*	-0.348	-0.279	0.069	-0.515*	-0.378	-0.117	-0.763**	-0.333	-0.314	-0.180
	(-0.991)	(0.876)	(1.390)	(0.136)	(0.151)	(-1.716)	(-0.809)	(-0.509)	(0.093)	(-1.824)	(-0.792)	(-0.195)	(-2.311)	(-0.612)	(-1.122)	(-0.288)
$UnderPrediction_{i,T_h-\Gamma}\times$	-0.098*	-0.069	-0.049	-0.046	-0.043	-0.087*	-0.054	-0.050	-0.029	-0.011	-0.025	0.001	-0.029	-0.025	-0.044	-0.043
$InstShareIndicator_{i,T_h-\Gamma}$	(-1.692)	(-1.461)	(-1.089)	(-0.868)	(-0.841)	(-1.776)	(-1.637)	(-1.476)	(-0.784)	(-0.233)	(-0.761)	(0.019)	(-0.467)	(-0.643)	(-1.355)	(-1.049)
Adjusted R^2	13.821	17.873	16.919	20.333	20.353	10.404	13.023	14.590	17.017	9.284	12.276	13.040	14.347	22.621	9.182	8.196
Obs. total	20,283	6,025	5,312	2,444	2,444	15,532	7,041	5,473	4,057	13,694	6,883	4,767	11,412	3,884	9,822	2,278
Obs. ForecastExposure > 0%	5,856	1,821	1,514	759	690	6,443	1,962	1,386	969	6,867	2,040	1,346	6,214	1,134	4,630	673
Obs. ForecastExposure ≥ 20%	483	201	149	82	69	1,254	183	130	80	1,835	152	102	1,713	98	1,003	40
Obs. UnderPrediction	1,098	1,387	1,526	1,130	1,199	859	1,531	1,695	1,622	980	1,643	1,481	739	1,203	1,118	757
Obs. UnderPrediction ≥ 20%	117	164	216	213	218	100	246	277	325	139	234	246	100	230	200	141
Obs. UnderPrediction ≥ 50%	51	67	100	96	99	45	115	123	153	56	106	106	43	104	93	67
NrHurricanes	27	8	7	3	3	20	9	7	5	18	9	6	15	5	13	3
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 12: Hurricane effects on implied volatility of insurance firms

This table reports the coefficients and test statistics when estimating the panel model in equation (9) for insurance firms. The dependent variable is the change (in percent) in the implied volatility of firm i from the day before the inception day of the hurricane T_h^* until 5 trading days after the landfall T_h. The independent variable measures the share of total premiums written by an insurance firm in states that experienced damage by a hurricane. For Panel A, a state is considered to have experienced hurricane damage if at least 10% of the counties experienced damage, and for Panel B, the threshold is 25% of the counties. To identify counties that have been damaged by a hurricane we use FEMA data and FEMA data enhanced with SHELDUS data. The data are from 2002 to 2017. Results are also shown for the subsample from 2007 to 2017, which corresponds to the time period for which we have hurricane forecast data used in the previous analysis. The values in parentheses are the t-stats. The standard errors are clustered by insurance firms. Time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: State considered hit if 10% or more of the counties were damaged

Dependent variable: Change in IV, $log\left(IV_{i,T_h+5}/IV_{i,T_h^*}\right)$	2002-2017		2007-2017	
	FEMA	FEMA+SHELDUS	FEMA	FEMA+SHELDUS
$StateDamageExposure_{i,T_h}$	0.348^{**}	0.328^{*}	0.357^{*}	0.360^{**}
	1.985	1.963	1.850	1.995
Adjusted R^2 (%)	36.894	34.790	33.399	33.597
Obs. total	400	418	238	238
Obs. insurance firm exposure $> 0\%$	356	374	207	207
Obs. insurance firm exposure $\geq 20\%$	49	88	30	53
Obs. insurance firm exposure $\geq 50\%$	11	11	11	11
Hurricanes	18	19	9	9
Time FE	Yes	Yes	Yes	Yes

Panel B: State considered hit if 25% or more of the counties were damaged

Dependent variable: Change in IV, $log\left(IV_{i,T_h+5}/IV_{i,T_h^*}\right)$	2002-2017		2007-2017	
	FEMA	FEMA+SHELDUS	FEMA	FEMA+SHELDUS
$StateDamageExposure_{i,T_h}$	0.426^{*}	0.400^{**}	0.424^{*}	0.413^{**}
	1.922	2.329	1.733	2.251
Adjusted R^2 (%)	38.482	36.595	37.328	38.254
Obs. total	367	385	205	205
Obs. insurance firm exposure $> 0\%$	326	345	177	179
Obs. insurance firm exposure $\geq 20\%$	22	43	14	27
Obs. insurance firm exposure $\geq 50\%$	7	11	7	11
Hurricanes	17	18	8	8
Time FE	Yes	Yes	Yes	Yes

Table A.1: Hurricane effects on implied volatility with industry interactions

This table reports the coefficients and test statistics when estimating the panel model in equation (9) but including an industry interaction term. The dependent variable is the change (in percent) in the implied volatility of firm i from the day before the inception of the hurricane until 5 trading days after the landfall. The first independent variable measures how much of the geographic footprint of a firm is exposed to the disaster area. The second independent variable interacts the exposure to the disaster area with an industry indicator variable for industry g. The industry classification is based on SIC. For Panel A, the geographic footprint used to measure the exposure to a hurricane of a firm is based on establishments per county, and for Panel B, the geographic footprint is based on sales per county. The analysis is based on damage data from FEMA enhanced with SHELDUS data. The data are from 2007 to 2017. The values in parentheses are the t-stats. The standard errors are clustered by county (headquarter location). Industry and time fixed effects are used. The significance of the coefficient estimate is indicated by * for $p < 0.10$, ** for $p < 0.05$, and *** for $p < 0.01$.

Panel A: Firms' hurricane exposure based on establishments

Dependent variable: Change in IV from hurricane inception to 5 days after landfall, $log\left(IV_{i,T_h+5}/IV_{i,T_h^*}\right)$

Interaction industry	Construction	Manufacturing	Mining	Retail	Services	Transportation	Wholesale
$DamageExposure_{i,T_h}$	0.095***	0.102***	0.107***	0.100***	0.091***	0.079***	0.090***
	(4.892)	(5.080)	(4.024)	(5.041)	(4.804)	(3.313)	(4.437)
$DamageExposure_{i,T_h} \times I_{i \in Industry_g}$	-0.012	-0.021	-0.059	-0.132	0.022	0.102	0.072
	(-0.080)	(-0.587)	(-1.513)	(-0.947)	(0.342)	(1.397)	(1.112)
Adjusted R^2 (%)	0.133	0.133	0.133	0.133	0.133	0.133	0.133
Obs. total	8,421	8,421	8,421	8,421	8,421	8,421	8,421
Obs. exposure $> 0\%$	3,692	3,692	3,692	3,692	3,692	3,692	3,692
Obs. exposure $\geq 20\%$	701	701	701	701	701	701	701
Obs. $i \in Industry_g$	161	3,657	846	783	1,463	1,166	280
Number of hurricanes	10	10	10	10	10	10	10
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table A.1: Hurricane effects on implied volatility with industry interactions (continued)

Panel B: Firms' hurricane exposure based on sales

Dependent variable: Change in IV from hurricane inception to 5 days after landfall, $log\left(IV_{i,T_h+5}/IV_{i,T_h^*}\right)$

Interaction industry	Construction	Manufacturing	Mining	Retail	Services	Transportation	Wholesale
$DamageExposure_{i,T_h}$	0.059***	0.061***	0.065***	0.062***	0.072***	0.042**	0.056***
	(3.393)	(2.719)	(3.074)	(3.832)	(4.548)	(2.058)	(3.136)
$DamageExposure_{i,T_h} \times I_{i \in Industry_g}$	0.127	-0.003	-0.031	-0.084	-0.073	0.100**	0.065
	(0.923)	(-0.086)	(-1.062)	(-0.683)	(-1.512)	(2.082)	(1.257)
Adjusted R^2 (%)	13.206	13.203	13.209	13.212	13.231	13.257	13.212
Obs. total	8,411	8,411	8,411	8,411	8,411	8,411	8,411
Obs. exposure $> 0\%$	3,680	3,680	3,680	3,680	3,680	3,680	3,680
Obs. exposure $\geq 20\%$	690	690	690	690	690	690	690
Obs. $i \in Industry_g$	161	3,647	846	783	1,463	1,166	280
Number of hurricanes	10	10	10	10	10	10	10
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Hedging Climate Change News*

Robert Engle[†] Stefano Giglio[‡] Heebum Lee[§] Bryan Kelly[¶] Johannes Stroebel[||]

Abstract

We propose and implement a procedure to dynamically hedge climate change risk. We extract innovations from climate news series that we construct through textual analysis of newspapers. We then use a mimicking portfolio approach to build climate change hedge portfolios. We discipline the exercise by using third-party ESG scores of firms to model their climate risk exposures. We show that this approach yields parsimonious and industry-balanced portfolios that perform well in hedging innovations in climate news both in sample and out of sample. We discuss multiple directions for future research on financial approaches to managing climate risk. (*JEL* G11, G18, Q54)

*This version: May 7, 2019. We thank Harrison Hong, Andrew Karolyi, and Ross Valkanov; participants at the Climate Finance Workshop at Columbia University, the Climate Finance Conference at Imperial College, and the Volatility Institute Conference at NYU Stern; and a number of anonymous referees for helpful comments. The Norwegian Finance Initiative and the Global Risk Institute provided generous grant support. We thank Konhee Chang for outstanding research assistance. Send correspondence to Robert Engle, NYU Stern, 44 West 4th Street, New York, NY 10012; telephone: 212-998-0710. E-mail: rengle@stern.nyu.edu.

[†]NYU Stern and NBER. Email: rengle@stern.nyu.edu

[‡]Yale University, NBER, and CEPR. Email: stefano.giglio@yale.edu

[§]NYU Stern. Email: hlee2@stern.nyu.edu

[¶]Yale University, AQR Capital Management, and NBER. Email: bryan.kelly@yale.edu

[||]NYU Stern, NBER, and CEPR. Email: johannes.stroebel@nyu.edu

Introduction

Earth's climate is changing, but uncertainty around the trajectory and the economic consequences of climate change is substantial. As a result, investors around the world desire products that allow them to hedge against the realizations of climate risk. Because of the long run and nondiversifiable nature of climate risk, standard futures or insurance contracts in which one party promises to pay the other in the event of a climate disaster are difficult to implement. Indeed, no counterparty could credibly guarantee to pay claims during a climate disaster event that might materialize in many decades, in part because a bad outcome would mandate all contracts to be paid at the same time. Individual investors are therefore largely constrained to self-insure against climate risk.

In this paper, we propose an approach for constructing climate risk hedge portfolios using publicly traded assets. We follow a dynamic hedging approach similar to Black and Scholes (1973) and Merton (1973). In this approach, rather than buying a security that directly pays off in the event of a future climate disaster, we construct portfolios whose short-term returns hedge *news* about climate change over the holding period. By hedging, period by period, the innovations in news about long-run climate change, an investor can ultimately hedge her long-run exposure to climate risk. In the short run, such a portfolio differs from the Markowitz mean-variance efficient portfolio and will thus exhibit a lower Sharpe ratio; but, in the long run, the dynamic hedging approach will compensate investors for losses that arise from the realization of climate risk.

The primary objective of this paper is to provide a rigorous methodology for constructing portfolios that use relatively easy-to-trade assets (equities) to hedge against risks that are otherwise difficult to insure. We show that our approach, which uses tools from standard asset pricing theory, does indeed allow us to construct portfolios that can successfully hedge climate news out of sample. Having said that, we do not view our resultant hedge portfolios as the definitive best hedges against climate change risk, but instead as a starting point for further exploration. Along these lines, we will discuss many valuable directions for future research on using financial markets to hedge climate risk.

The first challenge to implementing a dynamic hedging strategy for climate risk is to construct a time series that captures news about long-run climate risk, and which can therefore help us to construct an appropriate hedge target. We start from the observation that when there are events that plausibly contain such information about changes in climate risk, this will likely lead to newspaper coverage of these events; indeed, newspapers may even be the direct source that investors use to update their subjective probabilities of climate risks. Our approach in this paper therefore is to extract a climate news series from textual analysis of news sources. A wide range of events covered in newspapers can carry potentially relevant information. Indeed, the list of topics that are often covered by newspapers in relation to discussions about climate risk includes extreme weather events (e.g., floods, hurricanes, droughts, wildfires, extreme temperatures), physical changes to the planet (e.g., sea level changes, glacial melting, ocean temperatures), regulatory discussions, technical progress in alternative fuel delivery, and the price of fossil fuels.

We construct two complementary indices that measure the extent to which climate change is discussed in the news media. The first index is calculated as the correlation between the text content of *The Wall Street Journal* (WSJ) each month and a fixed climate change vocabulary, which we construct from a list of authoritative texts published by various governmental and research organizations. The WSJ is among the most salient media outlets for market participants, and thus our index captures the intensity of climate change discourse that is accessible to the investment community at very low cost.

Our WSJ Climate Change News Index associates increased climate change reporting with news about elevated climate risk, based on the idea that climate change primarily rises to the media's attention when there is a cause for concern. An alternative approach is to directly differentiate between positive and negative news in our index construction. To this end, we construct a second news-based climate index that is designed to focus specifically on bad news about climate change. This index applies sentiment analysis to climate-related articles to measure the intensity of *negative* climate news in a given month.

In this paper, we do not try to distinguish between different *types* of climate change news. In particular, we do not distinguish between news about physical damages from climate change and news about regulatory risks that are related to climate change. These two risk measures might move independent from each other. For example, the Paris accord, which led to a pledge to reduce carbon emissions, might have represented an increase in regulatory risk and a decrease in physical risk. Separately measuring news series about physical and regulatory climate risk represents an interesting avenue for future research. Also, our focus in this paper is on *global* climate change news. Our indices ignore news about local climate events, which are not covered in the WSJ or in a large cross-section of newspapers.

The second step in implementing our dynamic hedging strategy is to construct portfolios that allow us to hedge innovations in these two news series. In particular, we seek to systematically explore which stocks rise in value and which stocks fall in value when (negative) news about climate change materializes. Then, by constructing a portfolio that overweights stocks that perform well on the arrival of such negative news, an investor will have a portfolio that is well-positioned to profit the next time when such news about climate change materializes. Continued updating of this portfolio based on new information about the relationship between climate news and stock returns will ultimately lead to a portfolio which is long the winners from climate change and short the losers.

Our econometric approach to forming such hedge portfolios follows standard methods in the asset pricing literature. If climate risk represents a risk factor for asset markets (i.e., if it is a factor that drives the comovement of different assets), it is possible to construct a well-diversified portfolio the return of which isolates the exposure to that risk factor. Investors can then hedge their climate risk exposure by trading this portfolio without changing their exposures to the other risk factors in their portfolios. Various approaches to construct such hedge portfolios have been proposed in the literature. The two main ones are cross-sectional regressions like Fama-MacBeth (in which the hedging portfolio is obtained through period-by-period cross-sectional regressions of asset returns onto exposures to the risk factors), and direct projections of the risk factors onto a set of asset returns (the so-called "mimicking portfolio approach").[1] Among the many prominent papers in this literature are Fama and MacBeth (1973), Chen et al. (1986), Huberman et al. (1987), Breeden et al. (1989), Lamont (2001), Balduzzi and Robotti (2008), Lönn and Schot-

[1]The literature on cross-sectional regressions, like Fama-MacBeth, typically focuses on estimating the risk premiums of the factor, but risk premiums are simply the average excess returns of the corresponding hedge portfolios.

man (2017), and Roll and Srivastava (2018). Giglio and Xiu (2018) study the asymptotic properties of the different estimators in large cross-sections, and investigate their robustness to model specification errors. In this paper, we will apply the mimicking portfolio approach, as advocated by Lamont (2001).

The challenge with implementing this mimicking portfolio approach is that we only observe a limited number of months of climate news realizations, but have a large set of assets that we could use to form hedge portfolios. This leads to concerns about data mining, where we might end up constructing hedge portfolios that perform very well in sample but that are not stable going forward. To address this concern, we use characteristics that proxy for a firm's exposure to climate risk to parsimoniously parameterize the weights of the hedge portfolios. For example, one such characteristic might be the carbon footprint of each firm. In particular, it might be that when there is news about increasing climate risk, individuals will buy low-carbon-footprint stocks and sell high-carbon-footprint stocks. If this were the case, one could construct a portfolio that increases in value when there is (negative) news about climate risk using thousands of long and short positions based on just one parameter, the firms' carbon footprints.

We implement this characteristics-based approach by using firm-level environmental performance scores constructed by the ESG ("**E**nvironmental, **S**ocial, and **G**overnance") data providers MSCI and Sustainalytics to proxy for firms' climate risk exposure.[2] In particular, we use these scores as characteristics on which to sort individual stocks to form portfolios. We then construct the final hedge portfolios by projecting innovations in our climate news indices onto these ESG-characteristic-sorted portfolios, together with standard Fama-French factor-sorted portfolios (market, size, and value).

When we compare our hedge portfolios to alternative hedge portfolios that add simple industry bets (such as positions in the energy exchange-traded fund XLE) to the standard Fama-French factors, we find that our ESG-characteristic-based mimicking portfolios pro-

[2] Again, there is a question of what *type* of climate change risk exposure these measures capture. Specifically, they may more closely capture regulatory risks than physical risks, and other characteristics could be added to the analysis to capture different types of climate change exposures. For example, one could perhaps proxy for firms' physical climate risk by the distance of firms' headquarters or production facilities from the sea. Exploring different firm-level measures of climate risk exposure (both physical and regulatory) constitutes an interesting avenue for future research.

cedure produces hedge portfolios that perform better than the alternatives in hedging innovations in climate risk. In particular, our portfolios deliver higher in-sample and out-of-sample correlations with those innovations. For example, the return of the hedge portfolio based on the Sustainalytics E-Scores achieves out-of-sample correlations with the WSJ index innovations as high as 30%. Our hedge portfolios also do not resemble industry bets; rather, they identify, both within and across industries, those firms with the largest exposures to climate change risk, yielding a climate hedge portfolio that is relatively industry-balanced.

Our work contributes to a burgeoning literature that studies how climate change affects asset markets, and how asset markets in turn may affect the dynamics of climate change. Andersson et al. (2016) propose a passive investment strategy tilted to low-carbon stock as a hedge against climate risk, while Choi et al. (2018) explore how investors update their information about climate risk. Hong et al. (2018) investigate whether international stock markets efficiently price drought risk, and Kumar et al. (2018) explore whether fund managers misestimate the risk of climate disasters. Baldauf et al. (2018), Bakkensen and Barrage (2018), Bernstein et al. (2018), Giglio et al. (2018), and Murfin and Spiegel (2018) explore the pricing of climate risk in real estate markets, while Giglio et al. (2015, 2018) use real estate pricing data to back out very long-run discount rates that are appropriate for valuing projects aimed at mitigating climate change. Daniel et al. (2015) apply standard asset pricing theory to calibrate the social cost of carbon.

1 Construction of the Hedge Portfolios: Theory

This section discusses our methodology to construct portfolios that hedge news about climate change. We denote by r_t an $n \times 1$ vector of excess returns over the risk-free rate of n assets at time t. We assume that these returns follow a linear factor model, in which asset returns are driven by innovations in climate news, which we denote by CC_t, as well as by p other (tradable or nontradable) risk factors v_t:

$$\underbrace{r_t}_{n\times 1} = (\underbrace{\beta_{CC}}_{n\times 1}\underbrace{\gamma_{CC}}_{1\times 1} + \underbrace{\beta_{CC}}_{n\times 1}\underbrace{(CC_t - E[CC_t])}_{1\times 1}) + (\underbrace{\beta}_{n\times p}\underbrace{\gamma}_{p\times 1} + \underbrace{\beta}_{n\times p}\underbrace{v_t}_{p\times 1}) + \underbrace{u_t}_{n\times 1}. \tag{1}$$

The vectors β_{CC} and β are risk exposures of the n assets to the climate news factor and the other p factors, respectively. Similarly, γ_{CC} and γ are the corresponding risk premiums for the climate news factor and the other risk factors. Finally, u_t is an idiosyncratic error term. In this basic setup, the risk exposures are constant; we relax this assumption below.

Our objective is to construct a hedge portfolio for CC_t. This is defined as a portfolio that has unit exposure (beta) to climate risk shocks CC_t, but no exposure to any of the other p factors v_t. This ensures that investors can change their exposure to climate risk by trading in this portfolio, without modifying their exposure to the other risk factors. The asset pricing literature has followed two main approaches to construct hedge portfolios: the Fama-MacBeth cross-sectional regression approach and the mimicking portfolio approach. Giglio and Xiu (2018) derive theoretical properties of the two estimators in large-dimensional settings.

In this paper, we follow the mimicking portfolio approach; for completeness, Appendix A.1 provides a review of the Fama-MacBeth procedure in our setting. In the mimicking portfolio approach, the climate risk factor CC_t is directly projected onto a set of excess returns of a set of portfolios, $\tilde{r}_t$:

$$CC_t = \xi + w'\tilde{r}_t + e_t. \tag{2}$$

The hedge portfolio for CC_t is constructed using the weights $\hat{w}$ estimated from this regression; its excess return is $h_t^{CC} = \hat{w}'\tilde{r}_t$. The vector e_t captures the measurement error in CC_t, so that this approach explicitly accounts for potential measurement error in the climate risk factor CC_t. A sufficient condition for this procedure to recover the desired hedge portfolio for climate news is that the returns of the portfolios used in the projection, $\tilde{r}$, span the same space as the true factors, (CC_t, v_t).[3]

[3]Formally, write the model in the following compact form by calling f the vector of all factors: $f_t \equiv (CC_t, v_t)$, with covariance matrix Σ_f and β_f the matrix of betas: $\beta_f = (\hat{\beta}_{CC}, \hat{\beta})$. Call η the $(p+1)\times 1$ vector

1.1 Implementation and construction of the hedge portfolios

To build hedge portfolios using the mimicking portfolio approach, we choose a set of projection portfolios which are well diversified, so that idiosyncratic error is approximately eliminated, and which at the same time capture different dimensions of risk, so that their returns $\tilde{r}_t$ span the factor space. The portfolios used in the projection need to satisfy one further requirement. In particular, the setup described in Equation 1 includes the assumption that the risk exposures of the assets used in the estimation are constant over time. We therefore need to construct the portfolios $\tilde{r}$ in such a way that their exposures to the underlying risk factors are constant. A standard approach to achieve this is to form portfolios by sorting assets on characteristics. Indeed, to the extent that risk exposures of individual assets directly depend on these characteristics, sorting the assets by characteristics will ensure that the resultant portfolios have constant risk exposures. We follow this approach and choose a matrix of firm-level characteristics Z_t, appropriately cross-sectionally normalized, to construct the portfolio returns as

$$\tilde{r}_t = Z'_{t-1} r_t,$$

where r_t are excess returns of individual stocks, and portfolio weights are equal to the normalized characteristics.[4] Substituting this expression into Equation 2, we write

$$CC_t = \xi + w' Z'_{t-1} r_t + e_t. \tag{3}$$

Equation 3 can be interpreted in two ways. It can either be thought of as a projection of the hedge target CC_t onto characteristic-sorted portfolios $Z'_{t-1} r_t$ that are assumed to have constant risk exposure and that span the entire factor space. Alternatively, it can be thought of as a constrained projection of CC_t on *all* individual asset returns r_t, but with

with 1 as the first element and 0 everywhere else, so that $CC_t = \eta' f_t$. The population vector of weights w is $Var(\tilde{r}_t)^{-1} Cov(\tilde{r}_t, CC_t)$. If returns $\tilde{r}_t$ span the same space as the true factors, this means there exists an invertible matrix H such that $\tilde{r}_t = H f_t$. We can then write $w = (H \Sigma_f H')^{-1} H \Sigma_f \eta = H'^{-1} \eta$. The return of this portfolio is $h_t^{CC} = w' \tilde{r}_t = w' H f_t = \eta' H^{-1} H f_t = \eta' f_t = CC_t$.

[4] Note that we are exclusively working with excess returns, so there are no theoretical constraints on portfolio weights.

time-varying weights $w'Z'_{t-1}$; the weights are modeled as a linear function of characteristics, so that any individual firm's weight depends on its risk exposure to the different factors. Equation 3 therefore performs a one-step dimension reduction that estimates the hedge portfolio, while modeling the time variation in risk exposures.

2 Hedging Climate Change News

In this section, we implement the mimicking portfolio approach to hedging climate risk that we described above. As we have highlighted in the Introduction, the relevant performance measure for the resultant hedge portfolios is how well they hedge innovations to climate news out of sample. However, given the relatively short time period for which we observe measures of both climate news and firm-level climate risk exposures, there are a limited number of out-of-sample test periods on which to evaluate the climate hedge portfolios.[5] As will become apparent below, there are many degrees of freedom in how to construct these hedge portfolios, including decisions about how to construct measures of firm-level climate risk exposures and about what other portfolios to include in regression 2. As a result, there is the danger of optimizing over these degrees of freedom to construct portfolios that provide optimal out-of-sample hedges to climate news over the short period we observe, but that may not be effective at hedging this news going forward.

To avoid such data mining concerns, we will clearly describe the various choices we encountered in the construction of the climate hedge portfolios. However, instead of optimizing over these degrees of freedom to find a portfolio that optimally hedges climate news over our short test sample, we make choices that appear reasonable to us, and that will hopefully lead to stable approaches to hedging climate news that is yet to occur. This discussion will highlight a number of important directions in which to further develop these climate hedge portfolios, and longer time series of measures of climate news and

[5]In addition, even if we could easily extend our time series further into the past, it is unclear whether the additional sample periods would help us with constructing climate hedge portfolios today. In particular, it is plausible that climate risk has only started to be priced in stocks in recent years as investors' attention to this risk has increased. Indeed, some indirect evidence for such a suggestion comes from the fact that demand for ESG measures has substantially increased over the past few years. As a result, it is unclear whether firms with different climate risk exposures have had different excess returns in response to climate news that materialized in, say, the 1990s.

climate risk exposures will allow for more systematic ways of testing the true out-of-sample performance of different climate hedge portfolios.

2.1 Measuring climate change news

The first step in our analysis is to construct an index that measures innovations in news about climate risk. A variety of choices must be made when constructing this hedge target. How should we identify the news sources that reflect the information investors use in their climate risk-based investment decisions? Once we identify the appropriate news, how do we measure its relative intensity over time? How do we quantify the extent of good news versus bad news? And should one differentiate among subtypes of climate news (such as news about physical climate risks versus news about regulatory risks)?

Below, we follow two alternative approaches to building a climate news index. We believe they have the virtues of breadth and simplicity and offer scope for comparing trade-offs in some of our construction choices. At the same time, our indices have obvious imperfections and leave much room for other researchers to propose adjustments. Indeed, different investors might want to make different choices to ours in order to optimally align their hedge targets with the overall climate exposures of the rest of their portfolios. For example, investors with a strong coastal real estate portfolio might want to focus more on news about physical climate risk (because such real estate is strongly exposed to rising sea levels), while investors with a strong exposure to the coal industry might want to focus more on news about regulatory interventions in response to climate risk.[6]

2.1.1 *Wall Street Journal* climate change news index.

The first index that we construct is based on climate news coverage in *The Wall Street Journal* (WSJ). Two considerations support our use of the WSJ. One is a desire to measure news that is relevant to and salient for investors concerned about climate risks, and the WSJ is among the most important media sources consumed by financial market partici-

[6]In addition, some researchers and investors may want to expand the list of publications they consider beyond our newspaper-based approach. Additional publications of interest could include coverage in scientific journals or social media posts.

pants. The second advantage is that we have access to the full text of WSJ articles since the early 1980s, which provides us with complete flexibility in choosing how to build the climate news index from raw news content.

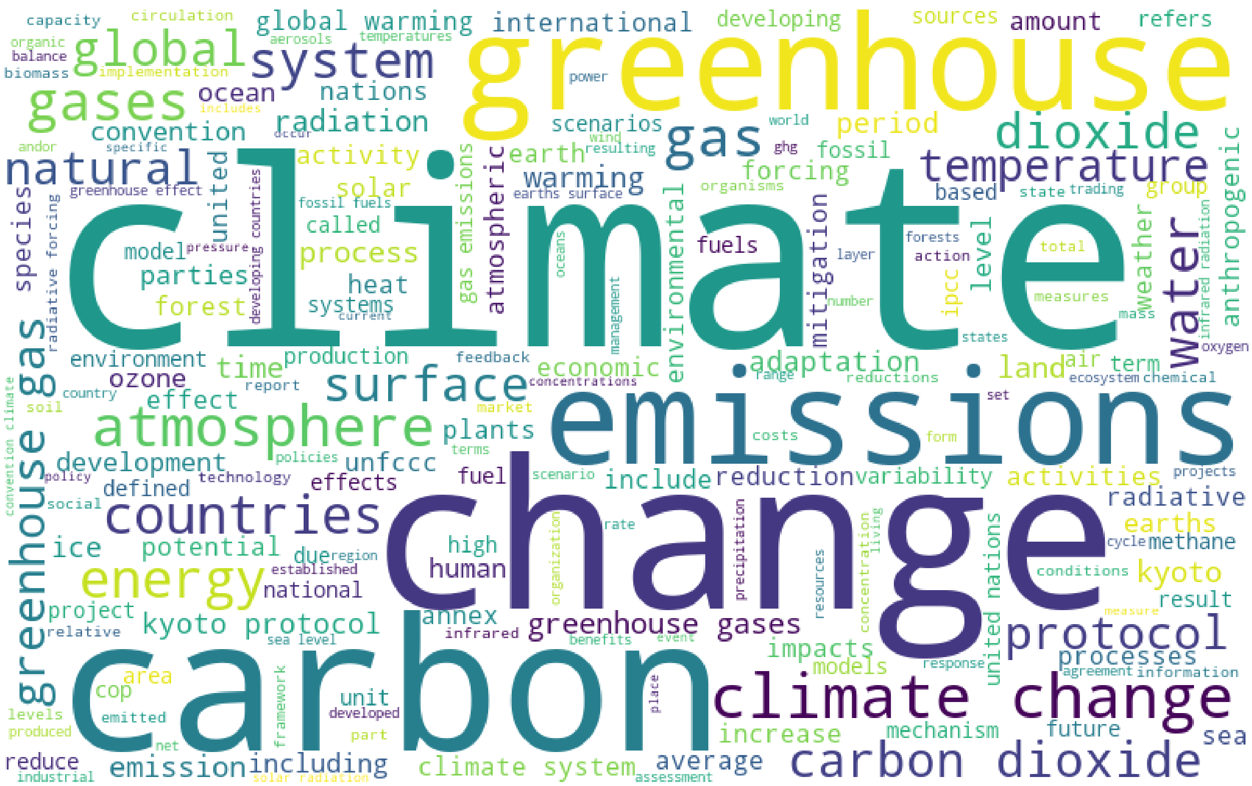
To quantify the intensity of climate news coverage in the WSJ, we compare the news content to a corpus of authoritative texts on the subject of climate change. In particular, we collect 19 climate change white papers from sources such as the Intergovernmental Panel on Climate Change (IPCC), the Environmental Protection Agency (EPA), and the U.S. Global Change Research Program. We complement these white papers with 55 climate change glossaries from sources such as the United Nations, NASA, the IPCC, the EPA, and others. Appendix A.2 presents the full list of these authoritative texts. We aggregate the seventy-four text documents into a "Climate Change Vocabulary (CCV)," which amounts to the list of unique terms (stemmed unigrams and bigrams) and the associated frequency with which each term appears in the aggregated corpus. Figure 1 provides an illustration of the CCV in the form of a word cloud, with term sizes proportional to their frequency.

We form an analogous list of term counts for the WSJ. Each (daily) edition of WSJ is treated as a "document," and term counts are tallied separately for each document. Next, we convert WSJ term counts into "term frequency–inverse document frequency," or *tf-idf*, scores. Common terms that appear in most documents earn low scores because they are less informative about any individual document's content (they have low *idf*), as do terms that are rare in a given article (they have low *tf*). The *tf-idf* transformation defines the most representative terms in a given document to be those that appear infrequently overall, but frequently in that specific document (see Gentzkow et al., 2018).

The main choice going into our index construction is to treat the CCV as our definition of phraseology associated with climate change discourse. That is, our CCV takes a stand on the specific terms, and their relative usage intensity, to identify news about the topic of climate change. Like with the WSJ, we convert Climate Change Vocabulary term counts into *tf-idf*. We treat the aggregated CCV as a single document when calculating term frequencies, and apply the inverse document frequency calculation from the WSJ corpus.[7]

[7]The choice to use the same *idf* for WSJ and CCV counts ensures that the document-frequency weights

Figure 1: Climate change vocabulary



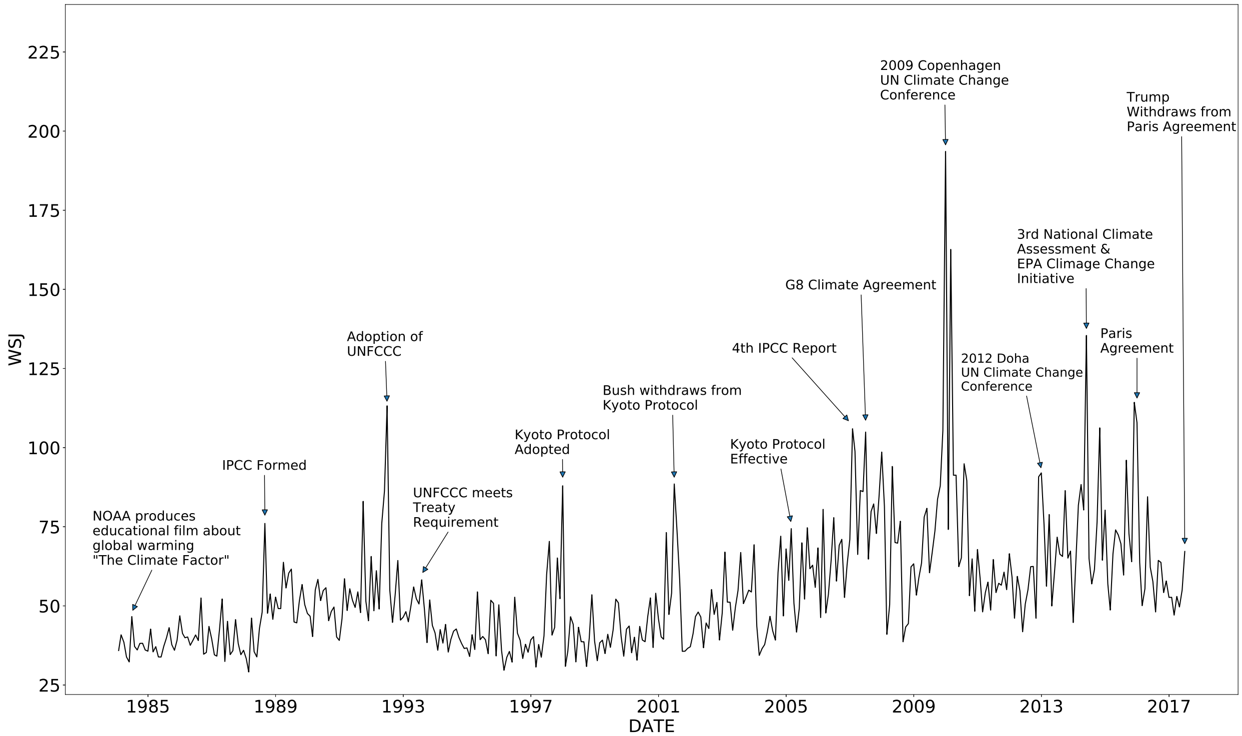
Word cloud summary of climate change vocabulary from a corpus of seventy-four authoritative climate change texts. Term sizes are proportional to their frequency in the corpus.

Finally, we construct our daily climate change index as the "cosine similarity" between the *tf-idf* scores for the CCV and each daily WSJ edition. Days in which the WSJ uses the same terms in the same proportion as the CCV earn an index value of one, while days in which the WSJ uses no words from the CCV earn an index value of zero. Approximately speaking, our raw WSJ Climate Change News Index describes the fraction of the WSJ dedicated to the topic of climate change each day, as defined by the texts that underlie the CCV. We scale this index by a factor of 10,000 to allow interpretation of the magnitudes of innovations in the index, which will represent our eventual hedge targets.

Figure 2 shows a time series of the WSJ Climate Change News Index since 1984. The figure shows that the intensity of climate news coverage has steadily increased since about the year 2000. In addition, the climate risk index spikes during salient climate

of CCV terms match the weights of WSJ terms. If we were to instead calculate *idf* based on the corpus of authoritative climate texts, we would down weight the most informative climate change terms and unduly distort the measurement of climate change discourse in the WSJ.

Figure 2: WSJ Climate Change News Index



This figure shows the WSJ Climate Change News Index from 1984 to 2017, annotated with climate-relevant news announcements.

events, such as the adoption of global climate treaties (e.g., the UNFCCC or the Kyoto protocol), or important global conferences to battle climate change (e.g., the 2009 UN Climate Change Conference in Copenhagen).

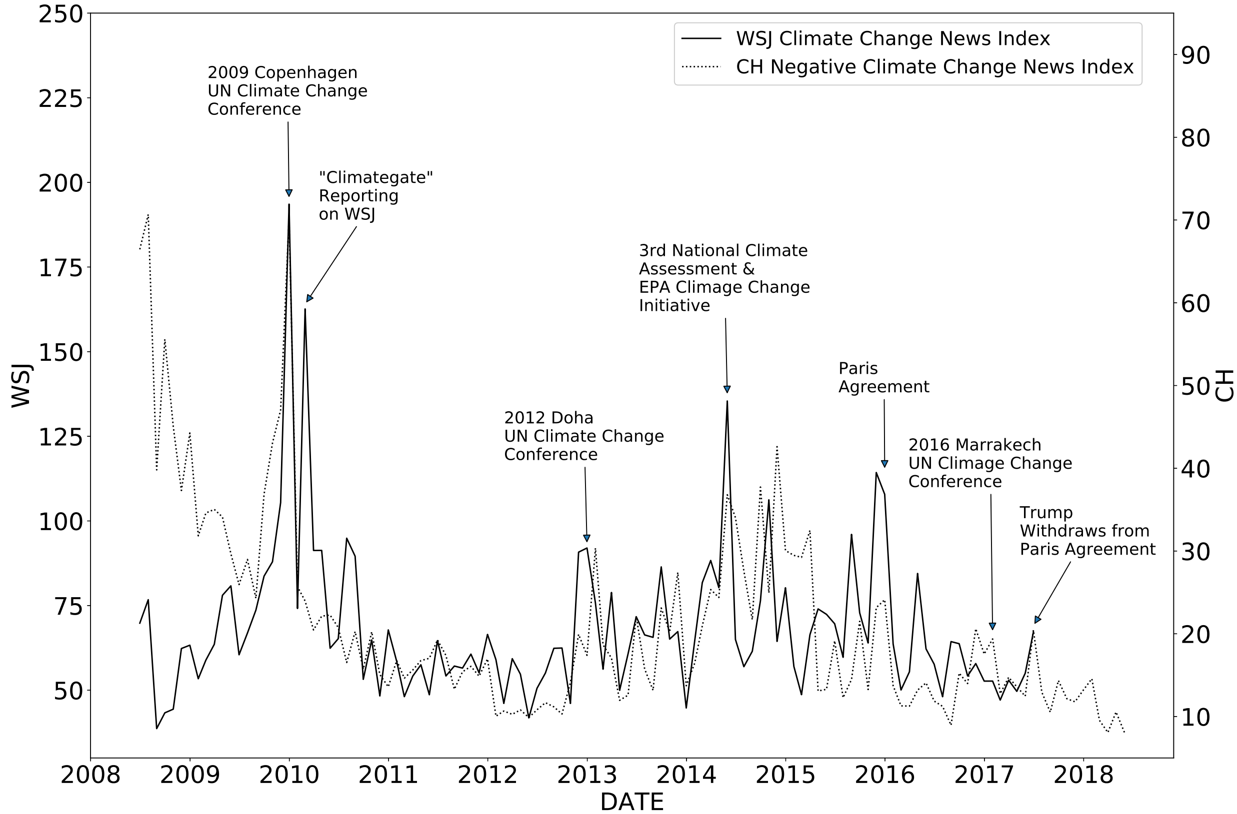
2.1.2 Crimson Hexagon's negative sentiment climate change news index.

Implicit in our construction of the WSJ Climate Change News Index is the assumption that the number of climate change discussions increase when climate risk is elevated. In other words, the WSJ index embeds the view that, when it comes to climate change, no news is good news. While we view this as a plausible assumption, there is a risk of inaccurately capturing discussions of positive climate news (e.g., news about new mitigation technologies) as increases in climate risk. A separate potential shortcoming of the WSJ index is that, being based on a single source, it may be too narrow in its quantification of climate discourse among investors.

To address these possible concerns, we study a second news-based climate risk index that is designed to focus specifically on negative climate news, and that is drawn from a much more expansive collection of news articles. For this purpose, we use the services of the data analytics vendor Crimson Hexagon (CH). Starting in May 2008, Crimson Hexagon has collected a massive corpus of over one trillion news articles and social media posts. The underlying news sources cover over 1,000 outlets, including the WSJ, *The New York Times*, *The Washington Post*, Reuters, BBC, CNN, and Yahoo News. Coverage in terms of total articles available expands over time. Cross-sectionally, the distribution of article counts is fairly evenly distributed across news outlets, with the top-100 outlets accounting for approximately 14% of the total article count. For a given user-provided search term, CH applies a variety of proprietary natural language processing analytics, such as sentiment analysis and topic modeling, to construct time series of the sentiment of coverage of that term across the sources it collects.

We provide CH with the search phrase "climate change" and restrict our analysis to discussions in the news media (i.e., we exclude social media). Based on these choices for terms and content sources, CH provided us with an array of indices that summarize the total number of articles that include climate change news, as well as the fraction of those summarized to contain positive and negative climate change news. It also provided indices for further sentiment subcategories (e.g., fear, joy, anger), as well as a topic decomposition of climate-related articles. Thus, there are many potential degrees of freedom in using Crimson Hexagon data to construct a climate news series. For example, we could tune our choice of search terms, or optimize across each of the finer indices that CH supplies for any given set of search terms. As described above, given the brevity of our data sample, we need to guard against data mining, and we do so in this case by restricting ourselves to the most obvious search term ("climate change") and focusing on the most obvious category that resolves our desire for "signed" news, namely those that CH categorizes as basic "negative sentiment." We calculate our CH Negative Climate Change News Index as the share of all news articles that are both about "climate change" and that have been assigned to the "negative sentiment" category; we multiply this measure by 10,000 in order to interpret the magnitudes of innovations in the index.

Figure 3: CH Negative Climate Change News Index



This figure shows the CH Negative Climate Change News Index from 2008 to 2017, overlaid against the WSJ Climate Change News Index, and annotated with climate-relevant news announcements.

Figure 3 plots the time series of the CH Negative Climate Change News Index, in addition to that of the WSJ Climate Change News Index for comparison. Both indices regularly spike around salient climate events, such as climate conferences. The initial level of the CH index is somewhat higher than that of the WSJ index, though this is during a period for which Crimson Hexagon has relatively little data; this is also a period that will not be included in our final analysis (as we discuss below, our empirical analysis starts in September 2009, the first month for which we observe complete coverage of firm-level climate risk exposures). Interestingly, the WSJ index spikes in a number of instances in which the CH index does not. One of these was in early 2010, a period during which the WSJ extensively reported on the "Climategate" controversy.[8]

[8]The Climategate controversy involved the publication of emails obtained through hacking a server at the Climatic Research Unit at the University of East Anglia. Several climate change "skeptics" alleged that these emails documented global warming to be a scientific conspiracy, with scientists manipulating data.

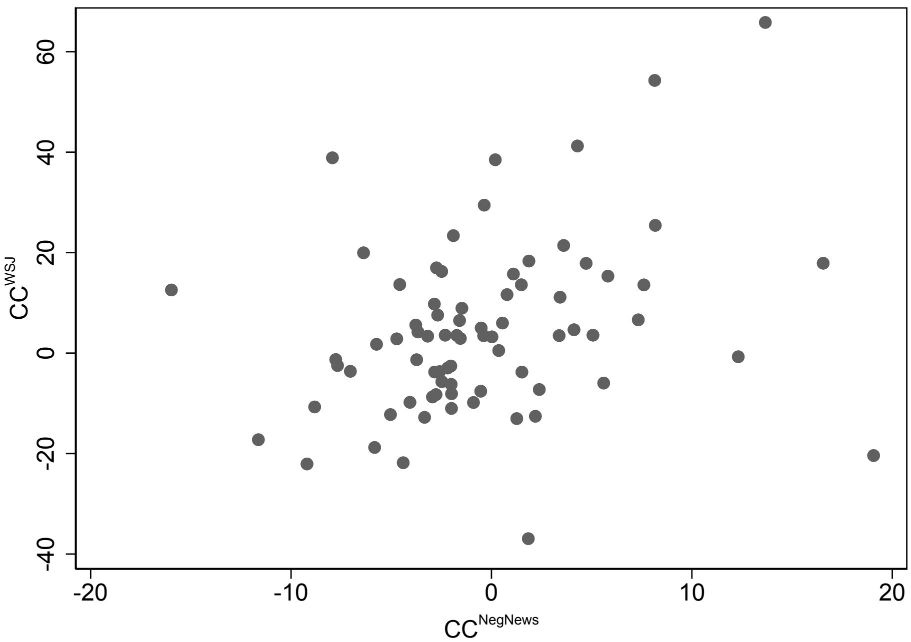
2.1.3 Constructing hedge targets.

To measure innovations in climate news, we average the daily values for the WSJ Climate Change News Index and CH Negative Climate Change News Index to the monthly level, and then construct values of CC_t as residuals from an AR(1) model. This gives us our two monthly hedge targets: CC_t^{WSJ}, which captures innovations in the WSJ Climate Change News Index, and $CC_t^{NegNews}$, which captures innovations in the CH Negative Climate Change News Index. Figure 4 shows the correlation across these measures across the 88 months that will be included in our final analysis, September 2009 to December 2016. The correlation coefficient is 0.3, which suggests that, although both measures capture common elements of climate risk, they are by no means identical. As we have discussed above, which of the two series (or any one of the potential alternative series that we could have constructed) represents the ideal hedge target depends on the precise application; as a result, we view the construction of alternative hedge targets as an exciting area for further research.

2.2 Potential assets in hedge portfolios

After defining the hedge targets, the second step in implementing the mimicking portfolio hedge approach described in Section 1 is to determine the universe of assets used to build the hedge portfolio. In this project, we focus on constructing hedge portfolios using U.S. equities as the underlying assets. We obtain monthly individual U.S. stock return data from CRSP. We include only common equity securities (share codes 10 and 11) for firms traded on the NYSE, AMEX and NASDAQ. Following Amihud (2002) and many others, we exclude penny stocks, defined as stocks with a price below \$5 at the time of portfolio formation. This is to avoid including stocks whose returns are dominated by market microstructure issues. We also drop microcap stocks, defined as stocks with a market capitalization in the bottom 20% of the sample traded on the NYSE, following the observation in Fama and French (2008) that the returns of hedge portfolios obtained from long-short positions can be distorted by the inclusion of such microcaps (see also the discussion in Hou et al., 2015).

Figure 4: Correlation across CC_t measures



This figure shows a scatterplot highlighting the correlation across our two climate hedge targets, CC^{WSJ} and CC^{NegNes}. Each observation corresponds to 1 month between September 2009 to December 2016. The correlation coefficient is 0.30.

2.3 Measuring climate risk exposures

Having identified the set of possible assets to include in the hedge portfolio, the next empirical challenge is to systematically measure different firms' exposures to climate risk, that is, to identify the characteristics in Z_t that drive such exposures. Our approach in this paper is to build on measures of firms' environmental exposures produced by third-party ESG data providers. Indeed, there has been a growing interest in ESG investing among investors who are increasingly demanding assets that fulfill certain environmental ("E"), social ("S"), and governance ("G") criteria.[9] Given this trend, measuring the ESG characteristics of firms has become an important task for investors, and firm-level ESG scores are available from numerous providers that collect raw data gathered from sources such as firms' disclosures, SEC filings, and reports by governments or NGOs. These raw

[9]According to The U.S. SIF Foundation, the dollar value of ESG assets owned by institutional investors grew to $4.73 trillion in 2016, an increase of 11% a year since 2005.

data are then translated into numerical ESG scores using proprietary algorithms.[10]

Our study uses information on firm-level ESG scores from two leading data providers, MSCI and Sustainalytics.[11] Both data providers construct various subscores that evaluate firms on different aspects of their ESG performance. From these subscores, we choose the broadest scores that plausibly proxy for firms' exposure to climate risk.

2.3.1 MSCI

We obtained from MSCI a data set of annual firm-level ESG scores between 1995 and 2016.[12] MSCI evaluates firms along several subcategories that capture either positive or negative environmental performance; Appendix A.3 presents the full list of subcategories. Each subcategory is either scored as a "1" when the firm satisfies a certain condition, or a "0" if the firm does not satisfy the condition. For instance, a "1" in the positive *"Climate Change - Energy Efficiency"* subcategory means that the company operates in a relatively energy-efficient way. The thresholds for satisfying each condition are determined by MSCI and are not disclosed with the data. Following Hong and Kostovetsky (2012), we calculate an overall environmental score for each firm by subtracting the total scores in the negative environmental subcategories from the total scores in positive environmental subcategories. We call the resultant variable the "MSCI E-Score," where a higher score suggests a firm is more environmentally friendly. In principle, it would be possible to also construct E-Scores from only a selection of all "E" subcategories, perhaps by focusing on those subcategories that are particularly relevant for climate change. The out-of-sample performance of hedge portfolios constructed using different combinations of "E" subcategories could then be compared to select the one with the best performance. However, given the relatively short time series to evaluate the performance of the resultant hedge

[10] As noted in the Introduction, ESG scores may capture specific notions of climate change exposure; for example, they may better capture exposure to regulatory risks than exposure to physical damages from climate risks. The methodology in this paper could be easily applied using other firm characteristics that may capture different types of climate risk exposures.

[11] The number of ESG data providers, including firms such as Arabesque and TruValue Labs, is growing. Analyzing which of these E-Scores results in the optimal hedge portfolio would be an interesting avenue for further research, but in the absence of longer time series is likely subject to concerns of data mining.

[12] These scores were formerly known as KLD scores. In 2010, following MSCI's acquisition of RiskMetrics, KLD scores were retooled into what are now known as MSCI KLD scores.

portfolios, even such an "out-of-sample" approach of finding the "best" E-Scores is naturally subject to data mining concerns. We hence decided to restrict ourselves to only analyzing the relatively broad overall E-Score, following prior approaches in the literature; we leave a more detailed exploration of the various subcategories to future research.

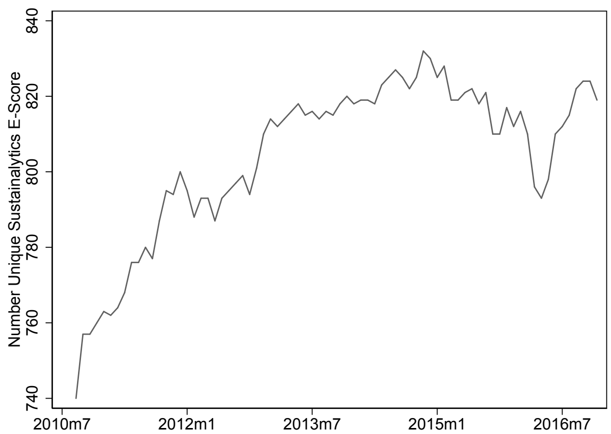
2.3.2 Sustainalytics

Sustainalytics provided us with monthly firm-level ESG scores beginning in September 2009. The broadest score in the data is the "Total ESG Score," which is the average of the "Total Environment Score," the "Total Social Score," and the "Total Governance Score." To determine each of the "E," "S," and "G" scores, Sustainalytics uses a number of subcategories and evaluates each firm's score by comparing it to peers in the same industry (Sustainalytics uses a nonstandard industry classification). For instance, the fifty-seven subcategories for the "Total Environment Score" include evaluations of a firm's efforts to reduce greenhouse gas emissions, increase renewable energy use, and reduce water use; Appendix A.3 presents the full list of subcategories. The scores in the subcategories are then aggregated by weighting them according to how exposed each industry is to each ESG risk, though this aggregation procedure is not well documented. Final scores are between 0 and 100. As before, a higher score suggests a firm is more environmentally friendly. We use the "Total Environment Score" in our empirical analysis.

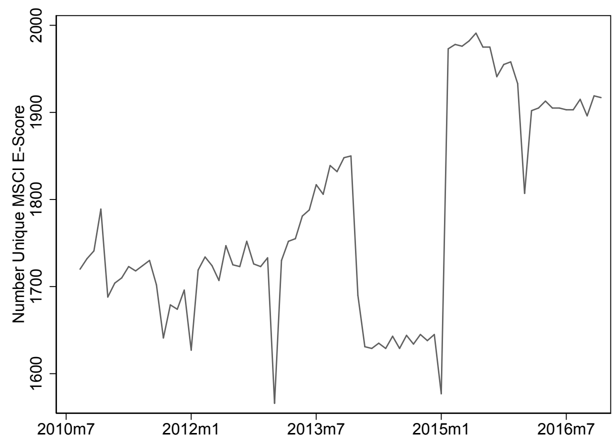
2.3.3 Summary Statistics

Our analysis of climate hedge portfolios focuses on the period between September 2009 and December 2016. This is a period for which we observe both measures of innovations of climate news, CC_t^{WSJ} and $CC_t^{NegNews}$, and both the Sustainalytics and MSCI E-Scores. We can therefore conduct a direct comparison of the performance of the various hedge portfolios for the two climate news series over this time horizon. For the MSCI E-Score, which is only reported annually, we assign the same score to all the months in the relevant year. Panels A and B of Figure 5 plot the number of firms in our pool of potential hedge assets for which we observe each E-Score over time. For Sustainalytics, we usually observe E-Scores for between 700 and 800 firms. MSCI E-Scores have broader coverage and are provided for between 1,700 and 1,900 firms.

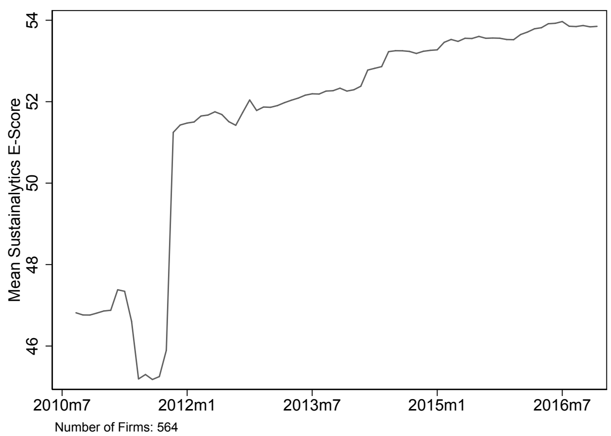
Figure 5: E-Scores: Summary statistics over time



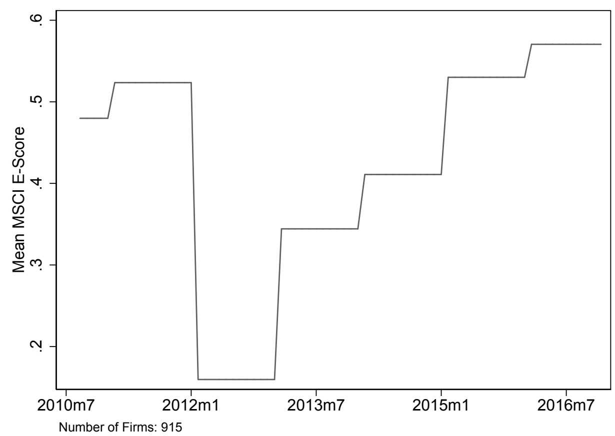
(a) Sustainalytics: Number of firms over time



(b) MSCI: Number of firms over time



(c) Sustainalytics: Mean over time



(d) MSCI: Mean over time

This figure provides summary statistics for our two E-Scores. The top row shows the number of firms in our sample for which we observe E-Scores. The bottom row shows the average E-Score over time across those firms that we observe in every period in our sample. The left column shows these statistics for the Sustainalytics E-Score, and the right panel shows the statistics for the MSCI E-Score.

Panels C and D of Figure 5 show the average values for each of the two E-Scores for a constant set of firms that we observe throughout the sample. The averages of each score contain a number of discontinuous breaks. For the MSCI E-Score, which is determined annually, these breaks could be either due to changes in firms' true ESG performance between years or due to changes in the modeling procedure. For Sustainalytics, which computes monthly scores, the discontinuous breaks are more likely due to changes in the modeling methodology over time, though we have been unable to obtain documentation on such changes that would allow us to verify this conjecture.[13] Such modeling changes would be problematic for building time-series models that perform well out of sample.

To minimize the complications from any modeling changes, we construct Z_t by cross-sectionally demeaning each E-Score in each month. However, this approach might still be problematic if changes to the model do not just shift the mean of the E-Scores over time, but also the cross-sectional dispersion. In that case, the meaning of absolute differences in the demeaned E-Score would change over time. As a second way to construct measures of Z_t, we therefore rank the E-Scores of all firms at each point in time, and then demean and rescale the ranked measure such that it ranges from -0.5 to +0.5. This approach preserves the ordinal content of the E-Scores but discards any information contained by the absolute differences between scores. Ranking-based approaches come with a number of issues. In particular, panels A and B of Figure 5 highlight that the number of firms for which E-Scores are available changes throughout the sample period. Firms added later in the sample are plausibly systematically different from those added earlier; for example, they might be less exposed to climate risk. The cross-sectional ranking of the same firm might therefore change over time without the true climate exposure of that firm changing. As a result, neither the demeaned absolute value nor the demeaned and rescaled ranked value of E-Scores are ex ante superior methods to construct climate exposures in Z_t. We will therefore present hedge portfolios using both approaches to constructing exposure measures and compare their relative performance.[14]

An interesting question is what firm characteristics are captured by the two E-Scores. A first hypothesis is that they primarily pick up industry-membership, whereby firms in

[13]Most uses of ESG scores by the financial services sector build on the cross-section of ESG scores at a given point in time, for example, by forming portfolios that have a relatively higher performance on these measures. Such use cases often do not require a stable meaning of the same numerical score over time.

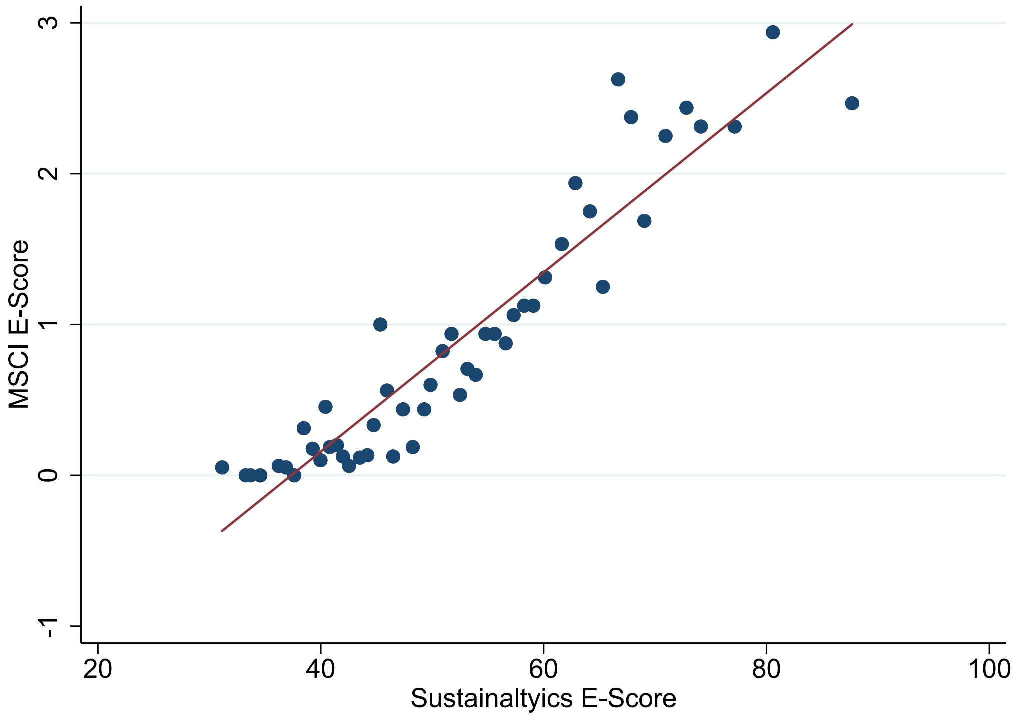
[14]The climate exposure measures in Z_t can be constructed from the various raw E-Scores in other ways. For example, one could cross-sectionally standardize each absolute measure to have a constant standard deviation over time. Alternatively, one could rank firms' E-Scores within industry rather than across all firms. However, in the absence of longer time series, a systematic analysis of which of these approaches obtains the best out-of-sample fit during our sample period is subject to the data mining concerns described earlier. As a result, we did not pursue these alternative approaches in this project.

"clean" industries, such as wind and solar energy, are assigned high E-Scores, and firms in "dirty" industries such as coal mining are assigned low E-Scores. To explore the extent to which the scores are primarily capturing a firm's industry, we begin by taking the firm-level E-scores in December 2016 (the last period in our data) and regressing them onto industry fixed effects. When regressing the absolute value of the Sustainalytics E-Score on 2-digit SIC code fixed effects, the adjusted R-squared of the regression is .103; it is .184 when regressing on fixed effects for 4-digit SIC codes. The measures of R-squared were similar when using the ranked measure of the Sustainalytics E-Score. When regressing the absolute value of the MSCI E-Score on 2-digit SIC codes (4-digit SIC codes), the adjusted R-squared of the regression is .099 (.203). These numbers show that, although there is some industry effect in determining E-Scores, most of the variation occurs within relatively narrow industries, rather than across industries.

Indeed, the three 2-digit SIC industries with the lowest Sustainalytics E-Scores are Personal Services (SIC code 72), Water Transportation (SIC code 44), and Motion Pictures (SIC code 78), probably not the first industries that come to mind when thinking of "dirty" industries. Similarly, the 2-digit SIC industries with the highest Sustainalytics E-Scores are Building Materials & Gardening Supplies (SIC code 52), Textile Mill Products (SIC code 22), and Furniture & Homefurnishings Stores (SIC code 57). When ranking by MSCI E-Scores, we similarly find that low-scoring firms are not necessarily those one would expect ex ante, such as those operating in the oil and gas sector.

A second question is the extent to which the MSCI and Sustainalytics E-Scores capture the same object. Figure 6 shows the correlation across the raw Sustainalytics and MSCI E-Scores in December 2016. They have a positive correlation of about 0.65, suggesting that they are both measuring aspects of the same object. However, enough independent variation occurs across the two measures to suggest that their usefulness in constructing climate hedge portfolios might vary. Indeed, we show below that the performance of the hedge portfolios varies noticeably when these hedge portfolios are constructed using the different E-Scores.

Figure 6: Correlation across E-Scores, December 2016



This figure shows a binned scatterplot that highlights the correlation across the Sustainalytics and MSCI E-Scores for all 796 firms in our sample that have both scores in December 2016. The correlation coefficient is 0.65.

2.4 Forming hedge portfolios

In this section, we construct hedge portfolios for innovations in climate news, CC_t, using the mimicking portfolio approach described in Section 1.1. As discussed above, we use two different approaches to transform the raw E-Scores into the characteristic vector Z_t:

(1) Using firms' cross-sectionally demeaned absolute value of the E-Score ("absolute scores", e.g., $Z_t^{SUS\_A}$)

(2) Ranking the firms cross-sectionally by their E-Score, and then standardizing these rankings to range between -0.5 and +0.5 ("ranked scores", e.g., $Z_t^{SUS\_R}$).

Recall that one of the conditions for the mimicking portfolio approach to isolate climate change risk (and to avoid picking up other potentially correlated risks in the economy) is that the projection portfolios have to span all the risk factors driving returns. In addition to portfolios sorted on the climate characteristics, we therefore also include in regression 2 three additional factors that might be correlated with climate risk and that are known to be important in explaining the cross-section of returns: size (using cross-sectionally standardized market value to create Z_t, so that half the firms, sorted by market value, have positive weight, and half have negative weight; note that this portfolio will be long large firms and short small firms), value (using cross-sectionally standardized values of book-to-market to create Z_t), and the market (setting Z_t to equal the share of total market value).[15] For example, when we use the absolute Sustainalytics E-Score to measure firms' climate risk exposures, regression 3 becomes

$$CC_t = \xi + w_{SUS} Z_{t-1}^{SUS\_A'} r_t + w_{SIZE} Z_{t-1}^{SIZE'} r_t + w_{HML} Z_{t-1}^{HML'} r_t + w_{MKT} Z_{t-1}^{MKT'} r_t + e_t, \quad (4)$$

where $w_{SUS}, w_{SIZE}, w_{HML}$ and w_{MKT} are scalars that capture the weight of the corresponding portfolios in the mimicking (hedge) portfolio for CC_t.

For comparability, we also analyze the performance of hedge portfolios constructed using returns of the exchange-traded funds (ETFs) XLE and PBD instead of the returns of portfolios of stocks sorted by their E-Scores. XLE is the ticker of the Energy Select Sector SPDR ETF, which represents the energy sector of the S&P 500. PBD is the ticker of the Invesco Global Clean Energy ETF, which is based on the WilderHill New Energy Global Innovation Index and comprises companies that focus on greener and renewable sources of energy and technologies facilitating cleaner energy. Constructing hedge portfolios based on those ETFs allows us to (a) analyze the extent to which our E-Score-based hedge portfolios simply represent a market tilt away from "brown energy" and toward "green energy" and (b) explore whether hedge portfolios based on XLE and PBD would have performed better than our E-Score-based hedge portfolios.[16]

[15] To maximize the number of stocks used to construct the hedge portfolios, we include stocks even if some of the characteristics Z_t are missing for that stock. To do so, we set all missing characteristics equal to zero.

[16] As before, there are many degrees of freedom for how to compute hedge portfolios based on ETFs, and we do not want to suggest that portfolios constructed using XLE and PBD constitute the "best" ETF-based portfolios for hedging climate risk. Indeed, we view the analysis of which ETFs and other funds are most helpful in hedging climate risk to be an exciting area for future research.

2.5 In-sample fit results

We begin by exploring the in-sample fit of various versions of regression 4 over the full sample period. Table 1 shows regressions when hedging innovations to the WSJ Climate Change News Index, CC_t^{WSJ}, described in Section 2.1. Columns 1 and 2 show that portfolios based on Sustainalytics E-Scores have a positive and significant relationship with CC_t^{WSJ}; in periods with more innovations in negative climate news, a portfolio that goes long firms with higher (more "green") E-Scores has relatively larger excess returns. The R-squared measures of these regressions show that the portfolios based on the Sustainalytics E-Scores can hedge 15%–18% of the in-sample variation in CC_t. Columns 3 and 4 show that portfolios based on the MSCI E-Scores also have higher excess returns during periods with innovations in negative climate news; the R-squared measures of the regressions are lower than those in Columns 1 and 2. Portfolios based on ranked versions of both E-Scores have a slightly higher in-sample fit than portfolios based on absolute demeaned values. In addition to the ESG scores, size appears to correlate with climate change exposure: larger firms appear more exposed than smaller firms to climate change news, in the sense that they perform worse when the amount of news coverage of climate change in the WSJ increases. Column 5 includes the returns of XLE and PBD instead of the return of a characteristic-sorted portfolio. The in-sample fit of this regression is lower than that of any of the regressions in Columns 1–4, even though we have fewer explanatory variables in those regressions. This suggests that the characteristic-weighted portfolios might have some advantages over a hedge approach that creates industry tilts using energy-related ETFs.[17] It also shows that most of the R-squared in Columns 1–4 is the result of the characteristics-weighted portfolios, and not of the other portfolios, which are also included in Column 5.

Table 2 presents the same set of regressions as Table 1, but hedges innovations in the CH Negative Climate Change News Index, $CC_t^{NegNews}$. As before, the in-sample fits of the hedge portfolios based on Sustainalytics E-Scores are higher than the fits of the

[17]The inclusion of the other factors in regression 4 make the resultant hedge portfolios in Column 5 of Table 2 different from a simple industry-tilt away from the market. Indeed, the resultant hedge portfolio will have a beta of 1 with CC_t, and a beta of zero with the other factors. Factor neutrality, not industry neutrality, is a desirable property of hedge portfolios.

Table 1: Full-sample regression: WSJ Climate Change News Index

	(1)	(2)	(3)	(4)	(5)
$Z_{t-1}^{SUS\_A'} r_t$	1.416***				
	(0.436)				
$Z_{t-1}^{SUS\_R'} r_t$		67.789***			
		(17.834)			
$Z_{t-1}^{MSCI\_A'} r_t$			12.658*		
			(6.849)		
$Z_{t-1}^{MSCI\_R'} r_t$				53.743*	
				(27.401)	
r_t^{XLE}					0.085
					(0.810)
r_t^{PBD}					0.208
					(0.630)
$Z_{t-1}^{HML'} r_t$	1.221	2.309	-5.862	-5.941	-6.772
	(7.019)	(6.873)	(6.878)	(6.858)	(8.093)
$Z_{t-1}^{SIZE'} r_t$	-5.680**	-6.034**	-5.511*	-5.459**	-2.765
	(2.350)	(2.289)	(2.773)	(2.696)	(2.474)
$Z_{t-1}^{MKT'} r_t$	0.783	0.789	0.841	0.789	0.091
	(0.642)	(0.628)	(0.692)	(0.680)	(1.285)
Constant	2.894	2.673	4.659*	4.891*	5.959**
	(2.681)	(2.613)	(2.700)	(2.669)	(2.897)
R-squared	.153	.187	.083	.088	.047
N	88	88	88	88	88

This table shows results from regression 4. The dependent variable captures innovations for the WSJ-Based Climate News measure. The unit of observation is a month, and the sample runs between September 2009 and December 2016. Standard errors are presented in parentheses. *p <.1; **p <.05; ***p <.01.

hedge portfolios based on MSCI E-Scores; similarly, the in-sample fits of the portfolios constructed using ranked E-Scores are marginally higher than those of the portfolios constructed using the absolute (demeaned) E-Score. Finally, the in-sample fits of all four portfolios based on E-Scores are somewhat higher than that of the portfolio based on XLE and PBD.[18] Overall, the relative performance of the various hedge portfolios is similar whether we are trying to hedge the WSJ Climate Change News Index or the CH Negative Climate Change News Index.

How would the hedge portfolios implied by these regressions look? To determine

[18]It is interesting to note that when hedging negative climate change news, the value-growth dimension seems to be aligned with the risk exposure. In particular, the table shows that value firms appear more exposed to climate news than growth firms.

Table 2: Full-sample regression: CH Negative Climate Change News Index

	(1)	(2)	(3)	(4)	(5)
$Z_{t-1}^{SUS\_A'} r_t$	0.266*				
	(0.141)				
$Z_{t-1}^{SUS\_R'} r_t$		12.286**			
		(5.864)			
$Z_{t-1}^{MSCI\_A'} r_t$			1.089		
			(2.173)		
$Z_{t-1}^{MSCI\_R'} r_t$				6.641	
				(8.696)	
r_t^{XLE}					-0.092
					(0.252)
r_t^{PBD}					0.036
					(0.196)
$Z_{t-1}^{HML'} r_t$	-4.536**	-4.390*	-5.934***	-5.919***	-5.520**
	(2.272)	(2.260)	(2.182)	(2.177)	(2.519)
$Z_{t-1}^{SIZE'} r_t$	-0.137	-0.179	0.210	0.100	0.501
	(0.761)	(0.753)	(0.880)	(0.856)	(0.770)
$Z_{t-1}^{MKT'} r_t$	0.315	0.314	0.287	0.295	0.297
	(0.208)	(0.206)	(0.219)	(0.216)	(0.400)
Constant	-0.115	-0.137	0.313	0.306	0.376
	(0.868)	(0.859)	(0.857)	(0.847)	(0.902)
R-squared	.125	.133	.090	.094	.089
N	88	88	88	88	88

This table shows results from regression 4. The dependent variable captures innovations for the *Newspaper-based negative climate news* measure. The unit of observation is a month, and the sample runs between September 2009 and December 2016. Standard errors are presented in parentheses. *p <.1; **p <.05; ***p <.01.

each firm i's weight in the hedge portfolio, we construct the following sum, where $Z_{i,t}$ values are taken as of December 2016: $\hat{w}_{SUS\_A} Z_{i,Dec16}^{SUS\_A'} + \hat{w}_{SIZE} Z_{i,Dec16}^{SIZE'} + \hat{w}_{HML} Z_{i,Dec16}^{HML'} + \hat{w}_{MKT} Z_{i,Dec16}^{MKT'}$, and where the various $\hat{w}$-terms represent the estimated coefficients from regression 4. This means that a firm's weight in the hedge portfolio is determined by its E-Score as well as its book-to-market ratio and its size. The resultant portfolio is the portfolio that an investor would form in December 2016 to hedge climate news in January 2017. Table 3 presents the average portfolio positions by 2-digit SIC code classification for the industries with the six largest negative average portfolio weights and the industries with the six largest positive average portfolio weights. We only present the portfolio positions based on the absolute E-Scores, because they look very similar to the positions in

Table 3: Largest average short and long positions (by 2-digit SIC code)

A. WSJ Climate Change News Index

Sustainalytics E-Score (absolute)		**MSCI E-Score (absolute)**	
Top negative portfolio weights	***SIC2***	***Top negative portfolio weights***	***SIC2***
Coal mining	12	Water transportation	44
Water transportation	44	Petroleum & coal products	29
Insurance agents, brokers, & service	64	Motion pictures	78
Mining non-metalic minerals, except fuels	14	Communications	48
Transportation services	47	Security & commodity brokers	62
Security & commodity brokers	62	Oil & gas extraction	13
Top positive portfolio weights	***SIC2***	***Top positive portfolio weights***	***SIC2***
Building materials & gardening supplies	52	Pipelines, except natural gas	46
Tabacco products	21	Tabacco products	21
Food & kindred products	20	Miscellaneous manufacturing industries	39
Paper & allied products	26	Lumber & wood products	24
Textile mill products	22	Paper & allied products	26
Furniture & homefurnishings stores	57	Textile mill products	22

B. CH Negative Climate Change News Index

Sustainalytics E-Score (absolute)		**MSCI E-Score (absolute)**	
Top negative portfolio weights	***SIC2***	***Top negative portfolio weights***	***SIC2***
General building contractors	15	General building contractors	15
Water transportation	44	Nondepository institutions	61
Coal mining	12	Auto repair, services, & parking	75
Insurance agents, brokers, & service	64	Communications	48
Holding and other investment offices	67	Water transportation	44
Insurance carriers	63	Insurance carriers	63
Top positive portfolio weights	***SIC2***	***Top positive portfolio weights***	***SIC2***
Railroad transportation	40	Chemical & allied products	28
Transportation by air	45	Textile mill products	22
Furniture & homefurnishings stores	57	General merchandise stores	53
Textile mill products	22	Lumber & wood products	24
Building materials & gardening supplies	52	Building materials & gardening supplies	52
Tobacco products	21	Tobacco products	21

This table shows the industries (2-digit SIC code) with the largest average short and long positions in the estimated hedge portfolios resulting from regressions presented in Tables 1 and 2. Panel A explores hedge portfolios based on regression 4 using innovations in the WSJ Climate Change News Index as CC_t, and panel B explores hedge portfolios based using innovations in the CH Negative Climate Change News Index as CC_t. All portfolios are constructed using the absolute demeaned value of the E-Scores. Within each portfolio, industries are arranged in ascending order of portfolio weights.

the hedge portfolio constructed using the ranked E-Scores. For the portfolio constructed using Sustainlytics E-Scores to hedge innovations in the CH Negative Climate Change News Index, for example, the largest short position is "General Building Contractors," followed by "Water Transportation." The largest long positions are "Building Materials & Gardening Supplies" and "Tobacco Products." This analysis highlights that the resultant hedge portfolios will not necessarily conform with common priors that the optimal way to hedge climate change news involves primarily going long green energy stocks and short oil companies; this is consistent with our observation that industry membership can only explain a small amount of the cross-sectional variation in firm-level E-Scores.

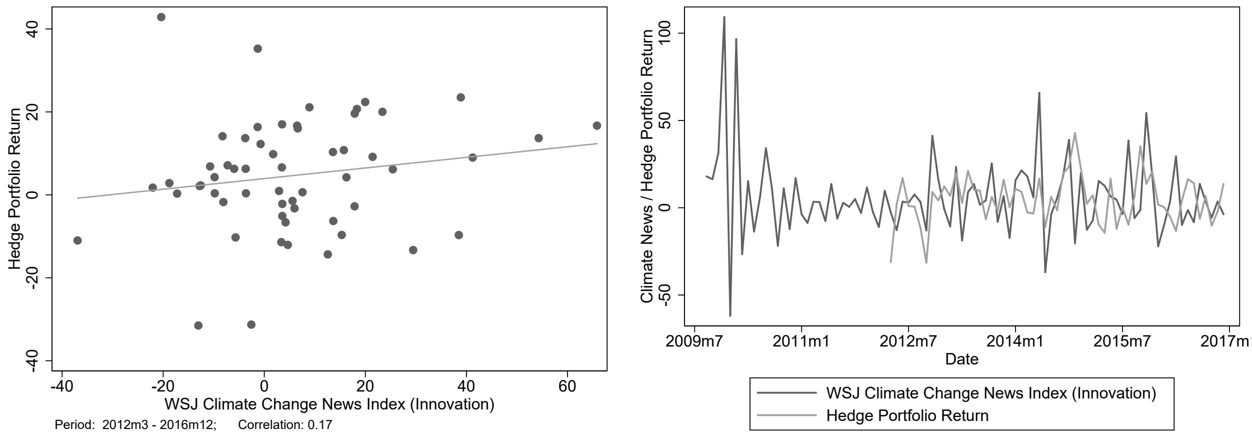
2.6 Out-of-sample fit results

The most important test of the hedge portfolios is their ability to hedge out-of-sample innovations to climate news, that is, to hedge innovations in months that were not included in the estimation of the portfolio weights. To construct a first measure of the out-of-sample performance of the hedge portfolios, for every period t we run regression 4 using data between periods t_{min} and $t-1$, where t_{min} corresponds to the first month for which we observe all climate exposures and CC_t series (September 2009). We then form the hedge portfolio based on these estimates and explore the correlation of the returns of that hedge portfolio in period t with CC_t. This corresponds to the approach one would have taken to hedge climate news in real time. Because we require a certain amount of data to estimate regression 4, we only compare the out-of-sample performance of the hedge portfolios starting in period $t_{min}+30$ (March 2012).[19]

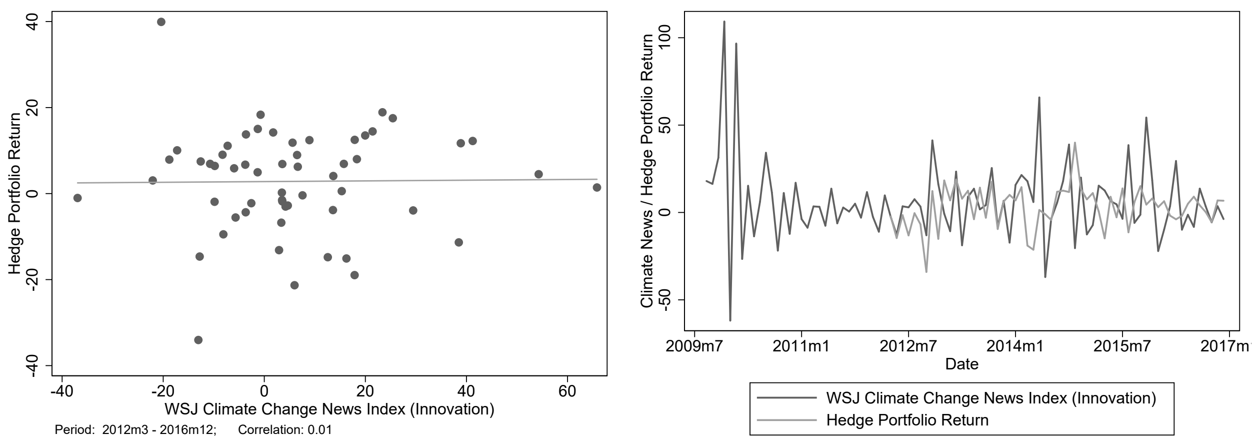
Figure 7 presents the out-of-sample performance of portfolios constructed to hedge innovations in the WSJ Climate Change News Index. The top panels show portfolios constructed using absolute values of the Sustainalytics E-Score, and the bottom panels show portfolios that build on the absolute values of the MSCI E-Score. The left columns present scatterplots of the out-of-sample returns of the hedge portfolios together with the

[19]Further reducing the number of portfolios onto which to project CC_t may improve the out-of-sample performance of the hedging portfolio. Given the short sample size available, in this paper we decided to not optimize the hedge portfolio further along this dimension.

Figure 7: Out-of-sample fit: WSJ Climate Change News Index



(a) Sustainalytics hedge portfolio



(b) MSCI hedge portfolio

This figure explores the out-of-sample performance of hedge portfolios constructed to hedge the *WSJ-Based Climate News Measure*. The top panel presents hedge portfolios built on the absolute values of the Sustainalytics E-Score, and the bottom panel presents portfolios built on the absolute values of the MSCI E-Score.

Table 4: Cross-correlations: WSJ Climate Change News Index

A. Out-of-sample fit

	CC^{WSJ}	$H_{OOS}^{SUS\_A}$	$H_{OOS}^{SUS\_R}$	$H_{OOS}^{MSCI\_A}$	$H_{OOS}^{MSCI\_R}$	H_{OOS}^{ETF}	r_t^{XLE}	r_t^{PBD}
CC^{WSJ}	1.000							
$H_{OOS}^{SUS\_A}$	0.174	1.000						
$H_{OOS}^{SUS\_R}$	0.206	0.973	1.000					
$H_{OOS}^{MSCI\_A}$	0.013	0.688	0.621	1.000				
$H_{OOS}^{MSCI\_R}$	0.019	0.677	0.624	0.988	1.000			
H_{OOS}^{ETF}	-0.005	0.427	0.349	0.861	0.852	1.000		
r_t^{XLE}	0.068	-0.138	0.004	-0.097	-0.039	-0.141	1.000	
r_t^{PBD}	0.111	0.185	0.272	0.294	0.350	0.190	0.656	1.000

B. Cross-validation fit

	CC^{WSJ}	$H_{Cross}^{SUS\_A}$	$H_{Cross}^{SUS\_R}$	$H_{Cross}^{MSCI\_A}$	$H_{Cross}^{MSCI\_R}$	H_{Cross}^{ETF}	r_t^{XLE}	r_t^{PBD}
CC^{WSJ}	1.000							
$H_{Cross}^{SUS\_A}$	0.244	1.000						
$H_{Cross}^{SUS\_R}$	0.300	0.976	1.000					
$H_{Cross}^{MSCI\_A}$	0.039	0.742	0.671	1.000				
$H_{Cross}^{MSCI\_R}$	0.067	0.733	0.676	0.982	1.000			
H_{Cross}^{ETF}	-0.069	0.454	0.390	0.678	0.651	1.000		
r_t^{XLE}	0.068	0.041	0.072	-0.009	-0.034	0.297	1.000	
r_t^{PBD}	0.111	0.272	0.266	0.310	0.298	0.470	0.656	1.000

This table shows cross-correlations of different portfolios and innovations in the WSJ Climate Change News Index. Panel A focuses on the performance of hedge portfolios from our out-of-sample approach, and panel B focuses on the performance of hedge portfolios from our cross-validation approach.

realizations of the innovation of climate news. The right panels plot the time series of the climate news series and the return series of the hedge portfolios. There is a clear, positive out-of-sample correlation with CC_t of 0.17 for the Sustainalytics hedge portfolio. In other words, the hedge portfolios indeed have higher returns during periods with positive innovations to climate news. Portfolios based on MSCI E-Scores or ETFs, on the other hand, have very little ability to hedge innovations in the WSJ Climate Change News Index, with an out-of-sample correlation of just 0.01.

Panel A of Table 4 provides additional information about the out-of-sample performance of the various portfolios designed to hedge innovations in the WSJ Climate Change News Index. The first column is the most important one, showing the correlation between the realizations of CC_t^{WSJ} and the returns of the various hedge portfolios (e.g., $R_{OOS}^{SUS\_A}$ corresponds to the out-of-sample returns of a hedge portfolio constructed using absolute values of the Sustainalytics E-Score). The hedge portfolios based on Sustainalytics E-Scores substantially outperform the hedge portfolios based on the MSCI E-Scores. In addition, hedge portfolios based on ranked E-Scores marginally outperform those based on absolute E-Scores, though the returns of portfolios based on absolute and ranked E-Scores from the same data provider are highly correlated. Finally, the out-of-sample performance of the Sustainalytics E-Score-based hedge portfolios is substantially better than that of portfolios based on ETFs. The returns of most hedge portfolios are *negatively* correlated with the returns to XLE, suggesting that these hedge portfolios are likely to hold short positions in the energy firms that constitute XLE. Similarly, we observe a positive correlation between the returns of all climate hedge portfolios and the returns of PBD, suggesting that the hedge portfolios likely hold long positions in many of the green energy firms that constitute PBD.

We also conduct a second test for the performance of the hedge portfolios based on a cross-validation approach. In particular, for every period t' we run regression 4 for all periods $t \neq t'$, and then use the resultant estimates to construct a hedge portfolio in a similar way as described above. The return of that hedge portfolio in period t' is then compared to $CC_{t'}$. Panel B of Table 5 explores the cross-validation performance of the various hedge portfolios. The hedge portfolios based on Sustainalytics E-Scores continue to outperform those based on MSCI E-Scores or ETFs substantially.

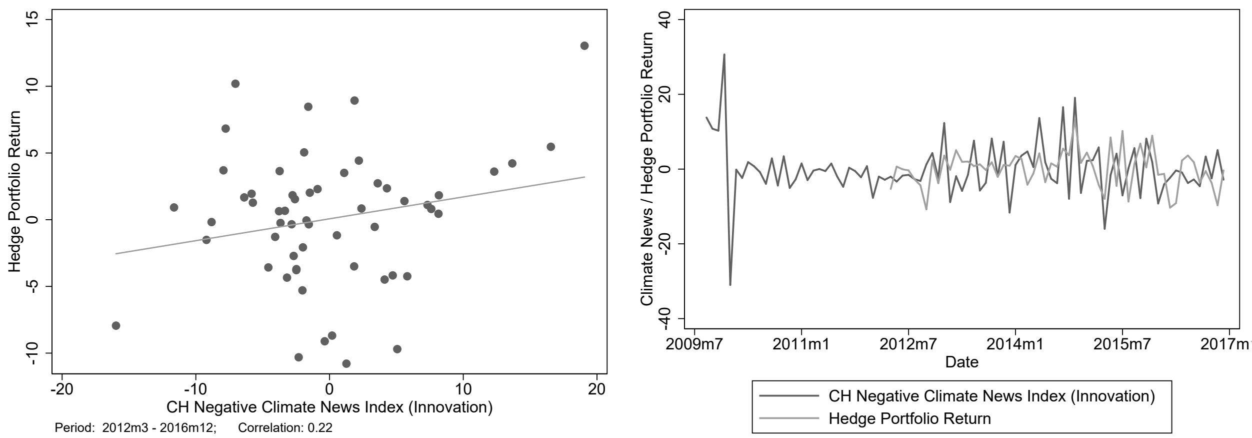
In sum, the hedge portfolios built using the Sustainalytics E-Score perform out of sample substantially better than any other hedge portfolio we have considered. The worse hedging performance of portfolios based on MSCI E-Scores highlights the importance of choosing characteristics that properly capture cross-sectional variation in exposure to climate change risks.

Figure 8 and Table 5 present results similar to those in Figure 7 and Table 4, but analyze the performance of portfolios designed to hedge innovations in the CH Negative Climate Change News Index. Portfolios based on Sustainalytics E-Scores have a similar ability to hedge this second climate news series as they had in hedging the CH Negative Climate Change News Index, both in the out-of-sample evaluation and in the cross-validation evaluation. The hedging ability of the MSCI indexes is in this case much higher than for the WSJ measure of climate change risks, suggesting that the MSCI E-Scores are more suited to capture negative climate change news as opposed to general coverage of climate change by the WSJ. Overall, the out-of-sample correlation between realization of climate change news and the hedge portfolios are 0.22 when using Sustainalytics E-Scores and 0.18 when using MSCI E-Scores.

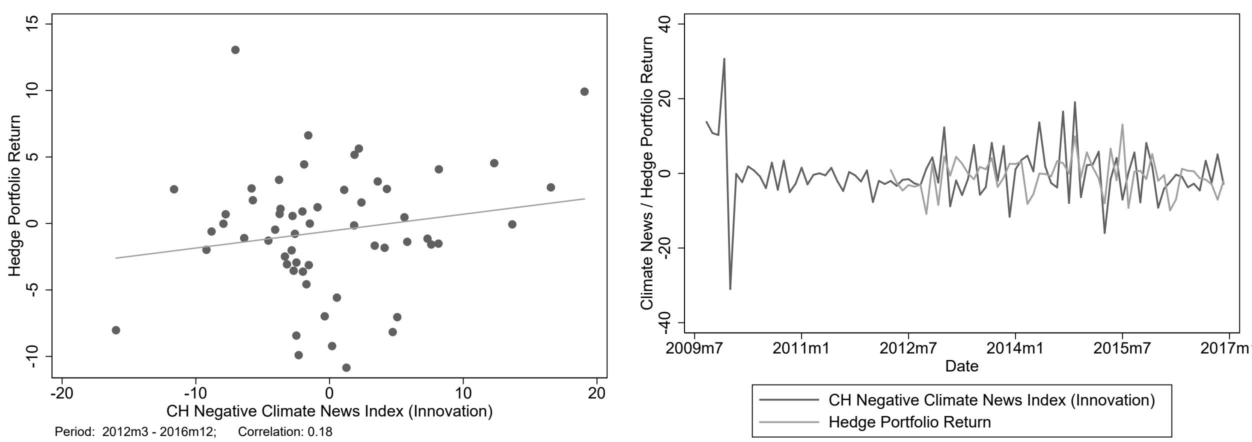
3 Conclusion and Directions for Future Research

We demonstrate how a mimicking portfolio approach can be successful in hedging innovations in climate change news across a number of out-of-sample performance tests. Across our two indices for climate news, the hedge portfolios based on Sustainalytics E-Scores have the best in-sample fit as well as the best out-of-sample and cross-validation performance. Portfolios based on MSCI E-Scores and ETFs have a lower (but still positive) ability to hedge innovations in climate news. There are no systematic differences in the relative performance of hedge portfolios based on absolute or ranked versions of the raw E-Scores. In general, however, the differences between the out-of-sample and cross-validation performance of some of the portfolios highlight that the portfolios we construct are somewhat sensitive to the exact time series on which our models are trained. This is

Figure 8: Out-of-sample fit: CH Negative Climate Change News Index



(a) Sustainalytics hedge portfolio



(b) MSCI hedge portfolio

This figure explores the out-of-sample performance of hedge portfolios constructed to hedge the *Newspaper-based negative climate news* measure. The top panel presents hedge portfolios built on the absolute values of the Sustainalytics E-Score, and the bottom panel presents portfolios built on the absolute values of the MSCI E-Score.

Table 5: Cross-correlations: CH Negative Climate Change News Index

A. Out-of-sample fit

	$CC^{NegNews}$	$H_{OOS}^{SUS\_A}$	$H_{OOS}^{SUS\_R}$	$H_{OOS}^{MSCI\_A}$	$H_{OOS}^{MSCI\_R}$	H_{OOS}^{ETF}	r_t^{XLE}	r_t^{PBD}
$CC^{NegNews}$	1.000							
$H_{OOS}^{SUS\_A}$	0.217	1.000						
$H_{OOS}^{SUS\_R}$	0.183	0.992	1.000					
$H_{OOS}^{MSCI\_A}$	0.179	0.869	0.852	1.000				
$H_{OOS}^{MSCI\_R}$	0.175	0.865	0.850	0.998	1.000			
H_{OOS}^{ETF}	0.157	0.780	0.767	0.961	0.960	1.000		
r_t^{XLE}	-0.066	-0.412	-0.353	-0.387	-0.367	-0.410	1.000	
r_t^{PBD}	0.063	0.061	0.112	0.096	0.127	0.119	0.656	1.000

B. Cross-validation fit

	$CC^{NegNews}$	$H_{Cross}^{SUS\_A}$	$H_{Cross}^{SUS\_R}$	$H_{Cross}^{MSCI\_A}$	$H_{Cross}^{MSCI\_R}$	H_{Cross}^{ETF}	r_t^{XLE}	r_t^{PBD}
$CC^{NegNews}$	1.000							
$H_{Cross}^{SUS\_A}$	0.148	1.000						
$H_{Cross}^{SUS\_R}$	0.154	0.991	1.000					
$H_{Cross}^{MSCI\_A}$	0.024	0.864	0.836	1.000				
$H_{Cross}^{MSCI\_R}$	0.048	0.885	0.861	0.993	1.000			
H_{Cross}^{ETF}	0.053	0.829	0.799	0.973	0.968	1.000		
r_t^{XLE}	-0.066	-0.208	-0.183	-0.205	-0.237	-0.223	1.000	
r_t^{PBD}	0.063	0.169	0.171	0.158	0.157	0.185	0.656	1.000

This table shows cross-correlations of different portfolios and innovations in the CH Negative Climate Change News Index. Panel A focuses on the performance of hedge portfolios from our out-of-sample approach, and panel B focuses on the performance of hedge portfolios from our cross-validation approach.

likely the result of only having a relatively few data points in each of our estimations. As we observe longer time series of E-Scores and climate news measures, our proposed method should deliver ever-better portfolios to hedge climate change news. Similarly, moving from hedging climate news that materializes over a monthly level to hedging on a daily level should allow researchers to substantially expand their training data, and thereby improve the out-of-sample performance of the hedge portfolios.

More generally, we view this article as providing a rigorous methodology for constructing portfolios that hedge against risks that are otherwise difficult to insure. We do not view our resultant hedge portfolios as the definitive best hedges against climate change risk, but instead as a starting point for further exploration. Indeed, future research could consider many valuable directions for climate finance, and we discussed a number of the dimensions that should be explored further, including the addition of more assets to the hedge portfolios (such as international stocks) and the formation of hedge portfolios based on both characteristic-sorted portfolios and ETFs.

One additional important direction for future work is to integrate more and better data to measure firm-level climate risk exposures. These data could come from commercial data providers or could be constructed by researchers themselves, for example, by including information such as geographical proximity to potential climate disasters (e.g., rising sea levels or hurricane-prone regions). Indeed, articles in this volume, such as Choi et al. (2018) and Kumar et al. (2018) make valuable progress toward developing new ways to quantify climate risk exposures.

Another direction for follow-on work is to develop alternative definitions of the climate change risks. One interesting question is whether it is important to differentiate between physical and policy-oriented climate risks. For example, a tax on greenhouse gas emissions, if comprehensively applied at an appropriate level, would reduce the demand for climate hedge portfolios and consequently the cost of insuring against climate change. Thus, good regulation will mean less need for climate hedges. But regulation itself creates winners and losers from regulatory risk, and one might therefore want to construct regulatory hedge portfolios. The stability of such regulatory hedge portfolios may well be sensitive to the prevailing political environment.

A related question pertains to the expected returns of the various hedge portfolios. Indeed, an increasing use of climate hedge portfolios by investors will increase the price (and thus reduce the expected returns) of those firms whose stock provides the most effective hedge against innovations in climate change news. This lower expected return corresponds to the insurance premium paid for the climate hedge portfolio. An interesting avenue for future work will be to quantify the cost of the climate hedge portfolios by looking at the associated risk premiums.[20] It is also interesting to study the general equilibrium effects resulting from the fact that a lower cost of capital for firms with high E-Scores might actually have a direct effect on the climate trajectory. For example, to the extent that green energy firms see a reduction in their cost of capital, this might allow them to achieve efficient scale faster, and thereby affect the path of greenhouse gas emissions. The design of structural asset pricing models that feature such general equilibrium feedback loops seems a promising direction for research.

[20]Note that this requires substantial time-series data, because realizations of negative climate news in sample might actually lead the hedge portfolios to outperform over any given period.

References

Amihud, Y. 2002. Illiquidity and Stock Returns: Cross-Section and Time-Series Effects. *Journal of Financial Markets* 5:31–56.

Andersson, M., P. Bolton, and F. Samama. 2016. Hedging Climate Risk. *Financial Analysts Journal* 72.

Bakkensen, L., and L. Barrage. 2018. Flood Risk Belief Heterogeneity and Coastal Home Price Dynamics: Going Under Water? *Working Paper* .

Baldauf, M., L. Garlappi, and C. Yannelis. 2018. Does Climate Change Affect Real Estate Prices? Only If You Believe In it. *Working paper* .

Balduzzi, P., and C. Robotti. 2008. Mimicking Portfolios, Economic Risk Premia, and Tests of Multi-Beta Models. *Journal of Business and Economic Statistics* 26:354–368.

Bernstein, A., M. Gustafson, and R. Lewis. 2018. Disaster on the Horizon: The Price Effect of Sea Level Rise. *Journal of Financial Economics, Forthcoming* .

Black, F., and M. Scholes. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy* 81:637–654.

Breeden, D., M. Gibbons, and R. Litzenberger. 1989. Empirical Tests of the Consumption-Oriented CAPM. *Journal of Finance* 44:231–262.

Chen, N.-F., R. Roll, and S. Ross. 1986. Economic Forces and the Stock Market. *Journal of Business* 59:383–403.

Choi, D., Z. Gao, and W. Jiang. 2018. Attention to Global Warming. *Working paper* .

Daniel, K., R. Litterman, and G. Wagner. 2015. Applying Asset Pricing Theory to Calibrate the Price of Climate Risk. *Working paper, Columbia University* .

Fama, E., and K. French. 2008. Dissecting Anomalies. *Journal of Finance* 63:1653–1678.

Fama, F., and J. MacBeth. 1973. Risk, Return, and Equilibrium: Empirical Tests. *Journal of Political Economy* 81:607–636.

Gentzkow, M., B. T. Kelly, and M. Taddy. 2018. Text as data. *Working Paper 23276, National Bureau of Economic Research* .

Giglio, S., M. Maggiori, and J. Stroebel. 2015. Very Long-Run Discount Rates. *Quarterly Journal of Economics* 130:1–53.

Giglio, S., M. Maggiori, J. Stroebel, and A. Weber. 2018. Climate Change and Long-Run Discount Rates: Evidence from Real Estate. *Working paper* .

Giglio, S., and D. Xiu. 2018. Asset Pricing with Omitted Factors. *Working Paper* .

Hong, H., and L. Kostovetsky. 2012. Red and blue investing: Values and finance. *Journal of Financial Economics* 103:1–19.

Hong, H., F. Li, and J. Xu. 2018. Climate Risks and Market Efficiency. *Journal of Econometrics, forthcoming* .

Hou, K., C. Xue, and L. Zhang. 2015. Dissecting Anomalies: An Investment Approach. *Review of Financial Studies* 28:650–705.

Huberman, G., S. Kandel, and R. Stambaugh. 1987. Mimicking Portfolios and Exact Arbitrage Pricing. *Journal of Finance* 42:1–9.

Kumar, N., A. Shashwat, and R. Wermers. 2018. Do Fund Managers Misestimate Climatic Disaster Risk? *Working Paper* .

Lamont, O. 2001. Economic Tracking Portfolio. *Journal of Econometrics* 105:161–184.

Lönn, R., and P. Schotman. 2017. Empirical asset pricing with many assets and short time series. *Working Paper, Maastricht University* .

Merton, R. 1973. An Intertemporal Capital Asset Pricing Model. *Econometrica* 41:867–887.

Murfin, J., and M. Spiegel. 2018. Is the Risk of Sea Level Rise Capitalized in Residential Real Estate. *Working paper* .

Roll, R., and A. Srivastava. 2018. Mimicking Portfolios. *Working Paper, California Institute of Technology* .

A Appendix

A.1 Review of the Fama-MacBeth approach

In this section, we review the Fama-MacBeth estimator for hedge portfolios in the context of our model. To apply the Fama-MacBeth procedure, the econometrician needs to take a stand on all the factors in the model: CC_t and v_t. Once the factors in the model are determined, the procedure follows two steps. In the first step, the risk exposures β_{CC} and β are estimated via time-series regressions of returns onto the factors, CC_t and v_t. In particular, for each asset i, $(\hat{\beta}^i_{CC}, \hat{\beta}^i)$ are estimated from the time-series regression:

$$r^i_t = \alpha^i + \beta^i_{CC} CC_t + \beta^i v_t + u_t.$$

In the second step, in each period t, hedge portfolios for all factors are obtained via cross-sectional regressions of returns r_t onto the estimated betas $(\hat{\beta}_{CC}, \hat{\beta})$:

$$r_t = h^{CC}_t \hat{\beta}_{CC} + h_t \hat{\beta} + e_t,$$

where $\hat{\beta}_{CC}$ and $\hat{\beta}$ are the betas estimated in the first step. The slopes of this regression in each period t are precisely the returns of the hedge portfolio in period t: h^{CC}_t (that hedges CC_t) and h_t (that hedges the remaining factors v_t). The hedge portfolios h^{CC}_t and h_t have, by construction, a beta of one with respect to the corresponding factors and zero with respect to all other factors. Their time-series means (the expected excess returns of the hedge portfolios) recover the risk premiums of the factors: $E[h^{CC}_t] = \gamma_{CC}$ and $E[h_t] = \gamma$.

The Fama-MacBeth procedure for constructing hedge portfolios has two potential drawbacks. First, it requires knowing all the factors in the model, CC_t and v_t. Second, the procedure is not robust to measurement error in the factor of interest, CC_t, which is a natural concern in many settings, including in ours (see the further discussion of omitted factors and measurement error in Giglio and Xiu, 2018).

A.2 Source of Climate Change Vocabulary (CCV)

To create the Climate Change Vocabulary, we collect twelve climate change white papers from various sources including the Intergovernmental Panel on Climate Change (IPCC), Environmental Protection Agency (EPA), and the U.S. Global Change Research Program. We complement this with fifty-nine climate change glossaries from sources such as United Nations, NASA, IPCC, and EPA.

Twelve climate change white papers: Table A1 reports the institution, title and published year of climate change white papers that we use to construct the CCV.

Fifty-nine climate change glossaries: We collect climate change glossaries, both words and their definition, from U.S. Environmental Protection Agency (EPA), BBC, United Nations(UN), Center for Climate and Energy Solutions Glossary of Key Terms, Intergovernmental Panel on Climate Change (IPCC), World Health Organization (WHO), European Climate Adaptation Platform, International Petroleum Industry Environmental Conservation Association(IPIECA), Lenntech, Wikipedia, Met Office, Integrated Regional Information Networks(IRIN), Climate Change in Australia, Guardian, International Rivers, Mekong River Commission, Exploratorium, *New York Times*, U.S. Forest Service, U.S. Department of Transportation, Durham Region, Classroom of the Future, Government of Canada, International Food Policy Research Institute (IFPRI), New Zealand Government, University of Miami, German Climate Finance, California Government, South West Climate Change Impacts Partnership (SWCCIP), Scent of Pine, Natural Climate Change, UN Climate Change Conference, Center for Strategic and International Studies(CSIS), Watts Up With That?, U.K. Climate Impacts Programme (UKCIP), Climate Change Zambia, Canadian Broadcasting Corporation(CBC), Auburn University, Global Warming Solved, REDD+, Climate Resilience Toolkit(CRT), What's Your Impact, The Nitric Acid Climate Action Group (NACAG), Garnaut Climate Change Review, Climate Policy Information Hub, Explaining Climate Change, Four Degrees Preparation, The European Initiative for Upscaling Energy Efficiency in the Music Event Industry (EE MUSIC), Regional Education and Information Centre (REIC), Ecology, Climate Reality Project, National Geographic, Agricultural Marketing Resource Center (AgMRC), Global Greenhouse Warming, and Conservation in a Changing Climate.

Table A1: The list of climate change white papers

Source	Title	Year
IPCC	IPCC Synthesis Report	1990, 1995, 2001, 2007, 2014
IPCC	IPCC Special Report: The Regional Impacts of Climate Change: an assessment of vulnerability	1997
IPCC	IPCC Special Report: Aviation and the Global Atmosphere	1999
IPCC	IPCC Special Report: Methodological and Technological Issues in Technology Transfer	2000
IPCC	IPCC Special Report: Safeguarding the Ozone Layer and the Global Climate System: Issues Related to Hydrofluorocarbons and Perfluorocarbons	2005
IPCC	IPCC Special Report: Carbon Dioxide Capture and Storage	2005
IPCC	IPCC Special Report: Renewable Energy Sources and Climate Change Mitigation	2011
IPCC	IPCC Special Report: Managing the Risks of Extreme Events and Disasters to Advance Climate Change Adaptation	2012
American Association for the Advancement of Science	What We Know: The Reality, Risks, and Response to Climate Change	2014
UC Berkley	American Climate Prospectus	2015
U.S. EPA	Climate Change Indicators in the United States (4th edition)	2016
Science	Social and Economic Impacts of Climate	2016
IMF	The Effects of Weather Shocks on Economic Activity	2017
U.S. Global Change Research Program	Our Change Planet: The U.S. Global Change Research Program for Fiscal Year 2017	2017
U.S. Global Change Research Program	Climate Science Special Report (4th National Climate Assessment, Vol. I)	2017

IPCC reports scientific and technical assessments of the current state of climate change. Generally, these reports comprise three volumes: one for each of the Working Groups of the IPCC. In addition to the main reports, Summaries for Policymakers and Synthesis Reports are provided. A Synthesis Report integrates materials covered by Assessment Reports and Special Reports. It is a nontechnical report targeting policy makers and addressing a broad range of policy-relevant but policy-neutral questions. Summary for Policymakers is an abridged version of the full Synthesis Report. In addition, IPCC Special Reports provide an assessment of a specific issue relating to climate change. They are generally structured similar to a volume of an Assessment Report. IPCC, Intergovernmental Panel on Climate Change; EPA, Environmental Protection Agency; IMF, International Monetary Fund.

A.3 Subcategories for "E" scores

A.3.1 MSCI

Positive indicators are Environmental Opportunities - Clean Tech, Waste Management - Toxic Emissions and Waste, Waste Management - Packaging Materials & Waste, Climate Change - Carbon Emissions, Property/Plant/Equipment, Environmental Management Systems, Natural Resource Use - Water Stress, Natural Resource Use - Biodiversity & Land Use, Natural Resource Use - Raw Material Sourcing, Natural Resource Use - Financing Environmental, Environmental Opportunities - Green Buildings, Environmental Opportunities in Renewable Energy, Waste Management - Electronic Waste, Climate Change - Energy Efficiency, Climate Change - Product Carbon Footprint, Climate Change - Insuring Climate Change Risk, Environment - Other Strengths.
Negative indicators are Regulatory Compliance, Toxic Emissions and Waste, Energy & Climate Change, Impact of Products and Services, Biodiversity & Land Use, Operational Waste, Water Stress, Environment - Other Concerns.

A.3.2 Sustainalytics

Subcategories are Formal Environmental Policy, Environmental Management System, External Certification of Environmental Management Systems (EMS), Environmental Fines and Non-monetary Sanctions, Participation in Carbon Disclosure Project, Scope of Corporate Reporting on GHG emissions, Programmes and Targets to Reduce GHG Emissions from Own Operations, Programmes and Targets to Increase Renewable Energy Use, Carbon Intensity, Carbon Intensity Trend, % of Primary Energy Use from Renewables, Operations Related Controversies or Incidents, Reporting Quality Non-Carbon Environmental Data, Programmes and Targets to Protect Biodiversity, Guidelines and Reporting on Closure and Rehabilitation of Sites, Environmental and Social Impact Assessments, Oil Spill Reporting and Performance, Waste Intensity, Water Intensity, Percentage of Certified Forests Under Own Management, Programmes & Targets to Reduce Air Emissions, Programmes & Targets to Reduce Air Emissions, Programmes & Targets to Reduce Water Use, Other Programmes to Reduce Key Environmental Impacts, GHG Reduction Programme, Programmes and Targets to Improve the Environmental Performance of Own Logistics and Vehicle Fleets, Programmes and Targets to Phase out CFCs and HCFCs[21] in Refrigeration Equipment, Formal Policy or Programme on Green Procurement, Environmental Supply Chain Incidents, Programmes to Improve the Environmental Performance of Suppliers, External Environmental Certification Suppliers, Programmes and Targets to Stimulate Sustainable Agriculture, Programmes and Targets to Stimulate Sustainable Aquaculture/Fisheries, Food Beverage & Tabacco Industry Initiatives, Programmes and Targets to Reduce GHG Emissions from Outsourced Logistics Services, Data on Percentage of Recycled/Reused Raw Material Used, Data on Percentage of Forest Stewardship Council (FSC) Certified Wood/Pulp as Raw Material, Programmes and Targets to Promote Sustainable Food Products, Food Retail Initiatives, Products & Services Related to Controversies or Incidents, Sustainability Related Products & Services, Revenue from Clean Technology or Climate Friendly Products, Automobile Fleet Average CO2 Emissions, Trend Automobile Fleet Average Fleet Efficiency, Products to Improve Sustainability of Transport Vehicles, Systematic Integration of Environmental Considerations at R&D Stage, Programmes and Targets for End-of-Life Product Management, Organic Products, Policy on Use of Genetically Modified Organisms (GMO) in Products, Environmental & Social Standards in Credit and Loan Business, Responsible Asset Management, Use of Life-Cycle Analysis(LCA) for New Real Estate Projects, Programmes and Targets to Increase Investment in Sustainable Buildings, Share of Property Portfolio Invested in Sustainable Buildings, Sustainability Related Financial Services, Products with Important Environmental/Human Health Concerns, Carbon Intensity of Energy Mix, Mineral Waste Management, Emergency Response Programme.

[21]CFCs refers to chlorofluorocarbons, and HCFCs refers to Hydrochloroflourocarbons.

Real Effects of Climate Policy:

Financial Constraints and Spillovers

Söhnke M. Bartram*, **Kewei Hou†**, **and Sehoon Kim‡**

Abstract

We document that localized policies designed to mitigate climate risk can lead to regulatory arbitrage by firms, resulting in unintended consequences. Using detailed plant level data, we investigate the impact of the most extensive regional climate policy in the United States, the California cap-and-trade program, on corporate real activities such as greenhouse gas emissions and plant ownership. We show that industrial plants governed by the policy reduce emissions in California when the parent company is financially constrained, but that these firms internally reallocate their emissions to plants located in other states. Similarly, constrained firms are more likely to reduce ownership in Californian plants and increase ownership in plants outside California. In contrast, unconstrained firms generally do not adjust plant emissions and ownership either in California or in other states. Overall, firms do not reduce their total emissions when part of their assets are affected by the regulation, but in fact increase them if financially constrained. The results document real spillover effects stemming from resource reallocations by constrained firms to avoid regulatory costs, undermining the effectiveness of localized policies. Our study has important implications for the current debate on global climate policy agreements.

Keywords: Climate policy, cap-and-trade, financial constraints, internal resource allocation, real effects, regulatory arbitrage, externalities, spillover effects, corporate environmental policy
JEL Classification: G18, G31, G32, Q52, Q54, Q58
First version: October 7, 2018
This version: February 11, 2019

---

* Professor of Finance, University of Warwick, Warwick Business School, Department of Finance, Coventry CV4 7AL, United Kingdom, Phone: +44 (24) 7657 4168, Fax: +1 (425) 952 10 70, E-mail: <s.m.bartram@wbs.ac.uk>, Internet: <http://go.warwick.ac.uk/sbartram/>.

† Ric Dillon Endowed Professor in Investments, The Ohio State University, 820 Fisher Hall, 2100 Neil Avenue, Columbus, OH 43210, Phone: +1 (614) 292-0552, E-mail: <hou.28@osu.edu>

‡ Assistant Professor of Finance, University of Florida, Warrington College of Business, PO Box 117168, Gainesville, FL 32611-7168, E-mail: <sehoon.kim@warrington.ufl.edu>.

The authors are Fellows of the Risk Institute at The Ohio State University Fisher College of Business, and acknowledge the Institute's support for this research. Söhnke Bartram gratefully acknowledges financial support by the British Academy/Leverhulme Trust. We thank Mark Flannery, Christopher James, Jay Ritter, and René Stulz for valuable comments and suggestions. Shema Mitali provided excellent research assistance.

Real Effects of Climate Policy:

Financial Constraints and Spillovers

Abstract

We document that localized policies designed to mitigate climate risk can lead to regulatory arbitrage by firms, resulting in unintended consequences. Using detailed plant level data, we investigate the impact of the most extensive regional climate policy in the United States, the California cap-and-trade program, on corporate real activities such as greenhouse gas emissions and plant ownership. We show that industrial plants governed by the policy reduce emissions in California when the parent company is financially constrained, but that these firms internally reallocate their emissions to plants located in other states. Similarly, constrained firms are more likely to reduce ownership in Californian plants and increase ownership in plants outside California. In contrast, unconstrained firms generally do not adjust plant emissions and ownership either in California or in other states. Overall, firms do not reduce their total emissions when part of their assets are affected by the regulation, but in fact increase them if financially constrained. The results document real spillover effects stemming from resource reallocations by constrained firms to avoid regulatory costs, undermining the effectiveness of localized policies. Our study has important implications for the current debate on global climate policy agreements.

Keywords: Climate policy, cap-and-trade, financial constraints, internal resource allocation, real effects, regulatory arbitrage, externalities, spillover effects, corporate environmental policy

JEL Classification: G18, G31, G32, Q52, Q54, Q58

"Climate policy advocates need to do a much better job of quantitatively analyzing economic costs and the actual, rather than symbolic, benefits of their policies. Skeptics would also do well to focus more attention on economic and policy analysis… We need to know what effect proposed policies have and at what cost. Scientific, quantifiable or even vaguely plausible cause-and-effect thinking are missing from much advocacy for policies to reduce carbon emissions."

David R. Henderson and John H. Cochrane, *July 2017, Wall Street Journal*

1 Introduction

Climate change is among the most intensely debated socio-economic issues of current times.[1] As a response to potential catastrophe risks from climate change, governments around the world are pushing for various forms of regulations to curb greenhouse gas emissions.[2] However, there is far from a consensus on what the optimal policy approach might be, and as a result climate policies are highly fragmented across the jurisdictions in which they are designed and implemented. More importantly, it is unknown whether such localized yet uncoordinated policies are able to internalize potential externalities that may impede addressing climate change as a global phenomenon or simply distort allocations in the economy. An example is the United States, where at the beginning of 2013, California became the first and only state to put an extensive mandatory carbon regulation in place in the form of a cap-and-trade system that applies universally to all industrial greenhouse gas emissions.[3] Exploiting the introduction of the California cap-and-trade rule, we investigate the internal resource allocation responses by firms and the real but unintended spillover effects of localized climate policies that arise from the importance of financial constraints. Our study helps understand the interplay between climate policy and firm behavior, and informs policy makers regarding the effectiveness of climate regulation.

Using a detailed plant level dataset on greenhouse gas emissions and parent company ownership made available by the US Environmental Protection Agency (EPA) and hand-matched to Compustat covering 2,806

[1] See Figure 1 for recent trends in carbon emissions from the use of fossil fuels and global temperatures.

[2] See Figure 2 for a map of implemented or planned carbon pricing regulations around the world.

[3] Most climate regulations in the United States thus far have left states with much discretion in implementing federal standards (e.g. Clean Air Act) or have largely been confined to the electricity production industry. Since 2009, nine states (Connecticut, Delaware, Maine, Maryland, Massachusetts, New Hampshire, New York, Rhode Island, and Vermont) have been part of the Regional Greenhouse Gas Initiative (RGGI), a cap-and-trade program that applies only to fossil fuel power plants generating 25MW or more. States have also been adopting varying versions of Renewable Portfolio Standards (RPS) requiring increased production of energy from renewable energy sources. From 2003 to 2010, the Chicago Climate Exchange (CCX) was available for voluntary emissions trading, but ceased trading due to inactivity.

industrial plants and 511 publicly listed non-utility and non-governmental firms over the sample period 2010 to 2015, we show that the 2013 California cap-and-trade rule has real spillover effects across the United States through firm financial constraints. Specifically, we employ a difference-in-differences (DID) framework and find that while financially constrained firms reduce greenhouse gas emissions from plants located in California by 35% relative to plants in other states, they significantly increase emissions from plants in other states by 29% more compared to those owned by firms without a presence in California. In contrast, we find no evidence that financially unconstrained firms adjust plant emissions in response to the new regulation, neither in California nor in other states.

We also find that compared to unconstrained firms, constrained firms are less likely to invest in plants in California (14.5% more likely to close and 8% less likely to open a plant), while they are more likely to invest in plants in other states (18% more likely to open and 6% less likely to close). However, these adjustments at the extensive margin are somewhat weaker than the emission shifts at plants already in place. The differences in responses between constrained and unconstrained firms are statistically significant across a host of financial constraint measures. Finally, we provide evidence that firms whose assets are partially affected by the regulation do not reduce their firm-wide emissions. In fact, constrained firms increase their total emissions by as much as 19%. Overall, our main results are consistent with the internal reallocation of corporate pollutive activities and resources to avoid regulatory costs in the face of limited access to external financing, and highlight the hidden costs of environmental policies through financial channels.

Our economic hypothesis is that financially constrained firms reallocate their greenhouse gas emissions and plant ownership away from California to other states in the face of heightened regulatory costs that alter the relative net expected returns across plants, because the costs of external capital render optimal levels of emissions in California infeasible and hence the net returns from internal reallocations more attractive. This conjecture is rooted in the literature in finance that studies the relationship between financial frictions and the value of internal capital allocation. This literature has shown that the contribution of internal capital markets to firm value and hence the value of corporate diversification is greater when external financial constraints are higher (see Billett and Mauer, 2003; Matvos and Seru, 2014; Matvos, Seru, and Silva, 2018). Empirically, the additional costs of greenhouse gas emissions to firms under the California cap-and-trade rule amounts to 15 basis points of the median firm's total assets and is comparable to the magnitude of the interest obligations on its short-term debt. We hypothesize that this increase in regulatory cost stemming from some of the firms' operations generates a wedge in the accessible set of net returns on capital between constrained and unconstrained firms, incentivizing constrained firms to reallocate even though emitting in California might have remained profitable had they been unconstrained. Looking at greenhouse gas emissions and ownership stakes in plants as outcome variables, we document that this is indeed strongly the case for reallocations of emissions across plants already in place, and to a lesser degree also true for changes in plant ownership.

Our conjecture and findings are consistent with criticisms by the media and small business owners that the increased regulatory costs from the cap-and-trade rule are not large enough to constitute significant deterrents to emissions for firms with deep pockets, but raise the burden for less financially capable players which may cause emission leakages.[4] Anecdotal evidence also supports the economic importance of the spillover effects we uncover. For example, Valero Energy, a major petroleum products company that was just recovering from large operating losses after the financial crisis in the early 2010s, strongly objected to the implementation of the cap-and-trade rule. It rallied other companies and warned citizens with placards at their California gas pumps that "the cap-and-trade rule would be a loss of two blue-collar jobs for every one green job created" and that "if the cap-and-trade legislation is passed, you will pay the price… Cap-and-trade will cost you 77 cents or more a gallon".[5] After the cap-and-trade rule went into effect at the beginning of 2013, Valero reduced the emissions by one of its largest Californian refineries in Los Angeles County by 8% over the next three years,

---

[4] In July 2017, as the cap-and-trade rule was about to be extended, the California state executive director of the National Federation of Independent Business (NFIB) stated on behalf of 22,000 small business members that as "California has been experimenting with cap-and-trade policies… jobs are moving to neighboring states with much more relaxed laws… Some believe cap-and-trade only impacts big businesses that buy and sell carbon credits, but the truth is that small businesses and consumers all pay the ultimate price." An October 2017 Wall Street Journal opinion piece, "The fatal flaw in California's cap-and-trade program" by Richard Sexton and Steven Sexton, criticized the cap-and-trade rule for its inability to effectively curtail carbon leakage and its failure to levy large enough burdens to large firms.

[5] See "Valero pumps up cap-and-trade debate" in CS News (September 2009), "Why did Valero launch a campaign against California's climate law?" in Los Angeles Times (October 2010), and "New energy outfoxes old in California" in The New York Times (November 2010) for media reference of Valero's reaction to California's plan for its cap-and-trade policy.

but sharply increased emissions by some of its largest refineries in other states, for example in New Orleans LA and Jefferson TX, by more than 10%.

We interpret our findings as optimal responses by firms to increased regulatory costs as a function of their financial constraints. Hence, we are comfortable with the fact that constrained and unconstrained firms are not randomly assigned their constraint characteristics, insofar as the assignment is not related to whether firms own plants covered by the California cap-and-trade rule in a way that permits confounding explanations for why firms respond to the rule in the way we document. Nevertheless, we exclude a number of alternative channels that may confound the interpretation of our results. To eliminate the possibility of reverse causality whereby financial constraints are affected by the introduction of the cap-and-trade rule or firm responses to it, or omitted variables simultaneously affecting constraints and firm responses, we measure financial constraints based on information reported strictly before our sample period and at least 3 years before the effective start date of the cap-and-trade rule. We also rule out explanations concerning observed or unobserved plant characteristics such as their industry purpose, maximum capacity, or technological obsoleteness by controlling for plant fixed effects, and preclude the effects of common time trends by controlling for year fixed effects. As further robustness checks, we also report results controlling for firm-by-plant fixed effects in the Appendix. Finally, we also control for firm characteristics that may be related to how much greenhouse gas firms are prone to release, such as the firm's property and plants or R&D stock. In short, we set a high bar to refute our conclusion that the cap-and-trade rule entails spillover effects due to the internal reallocation by financially constrained firms.

By providing concrete evidence on the real spillover effects from localized climate policies, our study makes an important contribution to the debate in economics about the impact of climate change and how to counter it with policy going back to the 1970s at the least (see Nordhaus, 1977a; 1977b). With the recent rise of social interest in climate change, more researchers are evaluating or prescribing energy related policies and are studying their implications (see Fabra and Reguant, 2014; Fowlie, Greenstone, and Wolfram, 2018; Khan, Knittel, Metaxoglou, and Papineau, 2016; Marin, Marino, and Pellegrin, 2018; Coria and Jaraite, 2015). An important aspect of climate policies, as with most energy policies in general, is that they are prone to externalities due to the fact that regulations are implemented locally while their outcomes impact climate globally. While much attention has been paid to the implications and potential distortions arising from coordination problems in climate change policy, most of these studies are based on model simulations using parameter estimates (see Nordhaus and Yang, 1996; Nordhaus, 2015; Martin, Muûls, De Preux, and Wagner, 2014; Fowlie, Reguant, and Ryan, 2016; Bushnell, Holland, Hughes, and Knittel, 2017). On the other hand, empirical evidence that speaks to the economic magnitudes and channels of externalities from coordination issues of climate change policy is scarce, and our study aims to fill this gap.

While empirical analysis of the economic impact and spillover effects of localized environmental policies across countries have garnered interest in the "pollution haven" literature in environmental economics, most of these studies have been limited to aggregate level cross-industry or cross-country analysis (see, among others, Ederington, Levinson, and Minier, 2005; Cole and Elliott, 2005; Levinson and Taylor, 2008; Wagner and Timmins, 2009; Kellenberg, 2009; Mulatu, Gerlagh, Rigby, and Wossink, 2010), indirect inference through the location and size of plants rather than pollutive activities per se (see Becker and Henderson, 2000), or self-reported survey data on CO_2 emissions aggregated at the firm-country level (see Ben-David, Kleimeier, and Viehs, 2018). Moreover, these studies have not focused on climate change policy, which differs from general environmental policies in that its explicit objective is to solve the problem of global warming rather than local pollution, the effectiveness of which critically hinges on the identification of unintended consequences across regulatory jurisdictions. Our paper has several key distinctions from these studies: (a) We utilize mandatorily reported data on plant level CO_2e greenhouse gas emissions and parent ownership, (b) we are able to exploit both within and between plant variation by focusing on the introduction of a local policy in the middle of our sample period whose clear mandate is to curb greenhouse gas emissions in one state but not others such that a DID analysis is possible, (c) we provide the first climate policy evaluation directly relevant for policy makers in the United States, (d) and most importantly, we emphasize firm financial constraints as an important economic channel for our results.

More specifically, our study contributes to the growing field of corporate environmental policies by focusing on the internal allocation of plant level emission activities and ownership within firms, thus providing a unique channel for the real effects of climate policy through the importance of firm financial constraints. Closely related to our work is a contemporaneous paper by Kim and Xu (2018) who show that financial constraints negatively affect corporate environmental policies due to the costs of waste management, and that this effect is stronger when regulatory monitoring is weak. Likewise, Shive and Forster (2018) argue that short-termist pressure for financial performance from outside investors force public firms to pollute more than private firms. In a similar vein, Akey and Appel (2018) find that firm subsidiaries are more likely to increase toxic emissions when parent companies have better liability protection for their subsidiaries' environmental clean-up costs, consistent with the binding effects of higher financial burdens associated with abatement. Complementing these studies, our findings highlight the reallocative effects of financial constraints inducing firms to internally shift their pollutive resources across plants under heightened regulatory costs, which in turn distort the outcome of regional environmental policies. Interestingly, while Akey and Appel (2018) find the effects of limited liability to be driven by lower "green" investments rather than by reallocation across plants, we show that the reallocation of greenhouse gas emissions and ownership stakes across plants are prominent responses by firms to climate policy, particularly when financial constraints are high.

Remedying policies to climate change risk are heatedly debated. Such policies have important implications for financial economists because among the primary subjects of regulation are private industrial firms. Since policies aimed at mitigating climate risk introduce regulatory frictions to firms, it is important to understand how firms react to and cope with such frictions. This will guide policy makers to internalize externalities that may otherwise result in unintended consequences and more effectively coordinate solutions to climate change. Given the importance of a sound evaluation of the efficacy and real effects of climate policy, it is surprising how little empirical work has been done on this front. This paper aims to take the debate on climate change, climate policy, and corporate environmental responsibility one step closer in this direction.

Our paper is organized as follows. In Section 2, we summarize the relevant details of the California cap-and-trade rule and develop our economic hypotheses. We describe our data sources and summarize the composition and characteristics of our sample in Section 3. Section 4 motivates our empirical research design, while Section 5 presents our results. We conclude in Section 6 with a brief summary and some policy guidance.

2 Background and Hypothesis Development

2.1 California Cap-and Trade

At the beginning of 2013, the state of California's Air Resources Board started enforcing a state-wide carbon cap-and-trade rule to reduce greenhouse gas emissions. Covering all electric power plants and industrial plants that emit 25,000 metric tons or more of carbon dioxide equivalents (CO_2e) per year, the California cap-and-trade was the first multi-sector cap-and-trade program in North America.[6] The California cap-and-trade rule is based on an allocation of capped allowances with specific year vintages and the market trading of those allowances. At the allocation stage, allowances are distributed to plants through a combination of quarterly held auctions and free allowances. Firms are then required to pay off their plants' emissions using these and additional allowances they may buy via market transactions, according to a vintage specific schedule laid out by the program.[7] Given this institutional structure, the question is whether the cap-and-trade rule constitutes a significant regulatory cost for firms with affected plants. We demonstrate in a number of ways that this is likely the case for firms that are financially constrained.

Table 1 presents publicly available aggregate data on quarterly allowance auctions (Panel A), free allocations (Panel B), and market transactions (Panel C) made available by the California Air Resources Board. Panel A shows that in every quarterly auction starting in 11/2012 for 2013 vintage allowances, current vintage allowances are completely sold out, there are more bids than available current vintages, and the settlement price for current vintages is always higher than the initial reserve price despite the reserve price being increased every year. Furthermore, Panel B indicates that the free allowance allocations leave substantial room for further incentives to bid in auctions or purchase at market prices. For example, in 2014, the average plant receives free

[6] In 2014, the California cap-and-trade program was linked with the cap-and-trade program in Quebec, Canada. As of 2015, total aggregate emissions covered by the rule in California (Quebec) was approximately 400 (60) million metric tons. In 2015, the program was extended to fuel distributors emitting more than 25,000 metric tons.

[7] For example, calendar year 2013 emissions are required to be paid off to at least a third by November 2014 and in full by November 2015. Calendar year 2014 emissions must be paid in full by November 2015. Calendar year 2015 (2016) emissions are to be paid in at least a third by November 2016 (2017) and in full by November 2018. Calendar year 2017 emissions are to be paid in full by November 2018, and the schedule goes on. No calendar year's emissions can be paid with future vintage allowances (i.e. firms cannot borrow from the future to pay today, though they can buy future vintages in advance).

allowances to emit 349 thousand metric tons of greenhouse gas. However, the median plant in our sample emits 690 thousand metric tons (see Panel B of Table 3), which means the plant must acquire the rights to emit the difference of 341 thousand metric tons either by bidding in auctions or buying them from other market participants. Assuming an average price of $12 per metric ton, the cost of doing so amounts to $4.1 million. Since the median firm owns three plants, the total costs to the firm from extra emissions beyond its free allowances amounts to $6.2 million, conservatively assuming half of its plants are located in California.[8] This is a non-trivial cost, which is in the order of 15 basis points of the firm's total assets, 3% of its net profits, and comparable to the interest obligations on its short-term debt assuming an interest rate of 2%. Finally, Panel C of Table 1 shows that the aggregate magnitudes of market transactions are comparable to those of the free allocations or auctions, and that the transaction prices not only hover above the contemporaneous auction settlement prices but also steadily increase over time. Figure 3, which plots the time series of emission allowance futures prices for each vintage, corroborates the evidence on price trends of market transaction.

Put together, Table 1 and Figure 3 suggest that the increase in costs of emitting greenhouse gases due to the introduction of the California cap-and-trade rule is substantial and sufficiently high for financial constraints to matter. However, this does not mean that these costs will affect unconstrained firms as well. Given the magnitude of the estimated costs, we conjecture that while it may be large for firms with high incremental financing costs, it may not be important for firms with deep pockets. This motivates our hypotheses for how the California cap-and-trade rule will affect greenhouse gas emissions and plant ownership by firms, and the role of financial constraints as the economic channel. We elaborate on the intuition in the following section.

2.2 Hypothesis Development

Economic theory posits that firms will allocate capital to places where marginal benefits and costs are equalized such that the net expected return on capital is maximized. Regional regulation, such as the state-wide cap-and-

[8] Given that there are 161 plants and 85 firms in the sample (see Table 2), the actual number of Californian plants owned by one firm with a presence in California is close to 2. The total costs to the median firm from extra emissions then amounts to $8.2 million, or 20 basis points of its total assets.

trade system in California, introduces perturbations to the distribution of net returns across regions and thus motivates capital reallocation for profit maximizing firms. Our hypotheses concern the direction and magnitude of this reallocation.

In our context, firms that have a plant presence both in California and in other states are geographically diversified, and thus can reallocate their internal resources when the profile of net expected returns change across their segments due to the increase in regulatory costs from the new cap-and-trade-rule. However, if firms have access to frictionless borrowing, they would accommodate the change without shifting resources across plants since their costs of external capital would be low enough to absorb the additional burden as long as their net expected returns remain sufficiently high. In contrast, one would expect that financially constrained firms that are geographically diversified reallocate resources away from plants that are subject to higher regulatory costs to plants they own elsewhere, as their costs of external capital would be too high to finance costly emissions and the net returns from internally reallocating their resources would be greater.

Figure 4 illustrates this intuition by plotting the revenues and costs from varying quantities of emissions. Suppose an imperfectly competitive market with downward sloping marginal (average) revenues *mr* (*ar*), and costs that depend on the locale of production. Firms that operate a plant in California face marginal (average) costs mc_{ca} (ac_{ca}) and an optimum point **I** with average costs ***a*** and emission quantity ***d***. The net return from the California plant is equal to the size of the green area bordered by **I** and ***a***. Once the California cap-and-trade rule is implemented, the cost functions move upward to mc'_{ca} and ac'_{ca} for quantities above the amount of the free allocations, shifting the optimum to **I'** where average costs are higher at ***b*** and quantity is lower at ***e***. The net return remains positive and large, equal to the size of the blue area bordered by **I'** and ***b***. Since the net return is still sufficiently high, firms with unlimited access to capital will continue to emit despite the higher costs.[9] However, **I'** is an infeasible equilibrium for financially constrained firms because they are unable to

[9] The assumption that the net return from emitting in California after the implementation of the cap-and-trade rule remains large enough is supported by state level GDP growth data. In Table 10, we document that California not only exhibits higher growth compared to other states by a large margin during the period when the cap-and-trade rule is in

produce at an average cost above the dashed horizontal line ("constraint binding cost"), so they reallocate their resources from California to other states where there are other investment opportunities with positive net returns. For example, if the costs from emitting in other states follow cost functions mc_{oth} and ac_{oth}, constrained firms will reallocate from **I** to **I''**. Empirically, these predictions imply that the cap-and-trade rule will push constrained firms to not only reduce emissions from plants in California by more than unconstrained firms (***d*** for constrained firms vs ***d*–*e*** for unconstrained firms), but also increase emissions from plants in other states by more (***f*** for constrained firms vs no increase for unconstrained firms).[10]

In other words, the value of internal reallocation would be greater for financially constrained firms when the costs of emissions are increased due to policy changes. The motivation of this hypothesis is grounded in the literature in finance on the value of internal capital markets in the face of financial frictions (for early studies, see Gertner, Scharfstein, and Stein, 1994; Lamont, 1997; Stein, 1997; Shin and Stulz, 1998). In this literature, it has been shown that the contribution of internal capital markets to firm value and hence the value of corporate diversification is greater when external financial constraints are higher, for example when there are large dislocations in financial markets (see Billett and Mauer, 2003; Matvos and Seru, 2014; Matvos, Seru, and Silva, 2018).

This economic rationale leads to two key research questions with regards to the effect of climate policy on firms that we investigate in this paper: (1) Do local climate policy changes (such as the introduction of the California cap-and-trade rule) affect firms' internal resource allocations and environmental policies such as their greenhouse gas emissions and ownership stakes across plants? (2) Are firms' reallocation responses to policy

---

effect, but also that the acceleration in GDP growth compared to the previous period is also greater in California than in other states.

[10] In Figure 4, the cost curve in other states lie below that of California. If this were not the case and mc_{oth} were identical to mc_{ca}, the figure would still suggest a sharper decrease in California emissions by constrained firms than by unconstrained firms, and a corresponding sharp increase in emissions from other states by constrained firms by the amount of ***d*** instead of ***f***. In this scenario, the central prediction that motivates our main hypothesis remains unchanged, and unconstrained firms would still not reallocate. Figure 4, however, raises the possibility that the overall level of firm emissions could increase as a result of the regulation due to the reallocation by constrained firms. We formally test this hypothesis in Section 5.4.

affected by their financial constraints? In the following sections, we describe the data and construction of our sample, and formulate the empirical methodology that we use to test these hypotheses.

3 Data and Sample

3.1 Data

In October 2009, the EPA published the Greenhouse Gas Reporting Program (GHGRP) mandating that sources that emit 25,000 metric tons or more of carbon dioxide equivalent greenhouse gases per year must report their emissions, compliant with the estimation methodologies prescribed by the EPA. Once the submitted information is verified by the EPA, the data is made publicly available through the Facility Level Information on GHGs Tool (FLIGHT), providing plant level information on the identity, geographic location, parent company ownership, NAICS industry code of the plant, as well as the quantity of greenhouse gas emissions on an annual basis starting in 2010 for large plants that meet the reporting requirements. Our sample period extends from 2010 to 2015 - three years before and after the beginning of the California cap-and-trade program - and the initial sample covers approximately 9,200 unique plants.[11]

To analyze the impact of financial constraints, we hand-match the EPA plant level dataset with annual financial accounting data from Compustat based on the names of parent companies. To be included in our merged sample, we require that firms have positive assets and sales greater than $10 million. While utilities and governmental firms might be significant greenhouse gas emitters, common measures of financial constraints may not matter for them in the same way as they do for typical industrial firms. For this reason, we exclude not only financial firms (SIC 6000–6999), but also utilities (SIC 4900–4999) and governmental firms (SIC 9000–9999). The final sample is an unbalanced panel of 2,806 plants and 511 firms over the sample period 2010 to 2015.

[11] We do not include the years 2016 and 2017, which include potentially confounding events such as the signing of the Paris Agreement and the subsequent withdrawal by the United States, as well as additional legislative packages signed by the state of California seeking to reduce greenhouse gas emissions and other air pollutants.

We collect standard variables from Compustat to be used as controls or to compute financial constraint measures such as total assets, PP&E, capital expenditures, short-term debt, long-term debt, cash, cash flow, profitability, Tobin's Q, dividends, repurchases, long-term (i.e. bond) and short-term (i.e. commercial paper) credit ratings. We take the difference between the observation year and founding year as firm age as in Jovanovic and Rousseau (2001). We also compute R&D stock using the perpetual inventory method, where we initialize R&D capital stock at zero and accumulate R&D expenses with a depreciation rate of 15% (see Hall, Jaffe, and Trajtenberg, 2005). All continuous financial variables are winsorized at the top and bottom 1% before a financial constraint index is computed. A detailed list of variable definitions is included in the Appendix of the paper.

3.2 Measuring Financial Constraints

Since an important part of our study is to establish an economic channel through which financial constraints determine how firms respond to climate policy, measuring financial constraints is a critical step in our paper. Based on the financial accounting information from Compustat, we employ six alternative measures of financial constraints commonly used in the literature. They are the Kaplan-Zingales index (see Kaplan and Zingales, 1997; Lamont, Polk, and Saá-Requejo, 2001), the Hadlock and Pierce (2010) index, the Whited and Wu (2006) index, firm size, payout, and credit (i.e., bond or commercial paper) ratings (see Almeida, Campello, and Weisbach; 2004). In addition, we combine the six variables into a composite indicator that is our primary measure of financial constraints.

For the Kaplan-Zingales, Hadlock-Pierce, and Whited-Wu indices, as well as firm size and payout, firms are assigned percentile rankings based on each measure every year. We then use the six years strictly before our sample period (i.e. fiscal years 2003–2008) to compute time-series average percentile rankings for each firm and each measure. Based on these average rankings, firms are categorized as financially constrained if they are above the median for the Kaplan-Zingales, Hadlock-Pierce, and Whited-Wu indices, and if they are below the median for firm size and payout.

For credit ratings, we first examine long-term bond ratings and short-term commercial paper ratings separately. If a firm did not have a bond (commercial paper) rating as of the most recent year of the 2003–2008 pre-sample period but had on average positive long-term (short-term) debt during this period, the firm is categorized as "long-term (short-term)" financially constrained. If the firm did have a bond (commercial paper) rating as of the most recent year of the six-year pre-sample period or had on average zero long-term (short-term) debt during this period, then the firm is "long-term (short-term)" unconstrained. If a firm is either long-term or short-term credit constrained, the firm is classified as constrained based on ratings and unconstrained otherwise.

Finally, we construct a composite measure of all six proxies of financial constraints. For the composite indicator, a firm is categorized as constrained if the majority of the six proxies classify the firm as being constrained; otherwise the firm is unconstrained. For all of our financial constraint measures, firms are classified strictly before they enter the sample period, thus ruling out reverse causality concerns or omitted variables simultaneously affecting the evolution of constraints and firm responses to policy.

3.3 Sample Statistics

Table 2 illustrates the geographical distribution of our sample of plants and firms owning these plants. The intersection of the EPA and Compustat universes covers firms and plants in virtually all states. Over the sample period, the average annual emissions per plant is approximately 289 thousand metric tons, implying the average annual aggregate amount to be 810 million metric tons. According to the EPA, the average amount of greenhouse gas emissions from the US industrial sector over this period was 1,430 million metric tons. Hence, approximately 57% of all industrial greenhouse gas emissions can be attributed to plants owned by our sample of public firms.

The focal state of our study, California, ranks third among all states in terms of the number of firms with a plant presence (i.e., 85 firms, or 17% of all firms, of which 70 also own a plant in other states), fourth in terms of the number greenhouse gas emitting plants (i.e., 161 plants), and seventh in terms of average annual emissions per plant (i.e., 398 thousand metric tons). In short, California is a significant source of greenhouse gas emissions and takes up a sizable portion of the plants and firms in our sample, despite its dominance in the high-tech industry. Understandably, the most important alternative state in the sample is Texas, a state whose weight in the US fossil fuel industry is disproportionately high. Approximately 14% of our sample firms (i.e., 70 out of 511) and 82% of firms with a plant in California (i.e., 70 out of 85) are geographically diversified in the sense they have a presence both in California and in other states. This final observation motivates our hypothesis that a policy curbing emissions in California alone could very well have spillover effects to other states that do not have such a comprehensive program in place.

Panel A of Table 3 describes the characteristics of the sample firms. The size of firms and amount of greenhouse gas they emit are both positively skewed, consistent with the fact that a smaller number of large firms own more plants that generate emissions. The average (median) firm emits 1,584 (277) thousand metric tons, and has total assets of $17.3 ($4.2) billion. R&D spending and R&D capital stock are both skewed to the right as well, as many industrial firms have little R&D. The firms in our sample tend to hold a significant portion of their assets in the form of property, plant, and equipment (47% of assets), have more long-term borrowing than short-term (28% versus 3% of assets), pay a substantial portion of their earnings out either as dividends or via repurchases (49% of earnings before extraordinary items), and on average are less than 30 years old.

Our sample is well balanced in terms of the composition of financially constrained and unconstrained firms, based on whether firms have long-term (> 1 year) or short-term (< 1 year) credit ratings. Approximately 45% of sample firms do not have a long-term bond rating and more than 75% do not have a short-term commercial paper rating, roughly consistent with Almeida, Campello, and Weisbach (2004). This bolsters our confidence in adopting the method of categorizing firms as financially constrained or unconstrained based on their cross-sectional rankings of constraint measures, which is commonplace in the literature but often subject to criticism that objectively unconstrained firms can be misclassified as relatively constrained.

Panel B of Table 3 shows the distribution of plant emissions and ownership. Plant emissions are skewed to the left: While the average plant emits 289 thousand metric tons of greenhouse gas, more than half of the plants emit substantially more given a median of 690 thousand metric tons per year. For almost all plants, ownership is concentrated in one firm. In other words, there are rarely cases where multiple firms share and operate the same plant. The average (median) firm owns six (three) plants; the positive skewness is consistent with the distribution of firm assets and emissions shown in Panel A.

Next, we formulate our empirical strategy to test the hypothesis that the California cap-and-trade rule differentially incentivizes financially constrained firms to reallocate emissions and reconfigure plant ownership using the EPA-Compustat merged sample dataset.

4 Empirical Methodology: Difference-in-Differences

We exploit the variation in treatment of the California cap-and-trade rule in the cross-section (i.e., plants in California versus other states) and time-series (i.e., before and after 2013) to implement difference-in-differences (DID) regressions at the firm-plant-year level. We first begin by comparing the emissions of plants in and outside of California without regards to heterogeneity in financial constraints, using the following regression specification:

$$Log(1+Emissions_{i,j,t}) = \alpha + \beta CalPlant_j \times After_t + \gamma' X_{i,t} + a_j + b_t + \varepsilon_{i,j,t} \tag{1}$$

$Log(1+Emissions_{i,j,t})$ is the logarithm of metric tons of greenhouse gases emitted by firm i at plant j. $CalPlant_j$ is an indicator variable equal to 1 if plant j is located in California and 0 otherwise. $After_t$ is an indicator equal to 1 if the year is 2013 or after and 0 otherwise. $X_{i,t}$ denotes a vector of firm level control variables such as PP&E and R&D stock. Finally, a_j and b_t each denote plant fixed effects and year fixed effects, respectively. The variables $CalPlant_j$ and $After_t$ are not included by themselves in the regressions as they are subsumed by the fixed effects. In the Appendix, we further report results controlling for firm-by-plant fixed effects as robustness checks.

This model evaluates the impact of the cap-and-trade scheme on the emissions by firms from plants they own in California as compared to plants located in other states. If the trends in emissions for treated plants (i.e., located in California) and non-treated plants (i.e., not located in California) are parallel prior to the implementation of the California cap-and-trade, the specification will plausibly isolate the effect of the rule itself, insofar as there are no confounding events that occur coincidentally with the introduction of the cap-and-trade rule.[12] The first two charts in Panel A of Figure 5 plot the time trends of firm-plant level emissions (in thousand metric tons). The first plot includes all firms and the second plot includes only firms that are geographically diversified (i.e., have plants both in California and in other states). The graphs show that emissions from California and non-California plants are closely aligned prior to treatment such that the trends are parallel. However, unconditionally there is also no visible subsequent divergence after the rule is implemented.

This picture changes dramatically when the sample of geographically diversified firms is split into financially constrained and unconstrained firms. In Panel A of Figure 6, the first two charts again show no apparent change from the parallel trends regardless of whether firms are constrained or not when the sample split is applied to all firms. The next two charts focus on plant emissions by geographically diversified firms that are relevant for our hypothesis on internal reallocation. For financially unconstrained firms, emissions from California and non-California plants move in parallel before the implementation of the cap-and-trade rule and largely maintain this pattern after 2013 as well. In sharp contrast, for constrained firms, the parallel trends before 2013 begin to diverge visibly afterwards. Post 2013, California plants owned by constrained firms reverse their prior upward trend and start reducing emissions, whereas non-California plants sharply increase emissions. These trends paint a clear picture that the California cap-and-trade rule has a large differential impact on how financially constrained and unconstrained firms allocate emissions across their plants located in California and in other states. To formally test whether these patterns are statistically significant, we estimate the difference-in-differences model in Equation (1) separately for constrained and unconstrained firms, and evaluate whether the coefficients on the interaction term $CalPlant_j \times After_t$ are significantly different in the two models.

[12] During our sample period from 2010 to 2015, the 2013 cap-and-trade rule was the only notable climate policy introduced to curb industrial greenhouse gas emissions. It was the first major regulation enforced to achieve the emission reduction objectives initially outlined and required by the landmark California state law AB 32. After 2015, AB 32 was strengthened by several subsequent legislative bills (e.g. SB 32 and AB 197 in 2016; AB 398 and AB 617 in 2017). Aside from AB 32, the governor of California signed SBX1 2 in 2011, requiring that one third of the state's electricity come from renewable sources by 2020, and in 2014, the energy efficiency requirements for newly constructed buildings were tightened pursuant to updated Green Building Standards. However, these policies are distinct from the cap-and-trade rule in their enforcement targets, intensity, and timing. Hence, the emission shifting between industrial plants that we identify around 2013 primarily correspond to the impact of the introduction of the cap-and-trade rule.

It is worth noting that anticipation about the cap-and-trade rule prior to its implementation is unlikely to be an issue in our setup, as there is no economic benefit to the firm from preemptively reallocating their emissions when profits from emitting in California are still high before the introduction of a regulatory cost. The absence of such anticipatory adjustments is also empirically evident in the parallel trends.

While Equation (1) provides for a test of whether the California cap-and-trade rule has a significant effect, the coefficients on the interaction term only capture the change in emissions from California plants relative to non-California plants. In other words, it does not disentangle whether the effect is coming from changes in the plants of interest located in California or from spillover effects to plants in other states. To tell these two effects apart, we run an additional DID regression, where we replace the plant level treatment dummy $CalPlant_j$ with a firm level dummy $DivFirm_{it}$ that is an indicator for whether a firm owns plants both in California and in other states during a given year or not. Since $DivFirm_{it}$ is not subsumed by fixed effects, it is also included as a regressor by itself.[13] We run this firm-plant-year level regression on a subsample of non-California plants. This specification effectively tests for spillover effects by studying whether plants outside of California exhibit changes after the cap-and-trade rule is implemented depending on whether the parent companies' assets are affected by the rule. This specification is written as follows:

$$Log(1+Emissions_{i,j,t}) = \alpha + \beta_1 DivFirm_{i,t} + \beta_2 DivFirm_{i,t} \times After_t + \gamma' X_{i,t} + a_j + b_t + \varepsilon_{i,j,t} \quad (2)$$

The third chart in Panel A of Figure 5 plots the emission trends of non-California plants owned by firms with and without a California presence. In this unconditional plot, the parallel trends assumption holds, but there are no visible changes in the post-trends either. In the plots that condition the sample based on financial constraints, shown in the third row of Panel A of Figure 6, the pre-trends are similarly parallel for both unconstrained and constrained firms. However, during the post 2013 period, constrained firms with California plants substantially increase emissions from their non-California plants, whereas there are no changes

[13] In robustness checks reported in the Appendix, where we control for firm-by-plant fixed effects, the term $DivFirm_{it}$ is dropped.

for plants owned by constrained firms without exposure to California or unconstrained firms regardless of their California exposure. Altogether, the illustrative evidence from Figures 5 and 6 validates the DID setup and is consistent with a strong spillover effect from financially constrained firms with assets exposed to the California cap-and-trade rule shifting their emissions to other states. Estimating Equation (2) separately for constrained and unconstrained firms and comparing the coefficients formally tests this spillover effect.

As an alternative to comparing coefficients from separate DID regressions on constrained and unconstrained subsamples, we run pooled regressions by including a *Constrained$_i$* dummy in an expanded triple difference framework. The triple difference specifications can be written as follows:

$$\begin{aligned} Log(1+Emissions_{i,j,t}) = \alpha &+ \beta_1 Constrained_i + \beta_2 Constrained_i \times After_t \\ &+ \beta_3 CalPlant_j \times Constrained_i + \beta_4 CalPlant_j \times After_t \\ &+ \beta_5 CalPlant_j \times Constrained_i \times After_t + \gamma' X_{i,t} + a_j + b_t + \varepsilon_{i,j,t} \end{aligned} \tag{3}$$

and

$$\begin{aligned} Log(1+Emissions_{i,j,t}) = \alpha &+ \beta_1 Constrained_i + \beta_2 DivFirm_{i,t} + \beta_3 Constrained_i \times After_t \\ &+ \beta_4 DivFirm_{i,t} \times Constrained_i + \beta_5 DivFirm_{i,t} \times After_t \\ &+ \beta_6 DivFirm_{i,t} \times Constrained_i \times After_t + \gamma' X_{i,t} + a_j + b_t + \varepsilon_{i,j,t} \end{aligned} \tag{4}$$

This method overcomes issues related to model fit or misspecification that may be compounded by comparing coefficients across multiple models, and enables the econometrician to control for differences across other coefficients in the model as well. We use both methods, separate and pooled regressions, for the analyses on emissions and focus on the pooled regression method in subsequent plant ownership tests or placebo tests.

In the following section, we present and discuss the results based on our data and empirical design. Motivated by the data patterns observed in Figures 5 and 6, we henceforth drop the analysis of all firms and focus our study on contrasting emission and plant ownership responses of California versus non-California plants for the sample of geographically diversified firms, and comparing the responses of non-California plants owned by geographically diversified versus undiversified firms. Hence, we concentrate our discussions on the internal reallocations by geographically diversified firms and the resulting spillovers to other states.

5 Results

In this section, we present empirical evidence pursuant to our economic hypotheses, DID methodology, and data outlined in the previous sections. We begin by presenting unconditional evidence on firms' emission responses to the introduction of the California cap-and-trade rule, and then move on to analyzing the impact of variations in financial constraints.

5.1 Unconditional Tests

In Table 4, we report results from unconditional tests without exploiting the heterogeneity in financial constraints across firms. These results help us understand the overall effects of the California cap-and-trade rule, and provide further motivation to explore the financial constraints channel through which they manifest.

We start with univariate results in Panel A. For each firm-plant, we first compute emissions growth as the difference between post–2013 and pre–2013 average emissions, scaled by the full sample period mean. We then divide our sample of plants into treatment and control groups, and compare the mean emissions growth between the two groups by reporting the difference of their means and its corresponding *t*-statistic. Treatment is defined under two alternative schemes, each corresponding to a subpanel in Panel A. In the first scheme, we define treatment at the plant level based on their location, where we define California plants as the treatment group and plants in other states as the control group. We focus on the subsample of geographically diversified firms that have plants both in California and in other states. The unconditional difference of emissions growth between California and non-California plants is –6%. While this raw difference is not statistically significant (*t*-statistic of –1.03), it is weakly significant in multivariate analyses that control for a host of control variables and fixed effects, as discussed below. In the second scheme, we focus on plants outside of California and define treatment at the firm level based on whether firms own plants in California, i.e., whether they have asset exposure to the cap-and-trade rule or not. The unconditional difference of the change in emissions between non-California plants owned by diversified and non-diversified firms is positive (5% pts) and statistically significant (*t*-statistic of 2.2). Overall, the univariate results indicate a modest decline in emissions from California plants and a sharp increase in emissions from plants located in other states as a result of the implementation of the California cap-and-trade rule.

In Panel B of Table 4, we report multivariate results from running the regressions in Equations (1) and (2), where we control for other observable characteristics and several dimensions of unobserved heterogeneities such as plant and year fixed effects. The first four columns (1)-(4) show results from running Equation (1) on the subsample of firms that are geographically diversified, where the dependent variable is the logarithm of emissions (*Log(1+Emissions)*) and treatment is whether a plant is located in California (*CalPlant*). Four variations of this specification are implemented: (i) without fixed effects or controls, (ii) with year fixed effects, (iii) with plant and year fixed effects, and (iv) with both fixed effects as well as controls (henceforth the main specification).[14] As controls, we include firm level PP&E and R&D stock, both scaled by total assets. The adjusted R^2 jumps from below 0.01 to over 0.85 as plant fixed effects are included, highlighting the fact that idiosyncratic differences across plants, such as industry classification, maximum capacity, or technological obsoleteness, explain an important portion of the variation in greenhouse gas emissions.

The key coefficient of interest is on the interaction term *CalPlant* × *After*, which captures the differential treatment effect of the introduction of the cap-and-trade rule on emissions. The sign on this coefficient is consistently negative across all four specifications, and the magnitude is also fairly similar even though the plant fixed effects subsume a small portion of its explanatory power. The coefficient on the interaction term is negative (–0.155) and significant at the 10% level controlling for both plant and year fixed effects as well as firm level controls. In terms of economic magnitude, the result indicates that firms reduce emissions from California plants by 15.5% more than from non-California plants.[15]

The next four columns (5)-(8) in Panel B examine whether part of this treatment effect can be explained by reallocations or spillovers to plants outside of California by focusing on the sample of plants located in other states. As in Equation (2), the *CalPlant* treatment dummy is replaced by a firm level *DivFirm* indicator for

---

[14] In Table A.1 of the Appendix, we also report results controlling for firm-by-plant and year fixed effects. The results are economically and statistically similar to those in Table 4.

[15] In the full sample, however, the coefficient is less than half in size (–0.064) and statistically insignificant. This implies that the impact of the cap-and-trade rule on California emissions is roughly confined to firms with plants in other states as well, but that the policy is not meaningfully effective in reducing emissions otherwise.

whether the parent firm has a plant presence in California and is thus exposed to the cap-and-trade regulation. The results indicate strong and significant spillover effects. The coefficient on *DivFirm* × *After* is positive and highly significant at the 1% level in all four specifications. Controlling for fixed effects and firm level variables, non-California plants owned by firms exposed to the California cap-and-trade rule increase emissions by 13.1% more than plants owned by non-diversified firms.

Overall, the results in Table 4 strongly suggests unintended consequences of the cap-and-trade rule in the form of spillover effects due to reallocation motives of firms whose assets are "partially" affected by the regulation. Moreover, we confirm in untabulated analysis that any significant reduction in emissions from California plants are mainly driven by the reallocation by geographically diversified firms who correspondingly increase their emissions from plants in other states. Otherwise, there is no meaningful evidence that the cap-and-trade rule effectively reduces emissions. An important economic question that arises from this finding is what frictions motivate firms to shift resources internally across their plants, because in a frictionless world firms could simply raise more capital to absorb the increased costs of emissions as long as operating in California yields sufficiently large and positive net returns. As discussed earlier, we hypothesize that financial constraints constitute an important friction that provides an economic channel for such reallocations and spillover effects.

5.2 The Impact of Financial Constraints

We now implement our DID regression strategy to explore the financial constraints channel. In Table 5, we run our DID regressions with plant and year fixed effects as well as firm controls, separately for financially constrained and unconstrained firms.[16] We use our composite indicator as well as six proxies (i.e. Kaplan-Zingales, Hadlock-Pierce, Whited-Wu, firm size, payout, and credit rating) for financial constraints, and report results for each of these measures. To statistically compare the effects across constrained and unconstrained groups, we test for differences in the coefficients on the *CalPlant* × *After* or *DivFirm* × *After* interaction terms.

[16] In Table A.2 of the Appendix, we also report results controlling for firm-by-plant and year fixed effects. The results are similar to Table 5, both economically and statistically.

In Panel A, we take the sample of geographically diversified firms that operate plants both in and outside of California and run regressions based on Equation (1), where we compare the changes in emissions from California plants with non-California plants using the interaction term *CalPlant* × *After*. These regressions show that constrained firms reduce their emissions from California plants more compared to plants in other states, whereas unconstrained firms do not. According to our composite constraint measure, constrained firms reduce emissions from California plants by 35% more (significant at 10%) compared to non-California plants, whereas this effect is a statistically insignificant 3% for unconstrained firms. The difference between the responses by constrained and unconstrained firms is statistically significant at 10% with a *p*-value of 0.07. This result is economically robust across all six constraint proxies, and statistically robust for the Hadlock-Pierce index and payout ratio in particular.

In Panel B, we estimate Equation (2) with *DivFirm* as the treatment dummy, where we compare emissions between plants outside California owned by firms with and without a California presence. We run this regression again separately for the sample of constrained and unconstrained firms, and formally compare the coefficients on *DivFirm* × *After* across the two models. The results are consistent with a strong spillover effect where financially constrained firms significantly increase their emissions from plants outside California if part of their assets are exposed to the increased regulatory burden of the California cap-and-trade rule. Under our composite measure, constrained firms with plants in California increase emissions by 29% more (significant at 1%) than those without plants in California. For unconstrained firms, the relative change in emissions is only –9% (only weakly significant at 10%). The difference between the responses by constrained and unconstrained firms is highly significant with a *p*-value close to zero. This finding is economically and statistically robust across all measures of financial constraints except for the Whited-Wu index, for which the result is still economically consistent. The differences of the effects between constrained and unconstrained firms are significant at 1% for the Kaplan-Zingales index, Hadlock-Pierce index, payout, and rating, and significant at 5% for firm size.

As discussed in the previous section, we also use alternative specifications where the sample of constrained and unconstrained firms are pooled together and a $Constrained_i$ dummy is included in a triple difference regression, instead of running separate regressions and comparing coefficients from the two models (see Equations (3) and (4)). This helps avoid biases of comparing coefficients from distinct regression models with varying degrees of model fit and misspecification, and it also enables us to control for differences in the other coefficients across the two models. Table 6 reports the results from these pooled triple difference regressions. Similar to the previous table, the results comparing emissions from California and non-California plants owned by geographically diversified firms based on the *CalPlant* indicator are reported in Panel A, and the tests for spillover effects comparing non-California plant emissions by diversified and non-diversified firms using the *DivFirm* indicator are reported in Panel B.[17]

In Panel A of Table 6, the main coefficient of interest is on the triple interaction term *CalPlant* × *After* × *Constrained*, which captures how firms change their emissions from plants in California relative to plants in other states depending on whether they are financially constrained or not. We expect the coefficient on this term to be negative, as constrained firms are expected to reduce emissions in California by more. Also relevant is the coefficient on *CalPlant* × *After*, which in this context measures how unconstrained firms behave. Since there are virtually no responses by unconstrained firms based on the results reported in Table 5, we expect this coefficient not to be significantly different from zero, or at most only weakly negative. The results confirm that this is the case. Panel A shows that for firms with plants both in and outside of California, the coefficient on the triple interaction term is economically large and negative (for all measures) and statistically significant (at 10% for the composite indicator, 5% for Hadlock-Pierce, 10% for size). The magnitude of the coefficient, for example –0.39 in the case of the composite indicator, is also consistent with the size of the difference between constrained and unconstrained firms in Table 5 (coefficients of –0.35 and –0.03, respectively). The coefficient on *CalPlant* × *After*, on the other hand, is small and insignificant for all constraints measures, consistent with our prior.

[17] In Table A.3 of the Appendix, we also report results controlling for firm-by-plant and year fixed effects. The results are economically similar to Table 6 for the results comparing emissions from California and non-California plants owned by geographically diversified firms (Panel A), and both economically and statistically robust for the tests on spillover effects to non-California plants by comparing geographically diversified and non-diversified firms (Panel B).

In Panel B of Table 6, we similarly examine the coefficients on *DivFirm* × *After* × *Constrained* and *DivFirm* × *After*. Drawing from the results in Table 5, we expect the triple interaction term to be positive and significant as constrained firms are more likely to shift their emissions to other states if their assets are exposed to the California cap-and-trade rule. We also expect the double interaction term not to be significantly different from zero as unconstrained firms should not exhibit differential changes in their plants outside of California. Our results strongly confirm these predictions. The coefficient on *DivFirm* × *After* × *Constrained* is positive and large in magnitude (for all measures) and also statistically significant (at 1% for the composite measure, 5% for Kaplan-Zingales, 10% for Hadlock-Pierce, 10% for size, and 1% for rating). In the case of the composite measure, for example, the magnitude of the coefficient, 0.34, closely matches the size of the difference in the coefficients of 0.29 and –0.09 for constrained and unconstrained firms in Table 5, respectively. The coefficient on *DivFirm* × *After* is indistinguishable from zero across all measures except for payout, also largely consistent with our prediction.

In summary, our results provide strong and consistent evidence that (a) firms owning plant operations both in California and in other states reduce emissions from their plants in California relative to plants in other states, (b) that these firms increase emissions from their plants in other states relative to firms with no presence in California, and (c) that these effects are almost exclusively due to their financial constraints and thus their incentives to internally reallocate emissions.

5.3 Plant Ownership Reallocation

A closely related question that arises from our results is whether firms not only shift emissions, but also go as far as to reconfigure the geographical distribution of their plant ownership profiles in response to higher regulatory costs in California due to the cap-and-trade rule. If the present value of all current and expected regulatory costs in the future are sufficiently high, then not only will firms shift their emission activities across plants they already have in place, but they should also be willing to incur the high fixed costs of closing or selling existing plants and opening or acquiring new ones. On the other hand, given the lumpy nature of financing and investment decisions, changes in variable and marginal operating costs due to the cap-and-trade rule may not suffice to induce responses such as large investments or divestments of fixed assets. While Figures 5 and 6 suggest that there are unlikely to be dramatic and large shifts in the level of ownership in fixed assets such as plants, it is theoretically possible that there will be marginal adjustments along this dimension as well. In this section, we test this prediction by using firm-plant level closure or opening decisions as categorical dependent variables, and we also examine whether firm level plant ownership is affected.

Table 7 presents the results from these tests. We employ three distinct methodologies. First, we define two binary variables, *Close* and *Open*, and use them as dependent variables in a linear probability model analogous to the pooled regression models in Equation (3) and (4).[18] *Close* is defined broadly as an indicator equal to 1 if one of the following are true (and 0 otherwise): In a given year, (a) the firm reduces its fractional ownership in a plant it has in place, or (b) a plant that was owned by a firm is no longer owned by that firm. *Open* is conversely defined to be equal to 1 (and 0 otherwise) if: (a) The firm increases its fractional ownership in an existing plant, or (b) a plant that was not owned by a firm is now owned by that firm. We include plant and year fixed effects in the linear probability model estimations, and report marginal effects as results.[19]

While the linear probability model has the advantage of being able to control for high dimensional fixed effects, it has the disadvantage of losing information when the available categories of decisions are more than binary. In our context of analyzing plant ownership decisions, for example, the relevant alternatives to closing an existing plant are twofold: To keep an existing plant, or open a new plant. To fully internalize this set of information in our analysis, we alternatively estimate multinomial logit models using a categorical variable equal to –1 for plant closure, 0 for no change, and +1 for plant opening. We set "no change" as the base case and estimate the probabilities of either a plant closure or opening with respect to that base case. While we keep the model specification analogous to Equations (3) and (4), we drop the fixed effects as logit models have difficulty converging for estimations with a high dimension of regressors. Marginal effects are reported. Finally,

---

[18] We obtain consistent results with regressions according to Equations (1) and (2) run separately for financially constrained and unconstrained firms. These results are untabulated to conserve space.

[19] In Table A.4 of the Appendix, we also report results controlling for firm-by-plant and year fixed effects. The results are economically similar to Table 7 for the results comparing closure/openings of California and non-California plants, and both economically and statistically robust for the tests on spillover effects to the closure/openings of non-California plants by comparing geographically diversified and non-diversified firms.

we also estimate firm level regressions with the number of plants owned by a firm as the dependent variable, controlling for firm and year fixed effects.

The first two columns of Panel A of Table 7 report results from estimating linear probability models of geographically diversified firms' plant closure and opening decisions in California relative to their decisions in other states, analogous to Equation (3). The results show that financially constrained firms are 15% more likely (significant at 1%) to close a plant in California, whereas unconstrained firms are unaffected in their likelihood of adjusting plant ownership. Under the linear probability model, there is no symmetric effect on the firms' decisions to open a plant. According to the multinomial logit estimations for the sample of diversified firms in the next two columns, constrained firms are 18% more likely to close (statistically not significant) and 8% less likely to open (significant at 5%) a plant in California. The results from both specifications can be summarized as firms being relatively more likely to close an existing plant in California rather than open a new one. As shown in the last column, there is however no significant evidence at the firm level that constrained firms own fewer plants in California, though the coefficient on *CalPlant* × *After* × *Constrained* is negative and thus has the predicted sign.

Panel B of Table 7 reports results from analyzing firms' plant closure and opening decisions in other states, as a function of their geographical diversification (i.e. presence in California) and financial constraints, analogous to Equation (4). Both results from linear probability models and multinomial logits are strongly consistent with constrained firms with a presence in California being significantly less likely to close (–6% in both models and significant at 1%) and also significantly more likely open new plants (13% and 18%, both significant at 1%) in states other than California. The magnitude of the coefficient on *DivFirm* × *After* × *Constrained* in the firm level regression is with 0.82 sizable but statistically insignificant, indicative of the average constrained firm having roughly 1 additional plant outside of California after 2013.

Overall, our analysis of plant ownership changes by firms in response to the California cap-and-trade rule suggests that beyond internal reallocations of greenhouse gas emissions, there are also adjustments along the extensive margin where firms show a higher probability of shifting their plant ownership profiles away from

California towards other states. While the levels of plant ownership do not exhibit much variation over time (see Figures 5 and 6), we document significant changes in the probabilities of marginal adjustments to the ownership of plants. Nonetheless, we take the results above as weaker evidence on plant ownership reallocation compared to our findings on emission reallocation across plants in place, consistent with the discrete and lumpy nature of financing and investment activities.

5.4 Firm-wide Total Emissions

A critical policy implication from the reallocation results thus far is that the California cap-and-trade rule may not lead to the desired reduction in global greenhouse gas emissions. To the contrary, it might result in an increase in emissions rather than a reduction, undermining the goal of climate policy to combat global warming as a consequence of climate change. For example, if the costs of emissions are lower in other states than in California as illustrated in Figure 4, the predicted reallocation may result in an overall increase in emissions. In this section, we test this possibility by examining firm level total emissions. Specifically, we aggregate plant emissions within firms and compare the change in total emissions before and after the implementation of the cap-and-trade rule between firms with and without assets affected by the rule (i.e. California plants). The results are reported in Table 8, where we run firm level regressions as follows:

$$Log(1 + Firm\,Total\,Emissions_{i,t}) = \alpha + \beta_1 After_t + \beta_2 Constrained_i \times After_t + \gamma' X_{i,t} + c_i + \varepsilon_{i,t}\,. \tag{5}$$

Log(1+Firm Total Emissions$_{i,t}$) is the logarithm of metric tons of greenhouse gases emitted by firm i in year t. $After_t$ is an indicator equal to 1 if the year is 2013 or later, and 0 otherwise. To test whether financially constrained and unconstrained firms increase or reduce emissions differently, we also include $Constrained_i$, a firm level dummy variable equal to 1 if firm i is financially constrained based on our composite measure and 0 otherwise, and its interaction with $After_t$. $X_{i,t}$ denotes a vector of firm level control variables such as PP&E and R&D stock. Finally, c_i denotes firm fixed effects. While we are interested in the coefficients for both $After_t$ and $After_t \times Constrained_i$ to infer increases or reductions in emissions, we also alter the specification to include year fixed effects and drop $After_t$ to ensure robustness of the rest of the coefficients.

We estimate this regression separately for a control group of firms that do not have plants in California, a treatment group of firms that do have plants in California, and a subsample of treatment firms that have plants both in California and in other states as well. We also run pooled regressions with a $Treat_{i,t}$ indicator for whether firm *i* is in the treatment group (or alternatively in the subsample treatment group) as below:

$$\begin{aligned} Log(1+Firm\,Total\,Emissions_{i,t}) = {} & \alpha + \beta_1 After_t + \beta_2 Treat_{i,t} + \beta_3 After_t \times Constrained_i \\ & + \beta_4 Treat_{i,t} \times After_t + \beta_5 Treat_{i,t} \times Constrained_i \\ & + \beta_6 Treat_{i,t} \times After_t \times Constrained_i + \gamma' X_{i,t} + c_i + \varepsilon_{i,t} \end{aligned} \quad . \tag{6}$$

Columns (1) and (2) of Table 8 show that control firms without any plants in California do not significantly change their overall emissions around the time the California cap-and-trade rule was implemented (i.e. coefficient on *After* of 0.08, not statistically significant), regardless of whether they are financially constrained or not (i.e. coefficients on *After* × *Constrained* of 0.12 and 0.14, not statistically significant). Columns (3) and (4) show that this is not different for treatment firms with plants in California (i.e., coefficients on *After* and *After* × *Constrained* of –0.06 and 0.11, respectively, all not statistically significant).

These results are also confirmed in regressions (5) and (6) where the control and treatment firms are pooled together. The coefficients on *After* (i.e., 0.08) and *Treat* × *After* (i.e., –0.17 and –0.15) are consistent with those on *After* in the separate regressions for the control and treatment samples, and the coefficients on *After* × *Constrained* (i.e., 0.15 and 0.17) and *Treat* × *After* × *Constrained* (i.e., 0.04 and 0.05) are also in line with those on *After* × *Constrained* in the separate estimations. In short, the results largely fail to show an overall reduction in firm level emissions.

In fact, the last four columns (7)-(10) of Table 8 provide suggestive evidence that a subset of treatment firms with plants both in and outside of California *increase* their total emissions. In column (7), the coefficient on *After* × *Constrained* is 0.26 and significant at 5%, whereas the coefficient on *After* is –0.07 and statistically insignificant. This implies that financially constrained firms significantly increase their firm-wide emissions by approximately 19% after the implementation of the cap-and-trade rule. This contrasts with the insignificant changes for both constrained and unconstrained firms in the control group. The difference between the treatment and control groups are also captured by the pooled regressions in the last two columns (9) and (10), albeit with weak statistical significance. For example, the magnitudes of the coefficients on *After* (i.e., 0.08) and *Treat* × *After* (i.e., –0.18) are consistent with the *After* coefficients in the separate regressions, and the coefficients on *After* × *Constrained* (i.e., 0.15 and 0.16) and *Treat* × *After* × *Constrained* (i.e., 0.14 and 0.15) are also consistent with those on *After* × *Constrained* in the separate estimations.

In short, we find no evidence that firms reduce their overall greenhouse gas emissions as a result of the introduction of the California cap-and-trade rule. To the contrary, there is weak evidence that a subset of financially constrained firms with plants both in California and in other states increase their total emissions, consistent with spillover effects from the cap-and-trade rule resulting in outcomes contradictory to climate policy objectives.

5.5 Impact on Sectoral GDP and Employment

We have thus far documented spillover effects from the California cap-and-trade rule with respect to plant level greenhouse gas emissions and ownership driven by firm financial constraints, and we have shown its impact on firm-wide total emissions. How is this related to broad economic outcomes such as economic activity and employment? This is an important question for economists and policy makers who are interested in the macroeconomic impact of such climate policies. To provide insight into this issue, we examine real GDP and employment data downloaded from the Bureau of Economic Analysis (BEA). We conduct state-sector level analyses to obtain sharper inference than at the state-level that can be confounded by differential changes over time across states. Specifically, we draw from our emission and ownership reallocation results thus far and hypothesize that the California cap-and-trade rule may differentially lower economic activity and employment in affected industries in California compared to other states. We also conjecture that this relative economic contraction from the "polluting" industry may be compensated for by growth from other industries that take its place.

We first define a plant's industry as the narrowest NAICS code with at least 50 plants in the entire cross-section each year, and map this to the narrowest available 2-4 digit NAICS industry classification for which the BEA publicly reports state level GDP and employment. We then collapse the data to state-sector-year level where we broadly categorize sectors as either "emission sector" or "non-emission sector". All BEA industries with greenhouse gas emitting plants are pooled together to comprise the emission sector, and all remaining industries are grouped as the non-emission sector. We then aggregate GDP (inflation adjusted with respect to 2009 dollars) and employment (total number of full- and part-time wage earning workers) up to each state-sector-year, and run the following regressions:

$$Y_{s,t} = Cal_s \times After_t + a_s + b_t + \varepsilon_{s,t} \tag{7}$$

$$\begin{aligned} Y_{s,k,t} = {} & Cal_s \times After_t + Cal_s \times After_t \times EmissionsSector_k + Cal_s \times EmissionSector_k \\ & + After_t \times EmissionSector_k + EmissionSector_k + a_s + b_t + \varepsilon_{s,t} \end{aligned} \tag{8}$$

where Equation (7) is a state-year level regression run for the emission sector and non-emission sector separately. $Y_{s,t}$ is either log(1+GDP) or log(1+Employment), Cal_s is a state level dummy indicating whether the state is California or not, and $After_t$ is an indicator for whether the year is 2013 or later. Equation (8) is a pooled state-sector-year level regression where we combine the two sector samples and include a sector dummy, $EmissionSector_k$. In both regressions we control for state fixed effects, a_s, and year fixed effects, b_t.

Table 9 reports the regression results. The first three columns show weak evidence on GDP growth. Column (1) shows that there is a marginal and statistically insignificant reduction in the economic output from the sector of industries impacted by the California cap-and-trade rule. On the other hand, column (2) shows that GDP in the non-emission sector increases significantly by 7% (significant at 5%). However, the pooled regression in the third column suggests that the difference between the emission and non-emission sectors is not statistically significant. While the coefficient on *Cal* × *After* is positive (i.e., 0.04) and *Cal* × *After* × *EmissionSector* is negative (-0.03), consistent with the signs from the separate regressions, neither of these is significant.

The next three columns document a sizable impact of the California cap-and-trade rule on sectoral employment. The negative coefficient on *Cal* × *After* in column (4) implies a 12% greater reduction in employment (significant at 10%) in the emission sector in California compared to other states. In sharp contrast, column (5) shows that employment increases by 10% more in the non-emission sector in California compared to other states. Column (6) confirms the statistical significance of the difference between the emission and non-emission sectors in a pooled regression. The coefficient on $Cal \times After$ of 0.09 and $Cal \times After \times EmissionSector$ of -0.20 are both statistically significant at 1% and 5%, respectively.

Overall, the results suggest that there is a macroeconomic tradeoff effect from the California cap-and-trade rule. Industries impacted by the regulation in California exhibit decreases in GDP and employment relative to other states, consistent with lower plant emissions and higher likelihood of plant closures in California documented in the previous sections. At the same time, there is a countervailing relative growth in GDP and employment from the non-emission sector comprised of "clean" industries. However, we are agnostic about the eventual welfare implications of these results and caution the reader that these macroeconomic outcomes should also be interpreted as relative reallocations not only across industries but also across regulatory jurisdictions.

5.6 Robustness Checks

5.6.1 Are Firms Reallocating to Chase Better Growth Opportunities?

One concern that could be raised is that our evidence on cross-state reallocation of emissions and plant ownership might be driven not by the higher regulatory costs of operating greenhouse gas emitting plants in California due to the introduction of the cap-and-trade rule, but simply by better growth prospects associated with some plants and not others. For example, if investment opportunities in California were waning whereas the economies of other states are growing faster, it would make sense for firms with limited access to external capital to shift their productive resources toward these more promising states. Our findings would be consistent with this alternative "opportunity chasing" story. To evaluate this argument, we construct and characterize measures of growth opportunities and evaluate the robustness of our results controlling for them.

The first measure is state level annual real GDP growth from private industries in the state the plant is located in, using GDP data from the BEA. While GDP growth captures the overall economic activity and growth within the plant's local economy at the state level, it reflects realized values rather than expectations and becomes a noisier estimate of expected growth opportunities at more granular levels (e.g., state-industry). Therefore, we construct a second forward looking measure as the median Tobin's Q of firms that own plants in the same state and industry as the plant and primarily operate in that industry, where industry is defined as the narrowest NAICS code with at least 50 plants each year. This market-based measure provides the added benefit of summarizing growth opportunities at the state-industry level. Panel A of Table 10 reports the population-weighted cross-state averages of these two measures separately for California and other states, each year over our sample period from 2010 to 2015.

According to GDP growth, California outperformed other states by a large margin in terms of economic growth during the post California cap-and-trade rule period of 2013 to 2015. The average annual growth rate of California over this period was 4.1%, the fourth highest of all U.S. states. In the period before the cap-and-trade rule from 2010 to 2012, by contrast, California's average growth rate was 2.1%, ranking below the twentieth fastest growing state. In other words, California was not only among the fastest growing states during the period after the introduction of its carbon trading scheme, but also among the states whose growth rates vastly improved compared to the period before the regulation (i.e., a significant increase of 2% points, in contrast to no significant increase in other states). According to median Tobin's Q, which better captures market assessments of the growth prospects specific to the plants' locale and industry, growth opportunities in California and other states were not very different before (i.e., 1.32 vs 1.36) or after (i.e., 1.38 vs 1.40) the introduction of the California cap-and-trade rule. At the minimum, there is no evidence that investment opportunities were better in other states compared to California during the latter half of the sample period.

For the opportunity chasing story to hold one must assume that firms responded to the lower realized growth rates during 2010 to 2012 and made reallocation decisions for the following years from 2013 to 2015 whilst oblivious to the improving prospects in California and ignoring the contemporaneous changes in cross-state economic growth during this period. One must also assume that Tobin's Q was uninformative for firms in gauging growth opportunities across their plants. These seem like unreasonable assumptions. Rather, the illustrative evidence largely goes against the alternative explanation that firms reallocated resources simply to capture better growth opportunities in other states, but is more consistent with constrained firms having reallocated *despite* higher expected growth in California due to their lack of financial flexibility to exploit such opportunities and increased regulatory costs. The growth trends also support our assumption that the net returns from emitting in California remain large enough such that unconstrained firms would have little incentive to shift emissions as constrained firms do.

In Panels B and C of Table 10, we employ regressions augmented from Equations (3) and (4) to examine how growth opportunities explain plant emissions and ownership, and whether our results are sensitive to controlling for them. For both GDP growth and Tobin's Q, we include the measure itself as well as its interaction with a *Constrained* dummy (based on our composite constraint measure) to allow constrained and unconstrained firms to respond to growth prospects differentially. As both of these measures vary across states and over time, we do not interact them with *CalPlant* or *After* (but do interact them with *DivFirm* in Panel C as detailed below).

Panel B compares emissions and ownership for California and non-California plants based on the sample of geographically diversified firms. In columns (1)-(3), regressions of plant emissions suggest that neither GDP growth nor Tobin's Q significantly affects emissions regardless of whether firms are constrained or not, and that our results concerning the effects of the cap-and-trade rule on emissions reported in Table 6 are robust to controlling for both growth measures as well as their interactions with *Constrained*. The coefficient on the triple interaction term *CalPlant* × *After* × *Constrained* is –0.37 and significant at 10%, comparable to –0.39 in Table 6. In columns (4)-(9), multinomial logit regressions of plant ownership changes show that unconstrained firms are less likely to close plants associated with better growth prospects (e.g., coefficient of –0.18 on *Q*, significant at 1%), while this is not the case for constrained firms (e.g., coefficient of 0.17 on *Q* × *Constrained*, significant at 1%), and that controlling for these factors does not affect our results on plant ownership documented in Table 7. The marginal probabilities of plant closures and openings in California by constrained firms after the cap-and-trade rule are 0.17 (not significant) and –0.07 (significant at 5%), respectively, comparable to 0.18 and –0.08 in Table 7.

Panel C studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. To allow geographically diversified firms to respond to growth prospects differentially, we further interact GDP growth and Tobin's Q with *DivFirm* × *Constrained* and *DivFirm*. The effects of growth opportunities documented in Panel B are similarly observed, except that the coefficient on *Q* × *Constrained* is positive and significant (i.e., 0.49, significant at 1%). For both GDP growth and Tobin's Q, their interactions with *DivFirm* load positively, while their interactions with both *DivFirm* and *Constrained* load negatively. Controlling for all of the growth opportunity variables and their respective interaction terms, the spillover effects documented in Tables 6 and 7 remain both economically and statistically robust. In the emissions regression, the coefficient on the triple interaction term *DivFirm* × *After* × *Constrained* is 0.30 and significant at 1%, comparable to 0.34 in Table 6. The marginal probabilities of non-Californian plant closures and openings by constrained firms with operations in California after the cap-and-trade rule are –0.05 (significant at 1%) and 0.15 (significant at 5%), respectively, comparable to –0.06 and 0.18 in Table 7.

In short, resource shifting by firms toward plants that face better growth opportunities does not explain our results on the spillover effects from the California cap-and-trade rule.

5.6.2 Placebo Tests

Another potential critique of our approach is that the DID design could be picking up spurious changes over time that happen to affect California more than the average state, but might also be occurring in other economically important states with a large presence of greenhouse gas emitting plants. We argue that this is unlikely the case, because in the absence of similarly extensive and costly climate regulation programs in other states, it is difficult to justify why there would be a significant change in the emissions and ownership across plants in one state relative to other states that is mainly driven by firms with operations in that one state, and why the reallocation would be conditioned by their financial constraints.

Notwithstanding this argument, we conduct placebo tests to rule out this concern. We use two alternative states that are the most important greenhouse gas emitters aside from California, i.e., Texas and Louisiana (see Table 2), as placebo states. We test whether geographically diversified firms (i.e., firms with a presence both in the placebo state and in other states) reduce plant emissions and ownership in the placebo state compared to other states, whether these firms create emission spillovers and increase plant ownership in other states, and whether these effects are related to firm financial constraints. To this end, we employ pooled triple difference regressions as in Equations (3) and (4) with our composite constraint indicator. To conserve space, we omit the full sample results and focus on the cross-state reallocation of geographically diversified firms and spillover effects to other states.

Panels A and B of Table 11 each report results from using Texas and Louisiana as placebo states, respectively. The first two columns present regressions of plant emissions in the same format as the first columns in Table 6. For both placebo states, we do not find results similar to Table 6. There is neither any indication that plants in a placebo state reduce emissions by more than plants in other states, nor any evidence that there are spillover effects from a placebo state to other states, nor any observable effects driven by financial constraints. In the remaining four columns, we conduct placebo tests for plant ownership using multinomial logit regressions as in Table 7 with a categorical dependent variable equal to –1 for a plant ownership reduction, 0 for no ownership change, and +1 for an ownership increase. For both placebo states, we find no evidence of changes in plant ownership that resemble the changes observed in Table 7. Given the large number of observations in the placebo tests that are comparable to those in our main analysis, the lack of evidence is unlikely a result of low statistical power. In short, our results do not seem to be driven by a general change over time coinciding with the introduction of the California cap-and-trade rule that affects heavy greenhouse gas emitter states in a similar way.

6 Conclusion

In this paper, we study the internal resource allocation responses by firms that lead to real, unintended spillover effects of localized climate policies driven by the financial constraints of firms who operate large greenhouse gas emitting plants. Using a detailed plant level dataset on greenhouse gas emissions and parent company ownership made available by the US EPA and hand-matched to Compustat, we show that the California cap-and-trade rule introduced at the beginning of 2013 has real spillover effects across other US states through firm financial constraints.

Motivated by the finance literature, we hypothesize that financially constrained firms reallocate their greenhouse gas emissions and plant ownership away from California to other states in the face of heightened regulatory costs that alter the relative net expected returns across plants. The economic mechanism is that because the costs of external capital render optimal levels of emissions in California infeasible for constrained firms, the net returns from internal reallocations become relatively more attractive. We document strong evidence for reallocations of emissions across plants already in place, and to a lesser degree also for changes in plant ownership. The overall consequence of this reallocation is that firms show no evidence of reducing their total emissions, and in fact that constrained firms strictly increase their emissions firm-wide. Our results are consistent with the internal reallocation of corporate pollutive activities and resources to avoid regulatory costs in the face of limited access to external financing, highlighting the hidden costs of environmental policies through financial channels.

Our study makes a significant contribution to the understanding of the interplay between climate policy and firm behavior, and provides a stepping-stone towards more effectively coordinated solutions to climate change by informing policy makers of the potential externalities from regionally segmented climate policies. If localized climate policies prove ineffective even within one country, they are unlikely to have the intended effect of reducing emissions on a global scale across countries. This paper also contributes to the growing field of corporate environmental policies by focusing on the internal plant level emission activities and resource allocations within firms, thus providing a unique channel for the real effects of climate policy through the importance of firm financial constraints.

Bibliography

Akey, Pat and Ian Appel, 2018. The limits of limited liability: Evidence from industrial pollution. Working Paper.

Almeida, Heitor, Murillo Campello, and Michael S. Weisbach, 2004. The cash flow sensitivity of cash. *The Journal of Finance*, Vol. 59, No. 4, pp. 1777–1804.

Becker, Randy and Vernon Henderson, 2000. Effects of air quality regulations on polluting industries. *Journal of Political Economy*, Vol. 108, No. 2, pp. 379–421.

Billett, Matthew T. and David C. Mauer, 2003. Cross-subsidies, external financing constraints, and the contribution of the internal capital market to firm value. *Review of Financial Studies*, Vol. 16, No. 4, pp. 1167–1201.

Boden, Thomas A., Gregg Marland, and Robert J. Andres, 2017. Global, regional, and national fossil-fuel CO_2 emissions. Carbon Dioxide Information Analysis Center, Oak Ridge National Laboratory, U.S. Department of Energy, Oak Ridge, Tenn., U.S.A.

Bushnell, James B., Stephen P. Holland, Jonathan E. Hughes, and Christopher R. Knittel, 2017. Strategic policy choice in state-level regulation: The EPA's clean power plan. *American Economic Journal: Economic Policy*, Vol. 9, No. 2, pp. 57–90.

Cole, Matthew A. and Robert J. R. Elliott, 2005. FDI and the capital intensity of "dirty" sectors: A missing piece of the pollution haven puzzle. *Review of Development Economics*, Vol. 9, No. 4, pp. 530–548.

Coria, Jessica and Jurate Jaraite, 2015. Carbon pricing: Transaction costs of emissions trading vs. carbon taxes. CERE Working Paper, Vol. 2015, No. 2, pp. 1–32.

Ederington, Josh, Arik Levinson, and Jenny Minier, 2005. Footloose and pollution-free. *Review of Economics and Statistics*, Vol. 87, No. 1, pp. 92–99.

Fabra, Natalia and Mar Reguant, 2014. Pass-through of emissions costs in electricity markets. *The American Economic Review*, Vol. 104, No. 9, pp. 2872–2899.

Fazzari, Steven M. and Bruce C. Petersen, 1993. Working capital and fixed investment: New evidence on financing constraints. *The RAND Journal of Economics*, Vol. 24, No. 3, pp. 328–342.

Fazzari, Steven M., R. Glenn Hubbard, and Bruce C. Petersen, 2000. Investment-cash flow sensitivities are useful: A comment on Kaplan and Zingales. *The Quarterly Journal of Economics*, Vol. 115, No. 2, pp. 695–705.

Fazzari, Steven M., R. Glenn Hubbard, Bruce C. Petersen, Alan S. Blinder, and James M. Poterba, 1988. Financing constraints and corporate investment. *Brookings Papers on Economic Activity*, Vol. 1988, No. 1, pp. 141–206.

Fowlie, Meredith, Mar Reguant, and Stephen P. Ryan, 2016. Market-based emissions regulation and industry dynamics. *Journal of Political Economy*, Vol. 124, No. 1, pp. 249–302.

Fowlie, Meredith, Michael Greenstone, and Catherine Wolfram, 2018. Do energy efficiency investments deliver? Evidence from the weatherization assistance program. *The Quarterly Journal of Economics*, Vol. 133, No. 3, pp. 1597–1644.

Gertner, Robert H., David S. Scharfstein, and Jeremy C. Stein, 1994. Internal versus external capital markets. *The Quarterly Journal of Economics*, Vol. 109, No. 4, pp. 1211–1230.

Hadlock, Charles J. and Joshua R. Pierce, 2010. New evidence on measuring financial constraints: Moving beyond the KZ index. *Review of Financial Studies*, Vol. 23, No. 5, pp. 1909–1940.

Hall, Bronwyn H., Adam Jaffe, and Manuel Trajtenberg, 2005. Market value and patent citations. *The RAND Journal of Economics*, Vol. 36, No. 1, pp. 16–38.

Jovanovic, Boyan and Peter L. Rousseau, 2001. Why wait? A century of life before IPO. *American Economic Association Papers and Proceedings*, Vol. 91, No. 2, pp. 336–341.

Kaplan, Steven N. and Luigi Zingales, 1997. Do investment-cash flow sensitivities provide useful measures of financing constraints? *The Quarterly Journal of Economics*, Vol. 112, No. 1, pp. 169–215.

Kellenberg, Derek K., 2009. An empirical investigation of the pollution haven effect with strategic environment and trade policy. *Journal of International Economics*, Vol. 78, No. 2, pp. 242–255.

Khan, Hashmat, Christopher R. Knittel, Konstantinos Metaxoglou, and Maya Papineau, 2016. Carbon emissions and business cycles. NBER Working Paper, No. 22294.

Kim, Taehyun and Qiping Xu, 2018. Financial constraints and corporate environmental policies. Working Paper.

Lamont, Owen, 1997. Cash flow and investment: Evidence from internal capital markets. *The Journal of Finance*, Vol. 52, No. 1, pp. 83–109.

Lamont, Owen, Christopher Polk, and Jesús Saá-Requejo, 2001. Financial constraints and stock returns. *Review of Financial Studies*, Vol. 14, No. 2, pp. 529–554.

Levinson, Arik and M. Scott Taylor, 2008. Unmasking the pollution haven effect. *International Economic Review*, Vol. 49, No. 1, pp. 223–254.

Marin, Giovanni, Marianna Marino, and Claudia Pellegrin, 2018. The impact of the European emission trading scheme on multiple measures of economic performance. *Environmental and Resource Economics*, Vol. 71, No. 2, pp. 551–582.

Martin, Ralf, Mirabelle Muûls, Laure B. De Preux, and Ulrich J. Wagner, 2014. Industry compensation under relocation risk: A firm-level analysis of the EU emissions trading scheme. *The American Economic Review*, Vol. 104, No. 8, pp. 2482–2508.

Matvos, Gregor and Amit Seru, 2014. Resource allocation within firms and financial market dislocation: Evidence from diversified conglomerates. *Review of Financial Studies*, Vol. 27, No. 4, pp. 1143–1189.

Matvos, Gregor, Amit Seru, and Rui Silva, 2018. Financial market frictions and diversification. *Journal of Financial Economics*, Vol. 127, No. 1, pp. 21–50.

Mulatu, Abay, Reyer Gerlagh, Dan Rigby, and Ada Wossink, 2010. Environmental regulation and industry location in Europe. *Environmental and Resource Economics*, Vol. 45, No. 4, pp. 459–479.

Nordhaus, William D. and Zili Yang, 1996. A regional dynamic general-equilibrium model of alternative climate-change strategies. *The American Economic Review*, Vol. 86, No. 4, pp. 741–765.

Nordhaus, William D., 1977. Economic growth and climate: The carbon dioxide problem. *The American Economic Review: Papers and Proceedings*, Vol. 67, No. 1, pp. 93–106.

Nordhaus, William D., 1977. Strategies for the control of carbon dioxide. Cowles Foundation Discussion Paper, No. 443.

Nordhaus, William D., 2015. Climate clubs: Overcoming free-riding in international climate policy. *The American Economic Review*, Vol. 105, No. 4, pp. 1339–1370.

Shin, Hyun-Han and René M. Stulz, 1998. Are internal capital markets efficient? *The Quarterly Journal of Economics*, Vol. 113, No. 2, pp. 531–552.

Shive, Sophie and Margaret Forster, 2018. Corporate governance and pollution externalities of public and private firms. Working Paper.

Stein, Jeremy C., 1997. Internal capital markets and the competition for corporate resources. *The Journal of Finance*, Vol. 52, No. 1, pp. 111–133.

Sulaeman, Johan and Abhishek Varma, 2018. The "green" geography: Corporate environmental policies and local institutional investors. Working Paper.

Wagner, Ulrich J. and Christopher D. Timmins, 2009, Agglomeration effects in foreign direct investment and the pollution haven hypothesis. *Environmental and Resource Economics*, Vol. 43, No. 2, pp. 231–256.

Whited, Toni M. and Guojun Wu, 2006, Financial constraints risk. *Review of Financial Studies*, Vol. 19, No. 2, pp. 531–559.

World Bank and Ecofys, 2016. Carbon pricing watch. A report prepared jointly by the World Bank and Ecofys, Washington, D.C.

Figure 1: Global Carbon Emissions and Temperature Changes

The figure shows the time-series of worldwide total carbon emissions from fossil fuel consumption and cement production (thick solid line, left axis) and the global land-ocean surface temperature index (thin line with markers, right axis). Total carbon emissions data is from Boden, Marland, and Andres (2017), and global temperature index data is from the NASA Goddard Institute for Space Studies (GISS). The temperature index is computed as deviations from the mean over a base period. Details regarding its computation can be found at https://data.giss.nasa.gov/gistemp/.

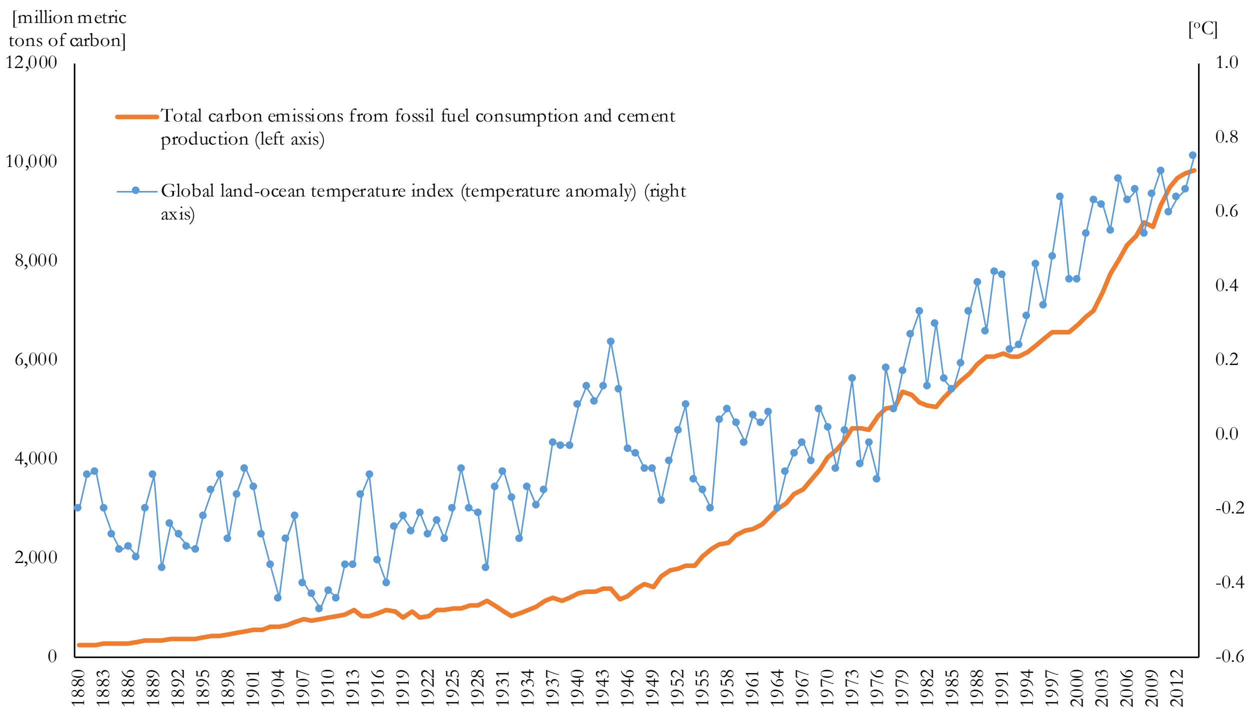
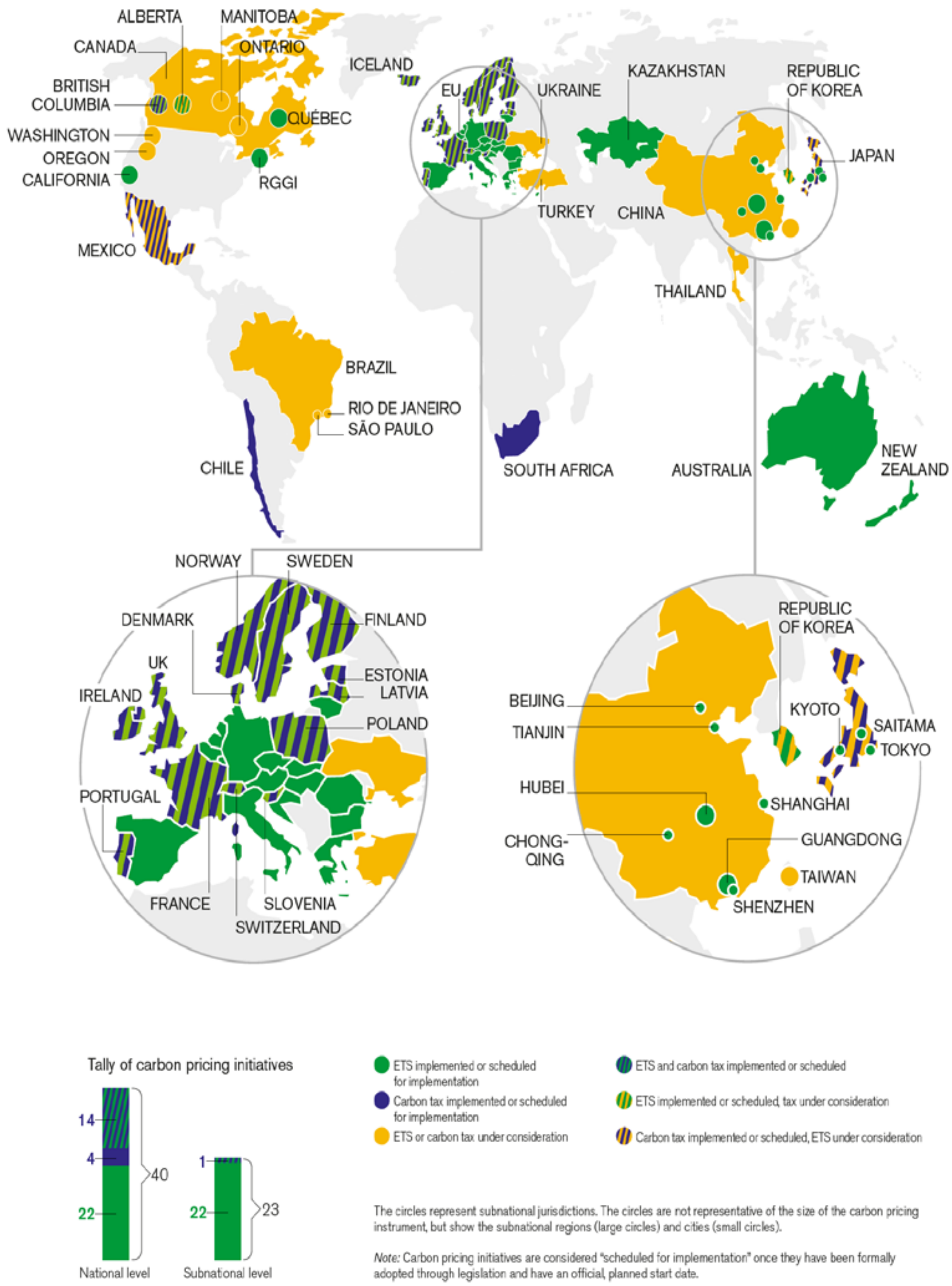


Figure 2: Climate Policies Around the World

The figure shows major climate policies such as carbon emission trading rules or carbon taxes implemented in various countries and states. Panel A shows a summary map of existing, emerging and potential regional, national and subnational carbon pricing initiatives (ETS and tax). Panel B shows the share of global emissions covered by regional, national and subnational carbon pricing initiatives. These figures are reproduced from World Bank and Ecofys (2016, pages 4 and 5).

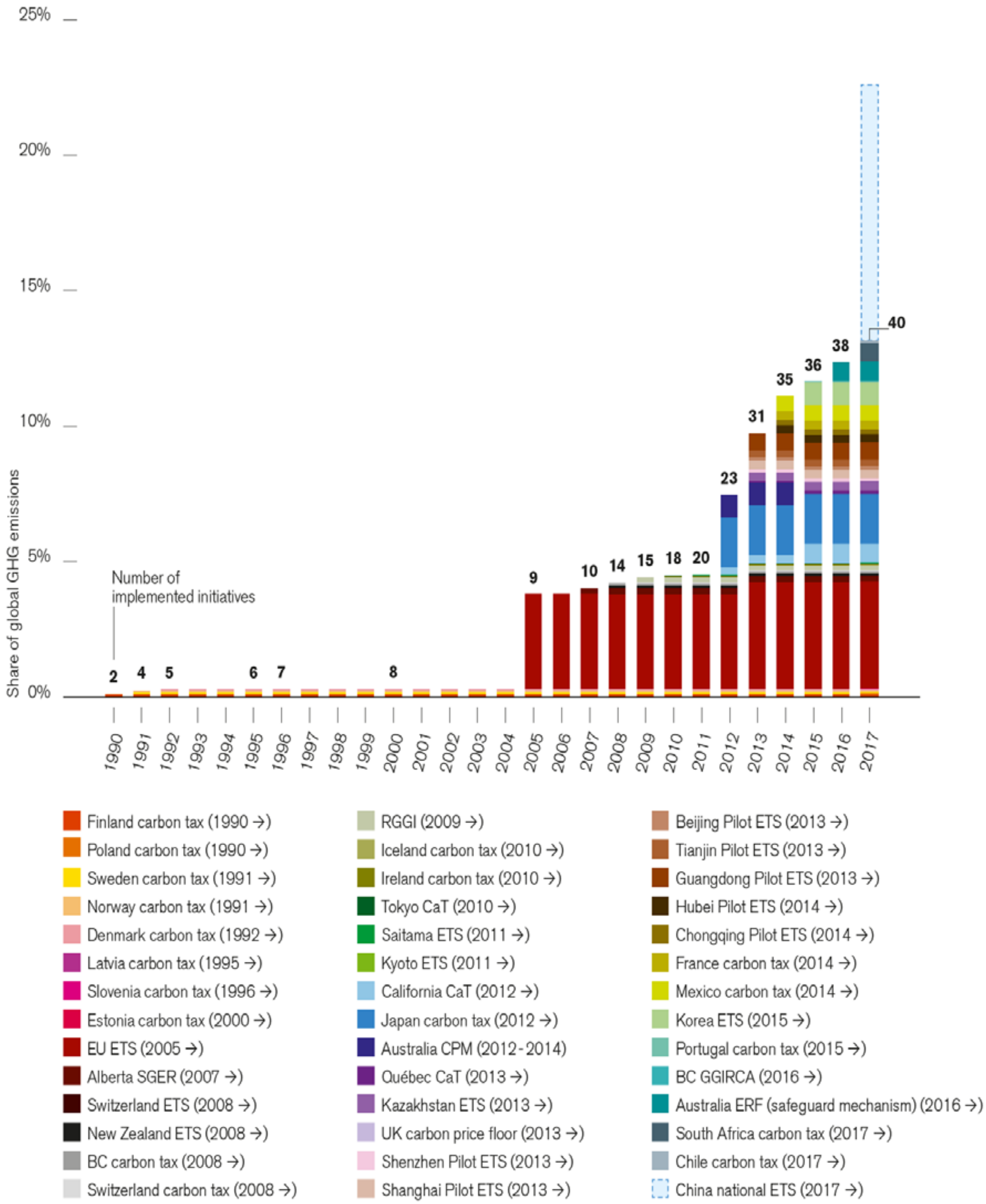
Panel A: Summary Map of Carbon Pricing Initiatives



(continued)

Figure 2: Climate Policies Around the World (continued)

Panel B: Share of Global Emissions Covered by Carbon Pricing Initiatives



Note: Only the introduction or removal of an ETS or carbon tax is shown. Emissions are given as a share of global GHG emissions in 2012. Annual changes in global, regional, national, and subnational GHG emissions are not shown in the graph. Data on the coverage of the city-level Kyoto ETS were not accessible and the British Columbia Greenhouse Gas Industrial Reporting and Control Act (GGIRCA) does not cover any emissions yet; their coverages are therefore shown as zero. The information on the Chinese national ETS represents early unofficial estimates based on the Chinese President's announcement in September 2015.

Figure 3: Transaction Prices and Volume of California Carbon Allowance Futures

The figure shows California carbon allowance future prices along with their trading volumes. Transaction prices (in \$/metric ton) are shown on the left axis, while trading volume (in thousands of metric tons of CO_2) is shown on the right axis. The graph shows data for futures contracts with different expiration dates (December 2013, December 2014, December 2015, December 2016). The vertical lines mark the periods in which the different futures contracts are traded, as well as the introduction of the California cap-and-trade system at the beginning of 2013. The data is from the Climate Policy Initiative & Intercontinental Exchange.

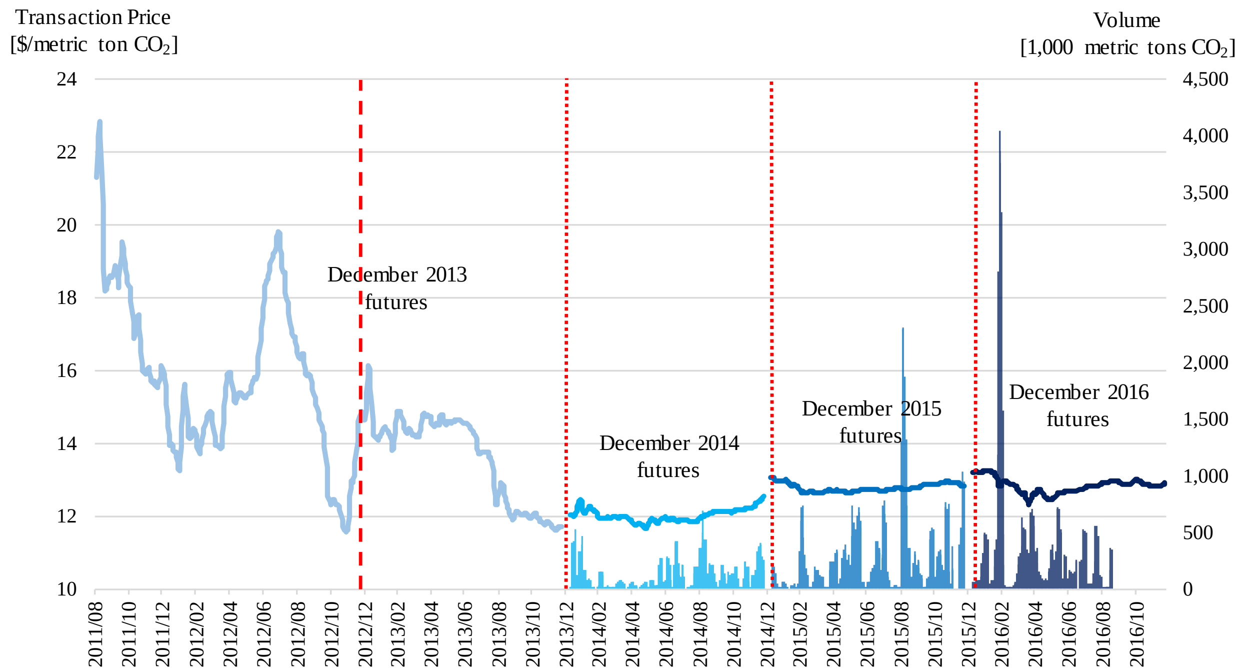


Figure 4: Economic Framework

The figure illustrates the economic channel of the main hypothesis. Revenues and costs (p) are plotted on the vertical axis, and emissions and production quantities (q) are plotted on the horizontal axis. The thick dashed horizontal line indicates the level of average costs above which constraints bind for financially constrained firms. Marginal (solid) and average (dotted) revenue curves, denoted *mr* and *ar*, are downward sloping consistent with an imperfectly competitive market. Marginal and average cost curves are plotted for three scenarios. mc_{ca} and ac_{ca} represent the pre-cap-and-trade costs of producing and emitting in California. mc'_{ca} and ac'_{ca} denote the post-cap-and-trade costs of emitting in California, which are tilted upward from the pre-policy curves for emission quantities above the free allocation amount. mc_{oth} and ac_{oth} are the cost curves should firms reallocate their emissions exceeding the free allocation amount to other states. *I*, *I′*, and *I″* each denote the equilibrium with the optimal amount of emissions in California before the cap-and-trade rule, in California after the cap-and-trade rule, and in other states, respectively.

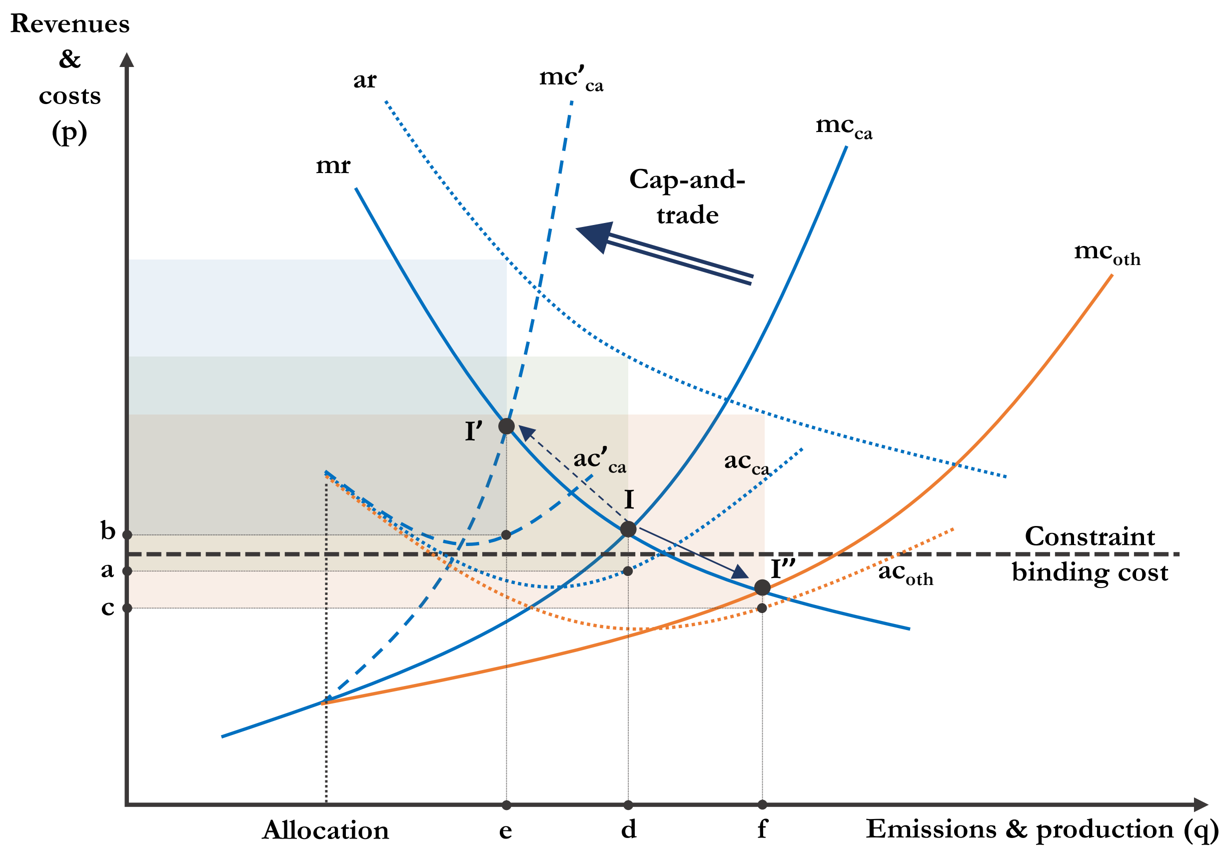
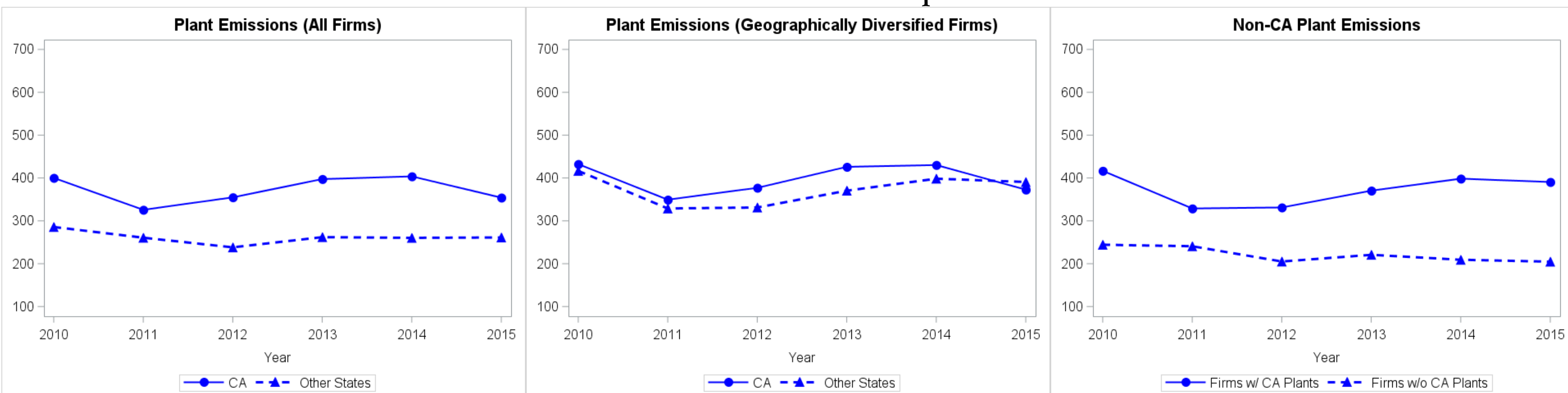


Figure 5: Unconditional Average Responses to Cap-and-Trade

The figure shows average plant emissions (Panel A) and plant ownership (Panel B) during the sample period 2010–2015, i.e. before and after the enactment of the California cap-and-trade program at the beginning of 2013. Outcome variables of the treatment and control group are plotted as solid and dotted lines, respectively. Panel A shows three graphs: Emissions (thousand metric tons) of plants in California and in other states based on all firms; emissions of plants in California and in other states based on geographically diversified firms; and emissions of non-California plants for firms with and without plants in California. Panel B shows three graphs: Plant ownership (number of plants owned by firm) of plants in California and in other states based on all firms; plant ownership of plants in California and in other states based on geographically diversified firms; and plant ownership of non-California plants for firms with and without plants in California.

Panel A: Plant Emission Responses



Panel B: Plant Ownership Responses

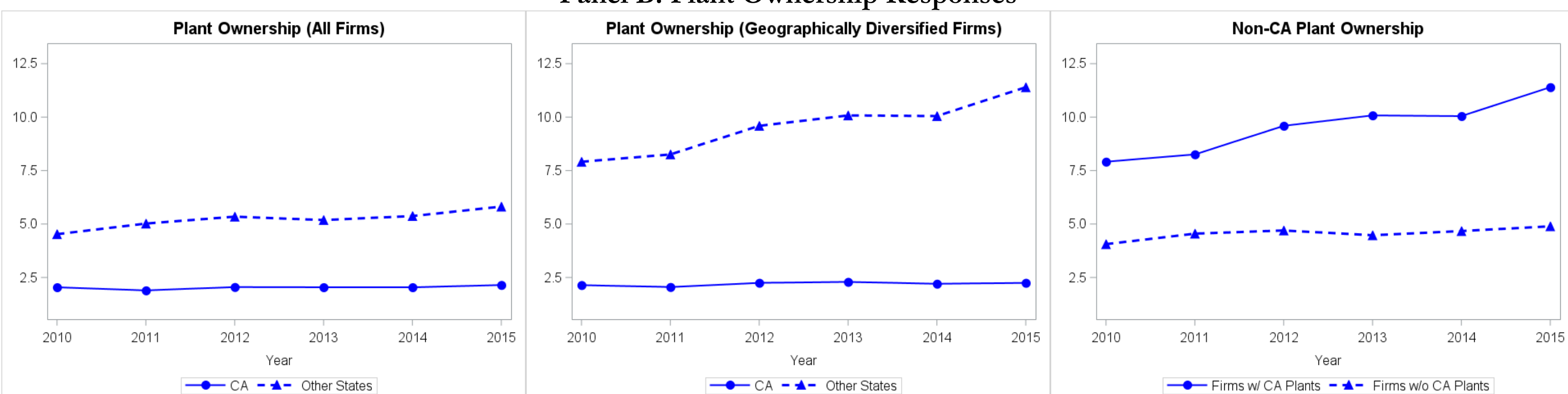
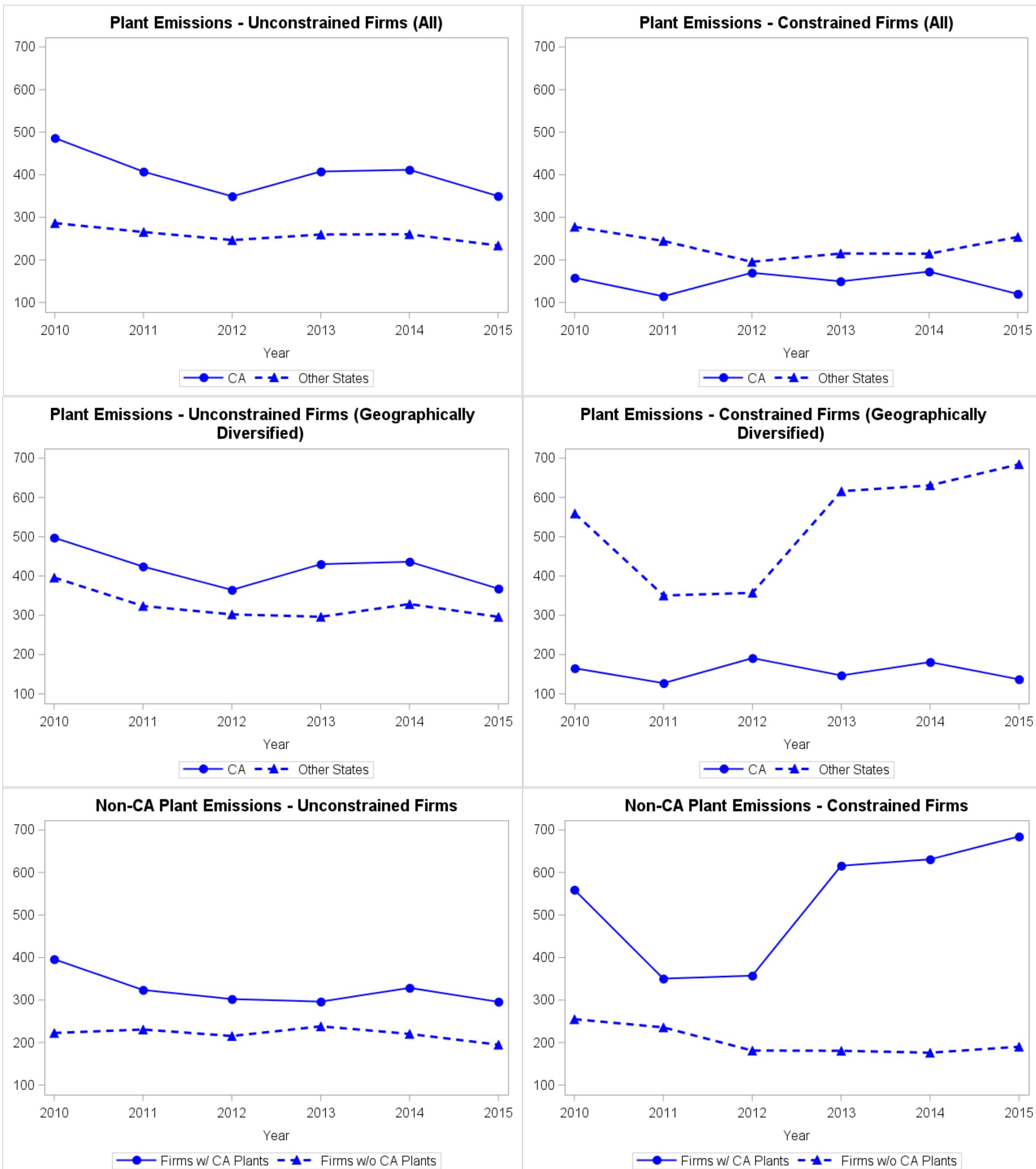


Figure 6: Average Responses of Constrained vs Unconstrained Firms

The figure shows average plant emissions (Panel A) and plant ownership (Panel B) separately for constrained and unconstrained firms during the sample period 2010–2015, i.e. before and after the enactment of the California cap-and-trade program at the beginning of 2013. Outcome variables of the treatment and control group are plotted as solid and dotted lines, respectively. Separately for constrained and unconstrained firms, Panel A shows three sets of graphs: Emissions (in thousand metric tons) of plants in California and in other states based on all firms; emissions of plants in California and in other states based on geographically diversified firms; and emissions of non-California plants for firms with and without plants in California. Separately for constrained and unconstrained firms, Panel B shows three sets of graphs: Plant ownership (measured as number of plants owned by a firm) of plants in California and in other states based on all firms; plant ownership of plants in California and in other states based on geographically diversified firms; and plant ownership of non-California plants for firms with and without plants in California.

Panel A: Plant Emission Responses



(continued)

Figure 6: Average Responses of Constrained vs Unconstrained Firms (continued)

Panel B: Plant Ownership Responses

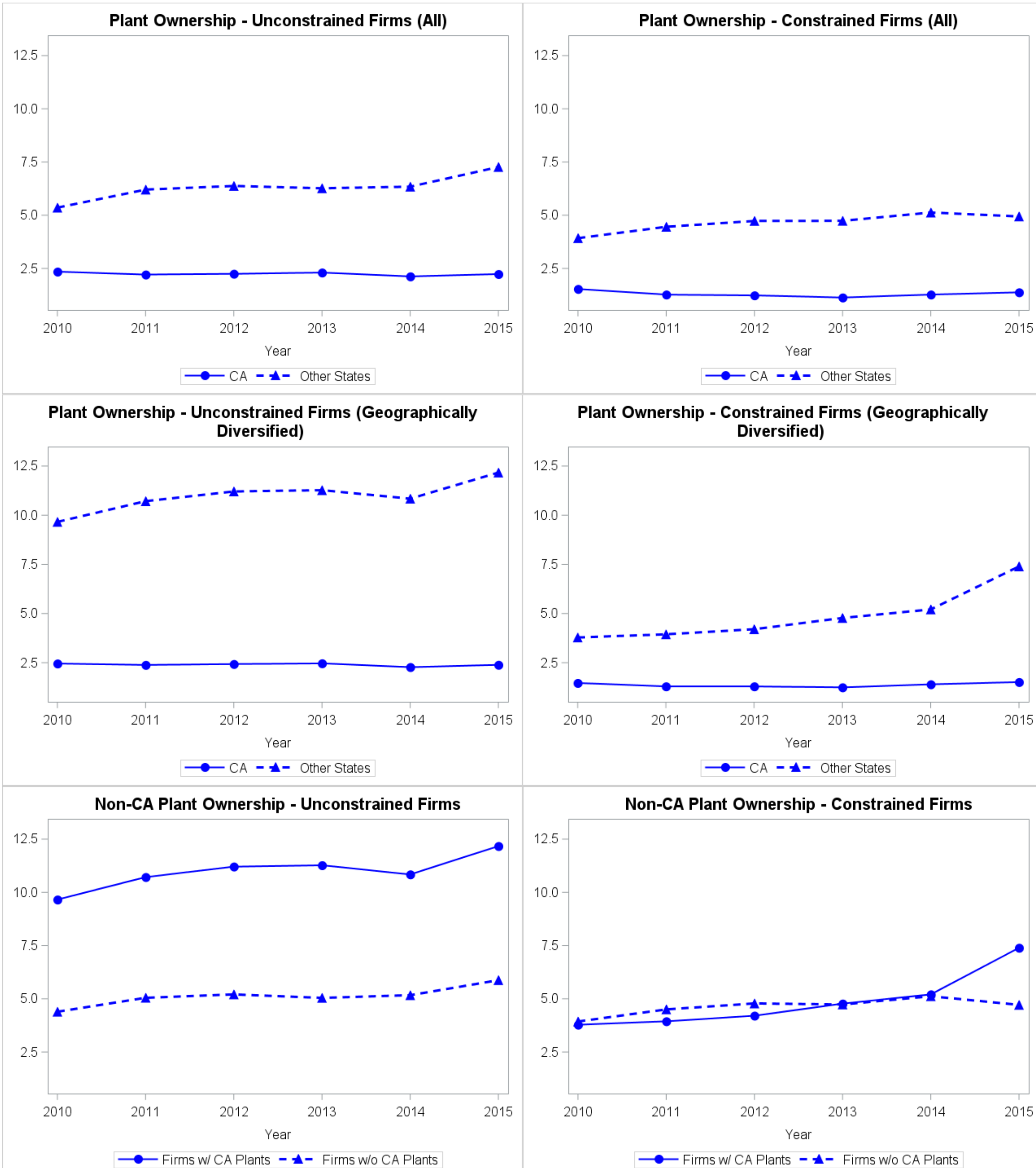


Table 1: Allowance Auctions, Allocations, and Transactions of California Cap-and-Trade

The table shows descriptive statistics on allowance auctions, allocations and transactions of California carbon allowances pursuant to the cap-and-trade program. With regards to allowance auctions, Panel A shows for different auction periods the number of bidders, available and sold quantities, the ratio of the number of bids to available quantities, the reserve price and settlement price. Panel B summarizes available data on the quantities of free allocations to industrial plants. Panel C shows for different years and allowance vintages the number of transactions, quantities and weighted average prices (for combined California and Quebec market). Data are from the California Air Resources Board.

Panel A: Allowance Auctions

Auction period		Number of bidders (organizations)	Available (thousand metric tons)	Sold (thousand metric tons)	Bids /Available	Reserve price ($/metric ton)	Settlement price ($/metric ton)
2012/11	Current vintage	73	23,126	23,126	1.06	10.00	10.09
	Future (3yr) vintage		39,450	5,576	0.14	10.00	10.00
2013/02	Current vintage	91	12,925	12,925	2.49	10.71	13.62
	Future (3yr) vintage		9,560	4,440	0.46	10.71	10.71
2013/05	Current vintage	81	14,522	14,522	1.78	10.71	14.00
	Future (3yr) vintage		9,560	7,515	0.79	10.71	10.71
2013/08	Current vintage	79	13,865	13,865	1.62	10.71	12.22
	Future (3yr) vintage		9,560	9,560	1.69	10.71	11.10
2013/11	Current vintage	77	16,615	16,615	1.82	10.71	11.48
	Future (3yr) vintage		9,560	9,560	1.64	10.71	11.10
2014/02	Current vintage	71	19,539	19,539	1.27	11.34	11.48
	Future (3yr) vintage		9,260	9,260	1.11	11.34	11.38
2014/05	Current vintage	74	16,947	16,947	1.46	11.34	11.50
	Future (3yr) vintage		9,260	4,036	0.44	11.34	11.34
2014/08	Current vintage	71	22,473	22,473	1.14	11.34	11.50
	Future (3yr) vintage		9,260	6,470	0.70	11.34	11.34
2014/11*	Current vintage	83	23,071	23,071	1.73	11.34	12.10
	Future (3yr) vintage		10,787	10,787	1.92	11.34	11.86
2015/02	Current vintage	87	73,611	73,611	1.14	12.10	12.21
	Future (3yr) vintage		10,432	10,432	1.02	12.10	12.10
2015/05	Current vintage	97	76,932	76,932	1.16	12.10	12.29
	Future (3yr) vintage		10,432	9,812	0.94	12.10	12.10
2015/08	Current vintage	88	73,429	73,429	1.28	12.10	12.52
	Future (3yr) vintage		10,431	10,431	1.78	12.10	12.30
2015/11	Current vintage	91	75,113	75,113	1.14	12.10	12.73
	Future (3yr) vintage		10,432	10,432	1.32	12.10	12.65

*: Joint auction with Quebec cap-and-trade from this point onward

(continued)

Table 1: Allowance Auctions, Allocations, and Transactions of California Cap-and-Trade (continued)

Panel B: Free Allocations to Industrial Plants

	2013	2014	2015
Allocation (thousand metric tons)	53,895	54,394	55,827
Number of plants	139	156	159
Per-plant allocation	388	349	351

Panel C: Market Transactions and Prices

Allowance vintage	Number of transactions	Thousand metric tons	Weighted avg price
2014 (Obligations from 2013 emissions due)			
2013	228	12,984	12.23
2014	338	33,588	11.98
Current total	**566**	**46,571**	**12.05**
2015	3	775	12.58
2016	35	12,012	11.92
2017	54	21,330	11.73
Future total	**92**	**34,117**	**11.82**
2015 (Obligations from 2014 emissions due)			
2013	87	6,385	12.51
2014	248	29,417	12.62
2015	444	112,921	12.68
Current total	**779**	**148,723**	**12.66**
2016	44	21,982	12.72
2017	60	20,699	12.65
2018	62	27,543	12.61
Future total	**166**	**70,223**	**12.66**
2016 (Obligations from 2015 emissions due)			
2013	23	1,237	12.50
2014	33	5,612	12.75
2015	431	65,652	12.72
2016	333	62,882	12.75
Current total	**820**	**135,383**	**12.74**
2017	21	11,352	12.88
2018	25	14,308	12.83
2019	8	2,820	12.77
Future total	**54**	**28,480**	**12.85**

Table 2: Number of Plants and Firms by State

The table shows the number of sample plants located in each state, the number of sample firms operating in each state, as well as the average plant emissions (in thousands of metric tons) and the average firm assets (in $ billions). States are sorted in descending order by the number of firms. The table also shows the totals across all states and firms with plants both in California and other states. The data is from the intersection of the EPA and Compustat databases. The sample period is 2010–2015.

State	Number of plants	Number of firms	Avg. emissions (thousand metric tons)	Avg. firm assets ($ billions)	State	Number of plants	Number of firms	Avg. emissions (thousand metric tons)	Avg. firm assets ($ billions)
Texas	587	174	300.53	20.51	Mississippi	24	23	304.41	17.17
Louisiana	225	104	326.50	28.09	New Jersey	19	21	394.75	50.88
California	**161**	**85**	**398.04**	**28.58**	Utah	29	20	180.06	35.49
Pennsylvania	133	73	276.87	24.47	Missouri	20	19	153.21	53.84
Illinois	88	70	707.61	21.42	Oregon	18	18	59.32	12.38
Ohio	95	68	371.01	24.22	Alaska	39	14	468.66	44.45
Oklahoma	170	59	222.70	19.49	North Dakota	16	13	224.44	18.93
Colorado	142	54	147.48	19.09	Nebraska	16	13	174.52	13.50
Indiana	61	50	529.68	24.70	Massachusetts	14	13	104.52	35.73
Michigan	67	48	246.03	30.99	Nevada	13	11	306.27	24.78
Alabama	59	47	254.41	22.52	Arizona	10	11	157.88	27.75
West Virginia	83	41	183.99	17.04	Idaho	16	10	51.44	24.78
Kentucky	53	37	314.86	16.78	Connecticut	13	10	121.26	51.72
Virginia	52	35	172.63	18.71	Maine	8	9	308.75	5.25
Tennessee	34	35	337.94	20.68	Montana	6	9	555.58	31.60
Minnesota	40	34	203.50	19.81	South Dakota	5	7	124.51	14.49
Kansas	36	33	293.43	18.47	Maryland	4	7	293.81	6.35
Georgia	36	30	158.02	27.75	Delaware	4	5	694.94	24.84
Wisconsin	34	30	111.38	14.63	Puerto Rico	4	5	70.97	39.58
Iowa	34	29	308.23	19.87	Hawaii	3	3	332.37	44.81
New Mexico	52	28	155.69	38.24	Vermont	1	1	39.33	11.04
Arkansas	44	28	125.15	14.49	Virgin Islands	1	1	36.10	34.67
New York	30	27	239.04	26.20	New Hampshire	1	1	18.24	104.57
North Carolina	39	26	370.64	12.92					
South Carolina	26	26	182.72	14.69	**All States**	**2,806**	**511**	**288.97**	**17.25**
Wyoming	65	24	191.91	23.54					
Florida	43	24	325.11	22.08	**Firms with Cal &**				
Washington	33	24	247.03	21.16	**Non-Cal Plants**	**948**	**70**	**424.03**	**29.22**

Table 3: Firm and Plant Characteristics

The table presents sample summary statistics of firm characteristics (Panel A) and plant characteristics (Panel B). In Panel A, emissions (in thousand metric tons) are summed across plants owned by a firm and reported as a firm-level measure. Total assets are in $ billions. Property, plant, and equipment (PP&E), capital expenditures, short-term and long-term debt, cash, and cash flow are shown as a fraction of total assets. Profitability is return on assets (ROA). Tobin's Q is the market value of assets divided by the book value of assets. R&D is scaled by sales. R&D stock is calculated using the perpetual inventory method (Hall, Jaffe, and Trajtenberg, 2005). Payout ratio is cash dividends plus repurchases divided by income before extraordinary items. Firm age is the difference between the observation year and founding year as in Jovanovic and Rousseau (2001). The panel reports the number of firm-year observations, average, standard deviation, minimum, 25th percentile, median, 75th percentile, and maximum values for these variables. The panel also reports the number of firm-year observations with and without a long-term or short-term bond rating. Panel B presents summary statistics for plant level carbon emissions and ownership structure. Carbon emissions are in thousand metric tons. Ownership structure is measured as the fraction of a plant owned by a firm, the number of plants owned by a firm, and the number of firms sharing ownership in a plant. The panel shows the number of plant-year observations, average, standard deviation, minimum, 25th percentile, median, 75th percentile, and maximum of each variable. Emissions and plant ownership data is from EPA. Accounting data are from Compustat. The sample period is 2010–2015.

Panel A: Summary Statistics of Firm Characteristics

	Firm-year obs.	Average	Std. dev.	Minimum	25th percentile	Median	75th percentile	Maximum
Emissions (thousand metric tons)	2,303	1,584.42	4,039.64	0.03	69.43	277.16	1,208.48	65,101.03
Total assets ($ billions)	2,303	17.28	27.85	0.03	1.50	4.18	18.96	111.03
PP&E	2,302	0.47	0.25	0.00	0.27	0.44	0.67	0.95
Capital expenditures	2,232	0.10	0.13	0.00	0.03	0.05	0.11	0.73
Short-term debt	2,302	0.03	0.05	0.00	0.00	0.01	0.03	0.59
Long-term debt	2,295	0.28	0.19	0.00	0.15	0.25	0.38	1.11
Cash	2,302	0.09	0.10	0.00	0.02	0.06	0.13	0.73
Cash flow	2,233	0.14	0.11	-0.67	0.09	0.13	0.19	0.66
Profitability	2,233	0.05	0.11	-0.95	0.01	0.05	0.09	0.55
Tobin's Q	2,086	1.46	0.57	0.49	1.07	1.33	1.69	5.03
R&D	2,303	0.02	0.05	0.00	0.00	0.00	0.01	0.90
R&D stock	2,303	0.11	0.44	0.00	0.00	0.00	0.07	8.00
Payout ratio	2,303	0.49	1.34	-6.03	0.00	0.28	0.79	9.52
Firm age	2,303	27.02	20.82	1.00	9.00	20.00	46.00	65.00
Bond rating (long-term, >1yr)								
No rating	1,047							
Rated	1,256							
Bond rating (short-term, <1yr)								
No rating	1,776							
Rated	527							

Panel B: Plant Level Carbon Emissions and Ownership Structure

	Plant-year obs.	Average	Std. dev.	Minimum	25th percentile	Median	75th percentile	Maximum
Carbon emissions (thousand metric tons)	13,679	288.97	628.75	0.00	339.71	690.18	2,038.08	4,466.38
Ownership								
Fraction of plant owned by a firm (percent)		91.12	22.86	0.00	100.00	100.00	100.00	100.00
Number of plants owned by a firm		5.94	10.54	1.00	1.00	3.00	6.00	142.00
Number of firms owning a plant		1.15	0.45	1.00	1.00	1.00	1.00	6.00

Table 4: Plant Emission Responses to California Cap-and-Trade Rule

Panel A presents univariate evidence on the differences in emissions growth around the introduction of the California cap-and-trade rule (calculated as [mean(after) – mean(before)]/mean) between California and non-California plants for geographically diversified firms, and between non-California plants owned by geographically diversified and undiversified firms, as well as their corresponding *t*-statistics. Panel B presents results from firm-plant level difference-in-difference (DID) regressions. Columns (1)-(4) compare California and non-California plants based on geographically diversified firms. Columns (5)-(8) study spillovers to non-California Plans comparing geographically diversified and non-diversified firms. The dependent variable is log (1+Emissions). The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. Control variables include PP&E, R&D Stock as well as plant and year fixed effects. Coefficients and their respective standard errors adjusted for clustering at the plant level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. Number of observations and adjusted R^2 are also reported.

Panel A: Univariate Evidence on Differences in Emissions Growth

California vs non-California plants (Geographically diversified firms)				Spillovers to non-California plants (Diversified vs undiversified firms)			
California	Other states	Diff.	t-stat.	Diversified	Undiversified	Diff.	t-stat.
-0.13	-0.07	-0.06	-1.03	-0.07	-0.12	0.05	2.20

Panel B: Plant Emission Responses to California Cap-and-Trade Rule

	California vs non-California plants (Geographically diversified firms)				Spillovers to non-California plants (Diversified vs undiversified firms)			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CalPlant x After	-0.210	-0.206	-0.161*	-0.155*				
	(0.132)	(0.132)	(0.093)	(0.094)				
CalPlant	0.034	0.030						
	(0.220)	(0.220)						
DivFirm x After					0.165***	0.156***	0.140***	0.131***
					(0.059)	(0.059)	(0.038)	(0.038)
DivFirm					0.170**	0.176**	-0.155	-0.159
					(0.081)	(0.081)	(0.106)	(0.105)
After	-0.083*				-0.249***			
	(0.048)				(0.032)			
PP&E				-0.203*				-0.295***
				(0.108)				(0.068)
R&D stock				-0.006				0.017
				(0.006)				(0.012)
Plant FE	No	No	Yes	Yes	No	No	Yes	Yes
Year FE	No	Yes	Yes	Yes	No	Yes	Yes	Yes
Observations	4,166	4,166	3,961	3,961	12,846	12,846	12,521	12,511
Adjusted R^2	0.001	0.001	0.862	0.862	0.008	0.010	0.745	0.746

Table 5: Firm Financial Constraints and Plant Emission Responses: Separate Regressions by Constraint Groups

The table presents results from DID regressions, separately for subsamples of financially constrained and unconstrained firms. A number of measures for financial constraints are used: Our composite measure, the Kaplan-Zingales index (following Kaplan and Zingales, 1997; Lamont, Polk, and Saá-Requejo, 2001), Whited-Wu (2006) index, Hadlock-Pierce (2010) index, size (firm assets), payout ratio, and rating. *p*-values from one sided *t*-tests comparing coefficients on the interaction term between constrained and unconstrained firms are reported as well. Results in Panel A compare California and non-California plants based on geographically diversified firms. Panel B studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. The dependent variable is log (1+Emissions). The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. Control variables include PP&E, R&D Stock as well as plant and year fixed effects. Coefficients and their respective standard errors adjusted for clustering at the plant level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. *p*-values from one-sided *t*-tests comparing the coefficients on the interaction term (CalPlant x After in Panel A, and DivFirm x After in Panel B) between constrained and unconstrained firms are reported as well. The table also reports the number of observations and the adjusted R^2.

Panel A: California vs Non-California Plants (Geographically Diversified Firms)

	Dependent Variable: Log(1+Emissions)													
	Composite		Kaplan-Zingales		Hadlock-Pierce		Whited-Wu		Size		Payout		Rating	
	High	Low	High	Low	High	Low	High	Low	Small	Large	Low	High	Unrated	Rated
CalPlant x After	-0.353*	-0.030	-0.209	-0.128	-0.548**	-0.043	-0.229	-0.115	-0.440	-0.056	-0.463**	-0.122	-0.115	-0.078
	(0.208)	(0.079)	(0.151)	(0.121)	(0.263)	(0.086)	(0.379)	(0.076)	(0.322)	(0.080)	(0.230)	(0.085)	(0.169)	(0.098)
PP&E	-0.626	1.072	-0.743**	1.478	-2.712***	0.835	-1.584**	0.904	-1.162	0.217	-0.662	0.639	-2.228**	1.171*
	(0.434)	(1.092)	(0.366)	(1.384)	(0.746)	(0.534)	(0.633)	(0.899)	(0.775)	(0.589)	(0.479)	(0.860)	(0.997)	(0.637)
R&D stock	2.532	-8.065**	1.283	-15.322***	44.981**	-5.794*	-4.017*	-8.680*	-2.601	-8.534**	-0.031	-0.405	2.381	-8.408**
	(3.476)	(3.902)	(1.580)	(5.851)	(18.339)	(3.128)	(2.144)	(4.628)	(2.568)	(3.925)	(2.691)	(3.228)	(2.903)	(3.932)
CalPlant x After: Con<Uncon?		0.073		0.338		0.034		0.384		0.124		0.082		0.425
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	973	2,187	1,456	1,604	469	2,685	440	2,646	279	2,866	994	2,152	1,062	2,096
Adjusted R^2	0.904	0.827	0.921	0.816	0.903	0.864	0.835	0.876	0.881	0.858	0.910	0.865	0.907	0.841

(continued)

Table 5: Firm Financial Constraints and Plant Emission Responses: Separate Regressions by Constraint Groups (continued)

Panel B: Spillovers to Non-California Plants (Diversified vs Undiversified Firms)

	Dependent Variable: Log(1+Emissions)													
	Composite		Kaplan-Zingales		Hadlock-Pierce		Whited-Wu		Size		Payout		Rating	
	High	Low	High	Low	High	Low	High	Low	Small	Large	Low	High	Unrated	Rated
DivFirm x After	0.292***	-0.089*	0.278***	0.069	0.264***	-0.003	0.129	-0.006	0.260**	0.003	0.350***	0.064	0.237***	-0.015
	(0.065)	(0.046)	(0.067)	(0.054)	(0.078)	(0.044)	(0.139)	(0.040)	(0.123)	(0.042)	(0.063)	(0.045)	(0.068)	(0.048)
DivFirm	-0.414**	-0.010	0.189	-0.165	-0.317**	-0.072	0.076	-0.142	-0.300	-0.098	-0.272***	0.137	-0.274**	-0.005
	(0.210)	(0.121)	(0.170)	(0.164)	(0.161)	(0.122)	(0.160)	(0.127)	(0.215)	(0.114)	(0.102)	(0.128)	(0.124)	(0.133)
PP&E	-0.447***	0.382	-1.295***	-0.262	-0.292*	-0.152	-0.550***	0.519	-0.355***	-0.147	-0.566***	0.373	-0.459**	0.100
	(0.132)	(0.423)	(0.419)	(0.234)	(0.153)	(0.200)	(0.148)	(0.351)	(0.114)	(0.318)	(0.152)	(0.297)	(0.178)	(0.230)
R&D stock	-0.014	-1.089	-0.009	-2.038	0.004	0.042	-0.001	-0.035	-0.025	0.048	-0.024	-0.438	-0.001	-0.246
	(0.035)	(1.329)	(0.036)	(1.569)	(0.050)	(0.053)	(0.042)	(0.059)	(0.037)	(0.077)	(0.039)	(0.921)	(0.036)	(0.806)
DivFirm x After: Con>Uncon?		0.000		0.008		0.001		0.175		0.024		0.000		0.001
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,929	4,969	4,437	5,823	5,078	5,801	4,832	5,527	4,280	6,550	5,230	5,491	5,984	4,820
Adjusted R^2	0.718	0.779	0.712	0.749	0.700	0.778	0.673	0.819	0.689	0.774	0.712	0.803	0.717	0.774

Table 6: Firm Financial Constraints and Plant Emission Responses: Pooled Regressions with Constraint Dummies

The table reports results from pooled triple difference regressions. Results in Panel A compare California and non-California plants based on geographically diversified firms. Panel B studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. The dependent variable is log (1+Emissions). The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. The firm-level dummy Constrained is an indicator for whether a firm is financially constrained according to each financial constraint measure, i.e. alternatively our composite measure, the Kaplan-Zingales (KZ) index, Whited-Wu (WW) index, Hadlock-Pierce (HP) index, size, payout ratio, and rating. Control variables include PP&E, R&D Stock as well as plant and year fixed effects. Coefficients and their respective standard errors adjusted for clustering at the plant level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations and the adjusted R^2.

Panel A: California vs Non-California Plants (Geographically Diversified Firms)

	Dependent Variable: Log(1+Emissions)						
	Composite	KZ	HP	WW	Size	Payout	Rating
CalPlant x After x Const.	-0.391*	-0.156	-0.605**	-0.206	-0.607*	-0.253	-0.162
	(0.219)	(0.193)	(0.299)	(0.397)	(0.350)	(0.243)	(0.179)
CalPlant x After	-0.021	-0.072	-0.013	-0.095	-0.067	-0.067	-0.072
	(0.077)	(0.118)	(0.092)	(0.075)	(0.081)	(0.085)	(0.102)
CalPlant x Constrained	0.341	-0.605	0.523	1.905**	1.022	-1.256**	2.493***
	(0.922)	(0.515)	(1.146)	(0.855)	(0.907)	(0.599)	(0.855)
After x Constrained	0.107	0.059	-0.045	-0.144	0.150	-0.148*	0.064
	(0.072)	(0.065)	(0.100)	(0.163)	(0.127)	(0.075)	(0.081)
Constrained	-1.710**	-1.458***	-0.597	-1.607**	-2.278***	0.494	-1.970**
	(0.805)	(0.479)	(1.107)	(0.780)	(0.804)	(0.551)	(0.768)
PP&E	0.058	-0.319	-0.799	-0.521	0.034	-0.725	-0.470
	(0.543)	(0.518)	(0.592)	(0.623)	(0.521)	(0.645)	(0.647)
R&D stock	-5.056	-6.917**	-3.256	-5.851*	-5.973*	-1.119	-4.567
	(3.219)	(3.261)	(3.206)	(3.296)	(3.255)	(2.791)	(3.243)
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,162	3,062	3,162	3,091	3,147	3,162	3,162
Adjusted R^2	0.855	0.859	0.851	0.853	0.857	0.853	0.853

(continued)

Table 6: Firm Financial Constraints and Plant Emission Responses: Pooled Regressions with Constraint Dummies (continued)

Panel B: Spillovers to Non-California Plants (Diversified vs Undiversified Firms)

	Dependent Variable: Log(1+Emissions)						
	Composite	KZ	HP	WW	Size	Payout	Rating
DivFirm x After x Const.	0.342***	0.221**	0.189*	0.166	0.228*	0.114	0.232***
	(0.080)	(0.087)	(0.103)	(0.143)	(0.131)	(0.089)	(0.084)
DivFirm x After	-0.056	0.071	0.010	0.020	0.025	0.084*	0.009
	(0.045)	(0.055)	(0.044)	(0.039)	(0.043)	(0.047)	(0.047)
DivFirm x Constrained	-0.696***	-0.121	-0.330	-0.234	-0.461*	-0.034	-0.502**
	(0.250)	(0.240)	(0.297)	(0.245)	(0.279)	(0.236)	(0.225)
After x Constrained	-0.359***	-0.267***	-0.289***	-0.295***	-0.302***	-0.234***	-0.242***
	(0.053)	(0.070)	(0.055)	(0.058)	(0.060)	(0.058)	(0.053)
DivFirm	0.060	-0.072	-0.032	-0.028	-0.041	-0.088	0.032
	(0.122)	(0.161)	(0.124)	(0.129)	(0.119)	(0.148)	(0.131)
Constrained	0.280**	-0.072	0.141	0.088	-0.001	0.331*	0.217*
	(0.125)	(0.076)	(0.086)	(0.149)	(0.126)	(0.170)	(0.123)
PP&E	-0.428***	-0.503***	-0.424***	-0.372***	-0.422***	-0.489***	-0.343**
	(0.120)	(0.140)	(0.123)	(0.121)	(0.126)	(0.133)	(0.161)
R&D stock	-0.014	0.016	0.002	-0.004	-0.013	-0.017	-0.017
	(0.036)	(0.036)	(0.036)	(0.035)	(0.039)	(0.038)	(0.040)
Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,993	10,373	10,987	10,468	10,938	10,775	10,993
Adjusted R^2	0.730	0.728	0.729	0.726	0.730	0.728	0.729

Table 7: Plant Ownership Responses to California Cap-and-Trade Rule

The table presents results from linear probability and multinomial logit regressions of plant closure and opening decisions as well as firm level plant ownership regressions. In linear probability models, the dependent variables are indicators for whether a firm closes (opens) a plant or not. In multinomial logits, the dependent variable is a categorical variable equal to –1 for plant closure, 0 for no ownership change, and +1 for plant opening. Panel A compares California and non-California plants based on geographically diversified firms. Panel B studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. Definitions of explanatory variables are as in Table 6. For multinomial logits, coefficients report marginal effects with respect to discrete changes from 0 to 1 for indicator variables, and unit changes at the means for continuous variables. Standard errors are adjusted for clustering at the plant level for firm-plant level regressions and at the firm-California operation level for firm level regressions. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations, adjusted R^2, and pseudo R^2.

Panel A: California vs Non-California Plants (Geographically Diversified Firms)

	Linear probability		Multinomial logits (base: no change)		Firm level # of plants
	Close	Open	Close	Open	
CalPlant x After x Const.	0.145***	-0.002	0.181	-0.080**	-0.207
	(0.055)	(0.070)	(0.129)	(0.032)	(1.083)
CalPlant x After	-0.036	0.036	-0.014	0.042	-1.209
	(0.036)	(0.044)	(0.024)	(0.045)	(0.898)
CalPlant x Constrained	-0.159	-0.140	0.001	0.008	3.406*
	(0.099)	(0.093)	(0.030)	(0.047)	(1.723)
After x Constrained	-0.053***	0.061**	-0.029	0.135***	0.081
	(0.020)	(0.029)	(0.019)	(0.046)	(0.704)
CalPlant			0.018	0.013	-7.861***
			(0.017)	(0.026)	(1.501)
After			-0.076***	-0.081***	
			(0.015)	(0.018)	
Constrained	0.093	0.015	-0.032***	-0.052***	
	(0.072)	(0.080)	(0.012)	(0.019)	
PP&E	-0.310***	0.185	0.072**	0.229***	-4.113
	(0.112)	(0.151)	(0.034)	(0.032)	(4.692)
R&D stock	-0.629	4.952***	-0.376	0.139**	3.976
	(0.640)	(1.500)	(0.267)	(0.055)	(4.702)
Plant FE	Yes	Yes	No	No	No
Firm FE	No	No	No	No	Yes
Year FE	Yes	Yes	No	No	Yes
Observations	2,518	2,735	3,062	3,062	528
Adjusted R^2	0.381	0.132			0.571
Pseudo R^2			0.050	0.050	

(continued)

Table 7: Plant Ownership Responses to California Cap-and-Trade Rule (continued)

Panel B: Spillovers to Non-California Plants (Diversified vs Undiversified Firms)

	Linear probability		Multinomial logits (base: no change)		Firm level # of plants
	Close	Open	Close	Open	
DivFirm x After x Const.	-0.063***	0.125***	-0.058***	0.182***	0.817
	(0.023)	(0.033)	(0.016)	(0.064)	(0.970)
DivFirm x After	-0.005	-0.010	-0.023	0.006	-0.521
	(0.016)	(0.021)	(0.017)	(0.022)	(0.727)
DivFirm x Constrained	-0.044	0.007	-0.063***	-0.058***	-5.395**
	(0.039)	(0.062)	(0.010)	(0.017)	(2.142)
After x Constrained	0.046***	-0.066***	0.038**	-0.016	0.239
	(0.013)	(0.016)	(0.018)	(0.016)	(0.492)
DivFirm	0.037**	-0.082***	0.061***	-0.016	5.738***
	(0.019)	(0.027)	(0.017)	(0.012)	(1.903)
After			-0.065***	-0.089***	
			(0.016)	(0.015)	
Constrained	0.080***	-0.116***	0.054***	0.009	
	(0.029)	(0.035)	(0.010)	(0.010)	
PP&E	0.131***	0.008	0.094***	0.260***	0.100
	(0.046)	(0.045)	(0.019)	(0.022)	(2.191)
R&D stock	0.859***	0.090***	0.128***	0.122**	0.131
	(0.054)	(0.018)	(0.041)	(0.059)	(0.092)
Plant FE	Yes	Yes	No	No	No
Firm FE	No	No	No	No	Yes
Year FE	Yes	Yes	No	No	Yes
Observations	8,353	9,524	10,053	10,053	1,905
Adjusted R^2	0.406	0.244			0.884
Pseudo R^2			0.054	0.054	

Table 8: Firm Level Total Emissions

The table presents results from firm level regressions testing whether firms affected by the California cap-and-trade rule increase their overall emissions, and whether financial constraints affect their overall responses. Columns (1) and (2) test whether a control group of firms without any plants in California change their overall emissions after the implementation of the cap-and-trade rule. Columns (3) and (4) test whether a treatment group of firms with plants in California change their total emissions. Columns (5) and (6) pool the control and treatment firms together and include a treatment dummy (Treat) to test the difference in emission changes across the two groups. Columns (7)-(10) repeat the previous four regressions with a narrower definition of treatment that firms have plants both in California and also in other states. The dependent variable is log(1+firm total emissions), where firm total emissions are computed by summing up emissions across all plants owned by a firm in a given year. After is an indicator equal to 1 if the time period is 2013 or onward and 0 otherwise. Constrained is an indicator for whether a firm is financially constrained according to our composite measure. Treat is a dummy equal to 1 if a firm owns a plant in California (or in California as well as in other states) in a given year, and 0 otherwise. Control variables include PP&E, R&D Stock as well as firm and year fixed effects. Coefficients and standard errors adjusted for clustering at the firm level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations and adjusted R^2.

	Dependent Variable: Log(1+Firm total emissions)									
	Control definition: Firms without plant in CA		Treatment definition: Firms with plants in CA				Treatment definition: Firms with plants in CA and other states			
	Control group only		Treatment group only		Treatment + Control		Treatment group only		Treatment + Control	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
After x Constrained	0.123	0.141	0.106	0.111	0.146	0.165	0.256**	0.262**	0.145	0.164
	(0.105)	(0.105)	(0.130)	(0.132)	(0.104)	(0.105)	(0.123)	(0.123)	(0.104)	(0.105)
Treat x After x Const.					0.044	0.050			0.137	0.151
					(0.184)	(0.182)			(0.211)	(0.207)
After	0.080		-0.055		0.075		-0.066		0.075	
	(0.094)		(0.089)		(0.093)		(0.094)		(0.093)	
Treat x After					-0.168	-0.154			-0.182	-0.175
					(0.129)	(0.128)			(0.131)	(0.131)
Treat					1.081***	1.005***			1.090***	1.017***
					(0.225)	(0.239)			(0.227)	(0.241)
Treat x Constrained					0.198	0.167			0.074	0.038
					(0.913)	(0.865)			(0.891)	(0.841)
PP&E	-0.188	0.113	-1.423*	-1.340*	-0.338	-0.045	-1.084	-1.016	-0.307	-0.014
	(0.379)	(0.378)	(0.718)	(0.704)	(0.344)	(0.342)	(0.722)	(0.699)	(0.345)	(0.342)
R&D stock	0.060	0.069*	0.614	0.640	0.061	0.069*	1.068	1.379	0.061	0.069*
	(0.049)	(0.040)	(0.774)	(0.754)	(0.050)	(0.041)	(1.975)	(1.970)	(0.050)	(0.041)
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	No	Yes	No	Yes	No	Yes	No	Yes	No	Yes
Observations	1,632	1,632	307	307	1,958	1,958	254	254	1,905	1,905
Adjusted R^2	0.849	0.856	0.977	0.976	0.871	0.877	0.971	0.971	0.867	0.874

Table 9: Impact on Sectoral GDP and Employment

The table examines whether the California cap-and-trade rule differentially impacts GDP and employment in affected industries in California compared to other states, and whether growth from other industries countervails this effect. A plant's industry is defined as the narrowest NAICS code with at least 50 plants in the entire cross-section each year, and mapped to the narrowest available 2-4 digit NAICS industry classification for which the BEA publicly reports state level GDP and employment. The data is collapsed to state-sector-year level where sectors are categorized as either "emission sector" or "non-emission sector". All BEA industries with greenhouse gas emitting plants are pooled together to comprise the emission sector, and all remaining industries are grouped as the non-emission sector. GDP (inflation adjusted with respect to 2009 dollars) and employment (number of wage earning workers) are aggregated up to each state-sector-year. Columns (1)-(3) report results with log(1+GDP) as the dependent variable, and columns (4)-(6) use log(1+Wage employment) as the dependent variable. For each dependent variable, separate regressions are run for the emission sector and non-emission sector, and then the two sectors are included together in a pooled regression. Cal is a state level dummy indicating whether the state is California or not, and After is an indicator for whether the year is 2013 or later. EmissionSector indicates whether the sector is comprised of industries with greenhouse gas emitting plants. State and year fixed effects are controlled for. Standard errors are adjusted for clustering at the state level. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations and adjusted R^2.

Dependent Variable	log(1+GDP)			log(1+Wage employment)		
	Emission sector	Non-emission sector	Pooled	Emission sector	Non-emission sector	Pooled
	(1)	(2)	(3)	(4)	(5)	(6)
Cal x After	-0.013	0.069**	0.038	-0.120*	0.104***	0.091***
	(0.039)	(0.026)	(0.042)	(0.069)	(0.007)	(0.016)
Cal x After x EmissionSector			-0.034			-0.197**
			(0.067)			(0.074)
Cal x EmissionSector			0.693***			1.224***
			(0.231)			(0.249)
After x EmissionSector			-0.127*			-0.064
			(0.067)			(0.074)
EmissionSector			-1.036***			-2.215***
			(0.231)			(0.249)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	299	287	588	299	288	589
Adjusted R^2	0.990	0.953	0.846	0.953	0.997	0.857

Table 10: Do Emissions and Plant Ownership Chase Growth Opportunities?

The table examines whether changes in emissions and plant ownership after the implementation of the California cap-and-trade rule are explained by variations in growth opportunities associated with plants. We employ two measures of growth opportunities: (1) annual private industry real GDP growth of the state the plant is located in, and (2) median Tobin's Q of firms that own a plant in the same state and industry as the plant and primarily operate in that industry. Panel A reports the population-weighted cross-state average real GDP growth and median Tobin's Q (first averaged within states) over our sample period from 2010 to 2015. The averages for the Before (2010–2012) and After (2013–2015) periods are shown, as well as the difference between the two and its corresponding *t*-statistic. State level GDP data is downloaded from the Bureau of Economic Analysis. Panel B compares emissions and ownership for California and non-California plants based on geographically diversified firms, controlling for GDP growth and Tobin's Q. The first three columns use log (1+Emissions) as the dependent variable. The first two columns each have either GDP growth or Tobin's Q as its explanatory variable as well as its interaction with the firm level Constrained dummy based on our composite constraint measure. The third column includes all growth opportunity variables and adds the main variables as in Table 7 and 8: CalPlant (equal to 1 if the plant is located in California and 0 otherwise), After (equal to 1 if the time period is 2013 or onward and 0 otherwise), Constrained (indicator for whether a firm is financially constrained according to our composite measure), and their interaction terms. In the next six columns, the dependent variable is a categorical variable equal to –1 for plant closure, 0 for no ownership change, and +1 for plant opening. The explanatory variables and set of results are analogous to the first three columns. Panel C studies spillovers to non-California plants comparing geographically diversified and non-diversified firms. The sample is restricted to plants located outside of California, and the variable CalPlant is replaced with DivFirm (indicates whether a firm owns plants both in California and in other states during a given year or not). GDP growth and Tobin's Q are further interacted with DivFirm x Constrained and Divfirm. Control variables include PP&E, R&D Stock as well as plant and year fixed effects. For multinomial logits, fixed effects are dropped and coefficients report marginal effects with respect to discrete changes from 0 to 1 for indicator variables, and unit changes at the means for continuous variables. Standard errors are adjusted for clustering at the plant level. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations, adjusted R^2, and pseudo R^2.

Panel A: Growth Opportunities in California and Other States

State	2010	2011	2012	2013	2014	2015	Before (2010-2012)	After (2013-2015)	After–Before	t-stat.
State GDP growth (%)										
California	1.60	1.50	3.10	2.90	4.40	4.90	2.07	4.07	2.00	2.52
Other States	2.70	2.01	2.43	1.99	2.68	2.79	2.38	2.49	0.11	0.34
Diff	-1.10	-0.51	0.67	0.91	1.72	2.11	**-0.31**	**1.58**	**1.89**	**3.00**
Median Tobin's Q										
California	1.29	1.36	1.31	1.34	1.42	1.38	1.32	1.38	0.06	1.94
Other States	1.34	1.41	1.34	1.35	1.43	1.43	1.36	1.40	0.04	1.04
Diff	-0.05	-0.05	-0.03	0.00	0.00	-0.06	**-0.04**	**-0.02**	**0.02**	**1.12**

(continued)

Table 10: Do Emissions and Plant Ownership Chase Growth Opportunities? (continued)

Panel B: California vs Non-California Plants (Geographically Diversified Firms)

	Dependent Variable: Log(1+Emissions)			Multinomial: -1=Close, 0=No change (base outcome), 1=Open					
				Close	Open	Close	Open	Close	Open
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
%ΔGDP x Constrained	-0.017		-0.009	0.005*	-0.007			0.001	-0.009
	(0.013)		(0.011)	(0.003)	(0.006)			(0.002)	(0.006)
%ΔGDP	-0.005		-0.004	-0.004**	-0.004*			-0.001	-0.004*
	(0.007)		(0.007)	(0.001)	(0.002)			(0.001)	(0.002)
Q x Constrained		-0.118	-0.134			0.171***	0.083*	0.186***	0.103**
		(0.166)	(0.160)			(0.050)	(0.047)	(0.044)	(0.049)
Q		0.084	0.092			-0.181***	0.011	-0.172***	0.007
		(0.087)	(0.091)			(0.036)	(0.031)	(0.032)	(0.031)
CalPlant x After x Const.			-0.369*					0.171	-0.070**
			(0.216)					(0.125)	(0.035)
CalPlant x After			-0.011					-0.014	0.056
			(0.078)					(0.022)	(0.047)
CalPlant x Constrained			0.325					-0.006	-0.012
			(0.913)					(0.025)	(0.040)
After x Constrained			0.110					-0.025	0.122***
			(0.073)					(0.019)	(0.046)
CalPlant								0.024	0.016
								(0.017)	(0.026)
After								-0.075***	-0.086***
								(0.014)	(0.017)
Constrained	-1.522***	-1.428***	-1.515*	-0.044***	0.018	-0.188***	-0.094**	-0.202***	-0.119***
	(0.452)	(0.509)	(0.818)	(0.012)	(0.020)	(0.045)	(0.044)	(0.044)	(0.040)
PP&E	0.020	0.012	0.074	0.063	0.219***	0.017	0.227***	0.040	0.263***
	(0.482)	(0.499)	(0.549)	(0.040)	(0.031)	(0.040)	(0.032)	(0.032)	(0.034)
R&D stock	-4.558	-4.470	-4.833	-0.479	0.160***	-0.264	0.132**	-0.203	0.135***
	(3.134)	(3.200)	(3.259)	(0.330)	(0.058)	(0.279)	(0.054)	(0.218)	(0.052)
Plant FE	Yes	Yes	Yes	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	No	No	No	No	No	No
Observations	3,162	3,162	3,162	3,062	3,062	3,057	3,057	3,057	3,057
Adjusted R^2	0.855	0.855	0.855						
Pseudo R^2				0.024	0.024	0.032	0.032	0.067	0.067

(continued)

Table 10: Do Emissions and Plant Ownership Chase Growth Opportunities? (continued)

Panel C: Spillovers to Non-California Plants (Diversified vs Undiversified Firms)

	Dependent Variable: Log(1+Emissions)			Multinomial: -1=Close, 0=No change (base outcome), 1=Open					
				Close	Open	Close	Open	Close	Open
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
%ΔGDP x DivFirm x Const.	-0.049**		-0.023	-0.023***	-0.007			-0.003	-0.019***
	(0.020)		(0.018)	(0.005)	(0.005)			(0.004)	(0.007)
%ΔGDP x DivFirm	0.026**		0.022*	0.007**	-0.000			0.002	0.002
	(0.011)		(0.013)	(0.003)	(0.003)			(0.003)	(0.003)
%ΔGDP x Constrained	0.001		-0.002	0.008***	0.006**			0.003	0.009***
	(0.011)		(0.012)	(0.003)	(0.003)			(0.002)	(0.003)
%ΔGDP	-0.000		0.001	-0.006***	-0.006***			-0.002	-0.007***
	(0.008)		(0.009)	(0.002)	(0.002)			(0.002)	(0.002)
Q x DivFirm x Constrained		-0.367**	-0.480**			-0.097***	0.011	0.257***	-0.068
		(0.168)	(0.213)			(0.015)	(0.013)	(0.084)	(0.066)
Q x DivFirm		0.044	0.285**			0.032***	-0.013*	-0.223***	0.020
		(0.088)	(0.142)			(0.008)	(0.007)	(0.063)	(0.039)
Q x Constrained		0.485***	0.513***			0.124***	0.098***	0.034	0.097***
		(0.105)	(0.116)			(0.031)	(0.025)	(0.030)	(0.029)
Q		-0.024	-0.136*			-0.140***	0.040*	-0.053*	0.026
		(0.066)	(0.075)			(0.030)	(0.022)	(0.030)	(0.025)
DivFirm x After x Const.			0.303***					-0.053***	0.154**
			(0.083)					(0.019)	(0.062)
DivFirm x After			-0.070					-0.029*	0.002
			(0.045)					(0.015)	(0.021)
DivFirm x Constrained			0.003					-0.123***	0.059
			(0.371)					(0.011)	(0.107)
After x Constrained			-0.336***					0.037**	-0.007
			(0.052)					(0.017)	(0.016)
DivFirm			-0.332					0.652***	-0.097***
			(0.240)					(0.165)	(0.034)
After								-0.063***	-0.088***
								(0.015)	(0.014)
Constrained	0.059	-0.544***	-0.398*	0.040***	-0.003	-0.084**	-0.132***	0.006	-0.159***
	(0.138)	(0.185)	(0.204)	(0.008)	(0.009)	(0.043)	(0.037)	(0.041)	(0.045)
PP&E	-0.439***	-0.394***	-0.377***	0.105***	0.262***	0.080***	0.270***	0.078***	0.276***
	(0.120)	(0.120)	(0.120)	(0.019)	(0.021)	(0.021)	(0.024)	(0.020)	(0.024)
R&D stock	0.020	-0.039	-0.058	0.150***	0.154***	0.166***	0.098	0.144***	0.082
	(0.035)	(0.052)	(0.050)	(0.042)	(0.056)	(0.057)	(0.076)	(0.052)	(0.072)
Plant FE	Yes	Yes	Yes	No	No	No	No	No	No
Year FE	Yes	Yes	Yes	No	No	No	No	No	No
Observations	10,993	10,892	10,892	10,053	10,053	9,954	9,954	9,954	9,954
Adjusted R^2	0.728	0.728	0.730						
Pseudo R^2				0.033	0.033	0.042	0.042	0.065	0.065

Table 11: Placebo Tests

The table reports results from placebo tests using Texas (Panel A) and Louisiana (Panel B) as placebo treatment states. In each panel, the first two columns use log (1+Emissions) as the dependent variable. The first column compares California and non-California plants based on geographically diversified firms. The next column studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. The next four columns show results from multinomial logits, where the dependent variable is a categorical variable equal to –1 for plant closure, 0 for no ownership change, and +1 for plant opening. The third and fourth columns compare California and non-California plants based on geographically diversified firms, and the last two columns test for spillovers to non-California Plants comparing geographically diversified and non-diversified firms. TreatPlant equals to 1 if the plant is located in the placebo state and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. DivFirm indicates whether a firm owns plants both in the placebo state and in other states during a given year or not. The firm-level dummy Constrained is an indicator for whether a firm is financially constrained according to our composite measure. Control variables include PP&E, R&D Stock as well as plant and year fixed effects. For multinomial logits, fixed effects are dropped and coefficients report marginal effects with respect to discrete changes from 0 to 1 for indicator variables, and unit changes at the means for continuous variables. Standard errors are adjusted for clustering at the plant level. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations, adjusted R^2, and pseudo R^2.

Panel A: Placebo State 1 - Texas

	Dependent Variable: Log(1+Emissions)		Multinomial: -1=Close, 0=No change (base outcome), 1=Open			
	Treat vs Control (Diversified firms)	Spillovers to Control Plants (Div vs undiv firms)	Treat vs Control (Diversified firms)		Spillovers to Control Plants (Div. vs undiv. firms)	
			Close	Open	Close	Open
TreatPlant x After x Const.	0.022		0.063	0.011		
	(0.146)		(0.051)	(0.040)		
TreatPlant x After	-0.133*		-0.030	-0.025		
	(0.071)		(0.023)	(0.025)		
TreatPlant x Constrained	-0.539		0.035	0.017		
	(0.349)		(0.032)	(0.027)		
TreatPlant			-0.019	0.038**		
			(0.018)	(0.018)		
DivFirm x After x Const.		-0.137			0.006	0.059
		(0.091)			(0.031)	(0.045)
DivFirm x After		0.086*			-0.004	0.010
		(0.051)			(0.025)	(0.030)
DivFirm x Constrained		-0.066			0.033	-0.048***
		(0.124)			(0.024)	(0.018)
DivFirm		0.403***			-0.003	0.035**
		(0.130)			(0.017)	(0.018)
After			-0.070***	-0.088***	-0.067***	-0.083***
			(0.016)	(0.018)	(0.024)	(0.029)
After x Constrained	-0.319***	-0.186***	0.062**	0.020	0.021	-0.022
	(0.067)	(0.064)	(0.025)	(0.025)	(0.026)	(0.028)
Constrained	0.358**	0.454***	0.016	-0.003	0.007	0.040**
	(0.180)	(0.156)	(0.014)	(0.016)	(0.016)	(0.017)
PP&E	-0.402*	-0.529***	0.164***	0.255***	0.144***	0.230***
	(0.234)	(0.128)	(0.026)	(0.030)	(0.022)	(0.023)
R&D stock	-0.067	-0.016	0.351***	-0.043	0.159***	0.149***
	(0.179)	(0.037)	(0.042)	(0.085)	(0.042)	(0.054)
Plant FE	Yes	Yes	No	No	No	No
Year FE	Yes	Yes	No	No	No	No
Observations	6,420	8,788	6,042	6,042	7,947	7,947
Adjusted R^2	0.728	0.748				
Pseudo R^2			0.063	0.063	0.054	0.054

(continued)

Table 11: Placebo Tests (continued)

Panel B: Placebo State 2 - Louisiana

	Dependent Variable: Log(1+Emissions)		Multinomial: -1=Close, 0=No change (base outcome), 1=Open			
	Treat vs Control (Diversified firms)	Spillovers to Control Plants (Div vs undiv firms)	Treat vs Control (Diversified firms)		Spillovers to Control Plants (Div. vs undiv. firms)	
			Close	Open	Close	Open
TreatPlant x After x Const.	0.019		0.089	0.061		
	(0.155)		(0.076)	(0.072)		
TreatPlant x After	-0.062		-0.043*	-0.076***		
	(0.122)		(0.025)	(0.028)		
TreatPlant x Constrained	0.476		-0.012	-0.039		
	(0.374)		(0.029)	(0.030)		
TreatPlant			0.014	0.032		
			(0.023)	(0.026)		
DivFirm x After x Const.		-0.001			-0.025	0.003
		(0.097)			(0.021)	(0.030)
DivFirm x After		-0.015			-0.016	0.076***
		(0.051)			(0.021)	(0.029)
DivFirm x Constrained		0.542***			0.002	-0.018
		(0.175)			(0.019)	(0.018)
DivFirm		-0.213			0.047***	0.017
		(0.131)			(0.017)	(0.014)
After			-0.089***	-0.051***	-0.064***	-0.124***
			(0.018)	(0.019)	(0.020)	(0.021)
After x Constrained	-0.296***	-0.312***	0.092***	-0.022	0.048**	0.007
	(0.080)	(0.060)	(0.029)	(0.023)	(0.023)	(0.022)
Constrained	0.274	-0.248*	-0.009	0.026	0.038***	0.021*
	(0.224)	(0.150)	(0.015)	(0.018)	(0.013)	(0.012)
PP&E	-0.096	-0.537***	0.108***	0.157***	0.096***	0.242***
	(0.266)	(0.138)	(0.029)	(0.038)	(0.017)	(0.018)
R&D stock	0.136	-0.020	0.479***	-0.394**	0.127***	0.142***
	(0.191)	(0.038)	(0.040)	(0.179)	(0.027)	(0.037)
Plant FE	Yes	Yes	No	No	No	No
Year FE	Yes	Yes	No	No	No	No
Observations	4,648	9,943	4,460	4,460	9,071	9,071
Adjusted R^2	0.749	0.727				
Pseudo R^2			0.061	0.061	0.055	0.055

Appendix A: Variable Names and Definitions

The table shows the names, definitions, and data sources of the variables used in the study.

Variable Name	Definition	Source
Emissions	Facility greenhouse gas emissions quantity by firm (metric tons × firm ownership in facility)	EPA
Close	Indicator equal to 1 if firm reduces fractional ownership in plant or ceases ownership in plant, and 0 otherwise	EPA
Open	Indicator equal to 1 if firm increases fractional ownership in plant or begins ownership in plant, and 0 otherwise	EPA
Multinomial close/open	Categorical variable equal to -1 if plant closed or ownership is reduced, 0 if there is no plant ownership change, 1 if plant opened or ownership is increased	EPA
CalPlant	Indicator equal to 1 if the plant is located in California, and 0 otherwise	EPA
TreatPlant	Indicator equal to 1 if the plant is located in placebo state, and 0 otherwise	EPA
DivFirm	Indicator equal to 1 if firm owns plants both in California and in other states, and 0 otherwise	EPA
After	Indicator equal to 1 if the time period is 2013 or onward, and 0 otherwise	
Total assets	Assets in $ billions (AT)	Compustat
Size	Log of total assets	Compustat
PP&E	Property, plant and equipment (gross)/Total assets (PPEGT/AT)	Compustat
Capital expenditures	Capital expenditures/Total assets (CAPX/AT)	Compustat
Short-term debt	Debt in current liabilities/Total assets (DLC/AT)	Compustat
Long-term debt	Long-term debt/Total assets (DLTT/AT)	Compustat
Cash	Cash and short-term investments/Total assets (CHE/AT)	Compustat
Cash flow	Operating income before depreciation/Total assets (OIBDP/AT)	Compustat
Profitability	Income before extraordinary items/Total assets (IB/AT)	Compustat
Tobin's Q	Market value of assets (Total assets (AT) + Market value of common equity (CSHO*$PRCC_F$) - Common equity (CEQ) - Deferred taxes (TXDB)) divided by 0.9·Book value of assets (AT)+0.1·Market value of assets	Compustat
R&D	Research and development expense/sales (XRD/SALE)	Compustat
R&D Stock	Perpetual inventory method with initial value of R&D capital stock set as zero and accumulating R&D expenses with a depreciation rate of 15%, scaled by total assets	Hall, Jaffe, and Trajtenberg (2005)
Payout Ratio	(Cash dividends + repurchases)/Income before extraordinary items ((DVP+DVC+PRSTKC)/ IB)	Compustat
Firm Age	Difference between observation year and founding year (annual, years)	Jovanovic and Rousseau (2001)
Long-term rating	Indicator equal to 1 if firm has rating on long-term (>1 year) obligations, and 0 otherwise	Compustat
Short-term rating	Indicator equal to 1 if firm has rating on short-term (<1 year) obligations, and 0 otherwise	Compustat
Kaplan-Zingales Index	-1.002·Cash flow + 0.283·Tobin's Q + 3.139·Total debt - 39.368·Dividends - 1.315·Cash	Kaplan and Zingales (1997); Lamont, Polk, and Saá-Requejo (2001)
Whited-Wu Index	-0.091·Cash flow - 0.062·Positive dividend dummy + 0.021·Long-term debt - 0.044·Size + 0.102·Industry sales growth - 0.035·Sales growth	Whited and Wu (2006)
Hadlock-Pierce Index	-0.737·Size + 0.043·$Size^2$ - 0.040·Age where Size is the log of Min(AT, $4.5 billion) and Age is Min(Firm age, 37 years)	Hadlock and Pierce (2010)
Financial constraints	For Kaplan-Zingales, Hadlock-Pierce, and Whited-Wu, size, and payout, firms are assigned percentile rankings based on each measure every year. Using six years strictly before the sample period (i.e. fiscal years 2003-2008) time-series average percentile rankings are computed for each firm and each measure. Based on average rankings, firms are categorized as constrained if they are above median for Kaplan-Zingales, Hadlock-Pierce, and Whited-Wu, and if they are below median for size and payout. For credit ratings, a firm is categorized as "long-term (short-term)" financially constrained if the firm did not have a bond (commercial paper) rating as of the most recent year of the 2003-2008 pre-sample period but had on average positive long-term (short-term) debt during this period. If the firm did have a bond (commercial paper) rating as of the most recent year of the six-year pre-sample period or had on average zero long-term (short-term) debt during this period, then the firm is "long-term (short-term)" unconstrained. If a firm is either long-term or short-term credit constrained, the firm is classified as constrained based on ratings and unconstrained otherwise.	Compustat
Composite	Indicator equal to 1 if firm is constrained according to majority of all six constraint measures, and 0 otherwise	Compustat

Table A.1: Plant Emission Responses to California Cap-and-Trade Rule

The table presents results from the firm-plant level difference-in-difference (DID) regressions. Results in the first four columns compare California and non-California plants based on all firms (geographically diversified firms). The last two columns study spillovers to non-California Plans comparing geographically diversified and non-diversified firms. The dependent variable is log (1+Emissions). The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. Control variables include PP&E, R&D Stock as well as firm x plant and year fixed effects. Coefficients and their respective standard errors adjusted for clustering at the plant level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations and the adjusted R^2.

Dependent Variable: Log(1+Emissions)						
	California vs Non-California Plants (All Firms)		California vs Non-California Plants (Geographically Diversified Firms)		Spillovers to Non-California Plants (Diversified vs Undiversified Firms)	
CalPlant x After	-0.061	-0.061	-0.176*	-0.171*		
	(0.089)	(0.090)	(0.095)	(0.096)		
DivFirm x After					0.169***	0.158***
					(0.035)	(0.035)
DivFirm					0.054	0.061
					(0.070)	(0.070)
PP&E		-0.405***		-0.099		-0.439***
		(0.097)		(0.075)		(0.102)
R&D stock		0.007		1.671		0.004
		(0.015)		(1.220)		(0.016)
Firm x Plant FE	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	12,784	12,773	3,825	3,825	11,996	11,985
Adjusted R^2	0.790	0.791	0.923	0.923	0.780	0.780

Table A.2: Firm Financial Constraints and Plant Emission Responses: Separate Regressions by Constraint Groups

The table presents results from DID regressions, separately for subsamples of financially constrained and unconstrained firms. A number of measures for financial constraints are used: Our composite measure, the Kaplan-Zingales index (following Kaplan and Zingales, 1997; Lamont, Polk, and Saá-Requejo, 2001), Whited-Wu (2006) index, Hadlock-Pierce (2010) index, size (firm assets), payout ratio, and rating. *p*-values from one sided *t*-tests comparing coefficients on the interaction term between constrained and unconstrained firms are reported as well. Results in Panel A compare California and non-California plants based on geographically diversified firms. Panel B studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. The dependent variable is log (1+Emissions). The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. Control variables include PP&E, R&D Stock as well as firm x plant and year fixed effects. Coefficients and their respective standard errors adjusted for clustering at the plant level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. *p*-values from one-sided *t*-tests comparing the coefficients on the interaction term (CalPlant x After in Panel A, and DivFirm x After in Panel B) between constrained and unconstrained firms are reported as well. The table also reports the number of observations and the adjusted R^2.

Panel A: California vs Non-California Plants (Geographically Diversified Firms)

	Dependent Variable: Log(1+Emissions)													
	Composite		Kaplan-Zingales		Hadlock-Pierce		Whited-Wu		Size		Payout		Rating	
	High	Low	High	Low	High	Low	High	Low	Small	Large	Low	High	Unrated	Rated
CalPlant x After	-0.387*	-0.048	-0.209	-0.138	-0.410	-0.094	-0.537	-0.059	-0.545	-0.079	-0.458**	-0.025	-0.304*	-0.081
	(0.216)	(0.068)	(0.155)	(0.108)	(0.281)	(0.082)	(0.373)	(0.062)	(0.346)	(0.078)	(0.229)	(0.067)	(0.169)	(0.097)
PP&E	-0.550	0.195	-0.501	0.858	-0.683	-0.067	-0.789	0.296	-0.998	-0.187	-0.409	-0.074	-0.241	-0.019
	(0.439)	(0.416)	(0.408)	(0.701)	(0.502)	(0.365)	(0.507)	(0.388)	(0.794)	(0.358)	(0.550)	(0.270)	(0.426)	(0.425)
R&D stock	2.359	2.427*	1.287	1.963	24.337	2.320**	-4.800**	4.281***	-3.200	1.583	0.422	2.742**	1.970	2.201
	(3.499)	(1.275)	(1.568)	(1.767)	(17.982)	(1.179)	(2.282)	(1.517)	(2.840)	(1.188)	(2.796)	(1.294)	(2.415)	(1.431)
CalPlant x After: Con<Uncon?		0.067		0.354		0.140		0.103		0.094		0.035		0.126
Firm x Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	967	2,150	1,451	1,566	466	2,651	436	2,610	273	2,829	993	2,124	1,055	2,062
Adjusted R^2	0.904	0.940	0.921	0.933	0.912	0.930	0.851	0.950	0.882	0.931	0.911	0.937	0.941	0.920

(continued)

Table A.2: Firm Financial Constraints and Plant Emission Responses: Separate Regressions by Constraint Groups (continued)

Panel B: Spillovers to Non-California Plants (Diversified vs Undiversified Firms)

	Dependent Variable: Log(1+Emissions)													
	Composite		Kaplan-Zingales		Hadlock-Pierce		Whited-Wu		Size		Payout		Rating	
	High	Low	High	Low	High	Low	High	Low	Small	Large	Low	High	Unrated	Rated
DivFirm x After	0.290***	-0.008	0.271***	0.117***	0.258***	0.056	0.135	0.050	0.227*	0.058	0.341***	0.038	0.259***	0.053
	(0.066)	(0.038)	(0.067)	(0.043)	(0.079)	(0.040)	(0.142)	(0.035)	(0.126)	(0.038)	(0.064)	(0.038)	(0.054)	(0.044)
DivFirm	-0.312	0.177**	0.146	0.049	-0.330*	0.149**	0.068	0.166**	-0.316	0.150**	-0.110	0.142	-0.191	0.165**
	(0.222)	(0.074)	(0.164)	(0.069)	(0.183)	(0.076)	(0.177)	(0.084)	(0.248)	(0.073)	(0.083)	(0.105)	(0.156)	(0.080)
PP&E	-0.590***	-0.035	-1.453***	-0.110	-0.475*	-0.532**	-0.687***	0.051	-0.593**	-0.343	-0.916***	-0.120	-0.720**	-0.274
	(0.223)	(0.268)	(0.443)	(0.163)	(0.245)	(0.256)	(0.262)	(0.236)	(0.250)	(0.277)	(0.264)	(0.253)	(0.304)	(0.190)
R&D stock	0.001	-0.603	-0.008	-1.326	0.002	0.042	0.032	-0.032	-0.019	0.071	-0.005	0.176	0.000	0.273
	(0.035)	(1.329)	(0.037)	(1.627)	(0.044)	(0.053)	(0.040)	(0.054)	(0.039)	(0.060)	(0.037)	(0.945)	(0.037)	(0.795)
DivFirm x After: Con>Uncon?		0.000		0.027		0.011		0.281		0.100		0.000		0.002
Firm x Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	5,756	4,916	4,417	5,694	4,910	5,756	4,721	5,468	4,142	6,476	5,068	5,387	5,906	4,766
Adjusted R^2	0.721	0.857	0.718	0.825	0.708	0.838	0.673	0.882	0.685	0.835	0.713	0.841	0.724	0.844

Table A.3: Firm Financial Constraints and Plant Emission Responses: Pooled Regressions with Constraint Dummies

The table reports results from pooled triple difference regressions. Results in Panel A compare California and non-California plants based on geographically diversified firms. Panel B studies spillovers to non-California Plants comparing geographically diversified and non-diversified firms. The dependent variable is log (1+Emissions). The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. The firm-level dummy Constrained is an indicator for whether a firm is financially constrained according to each financial constraint measure, i.e. alternatively our composite measure, the Kaplan-Zingales (KZ) index, Whited-Wu (WW) index, Hadlock-Pierce (HP) index, size, payout ratio, and rating. Control variables include PP&E, R&D Stock as well as firm x plant and year fixed effects. Coefficients and their respective standard errors adjusted for clustering at the plant level are reported. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations and the adjusted R^2.

Panel A: California vs Non-California Plants (Geographically Diversified Firms)

	Dependent Variable: Log(1+Emissions)						
	Composite	KZ	HP	WW	Size	Payout	Rating
CalPlant x After x Const.	-0.347	-0.075	-0.313	-0.477	-0.500	-0.431*	-0.224
	(0.225)	(0.187)	(0.293)	(0.374)	(0.368)	(0.237)	(0.194)
CalPlant x After	-0.038	-0.129	-0.091	-0.053	-0.079	-0.023	-0.076
	(0.070)	(0.106)	(0.082)	(0.063)	(0.078)	(0.067)	(0.096)
After x Constrained	-0.006	-0.025	-0.015	-0.148	0.011	0.007	0.049
	(0.059)	(0.052)	(0.068)	(0.148)	(0.112)	(0.051)	(0.047)
PP&E	-0.224	-0.132	-0.153	-0.340	-0.228	-0.170	-0.081
	(0.306)	(0.326)	(0.307)	(0.336)	(0.320)	(0.283)	(0.316)
R&D stock	2.322*	2.020*	2.327**	0.825	0.791	2.192*	1.934
	(1.215)	(1.219)	(1.172)	(1.654)	(1.144)	(1.228)	(1.210)
Firm x Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	3,117	3,017	3,117	3,046	3,102	3,117	3,117
Adjusted R^2	0.928	0.927	0.928	0.928	0.929	0.928	0.928

(continued)

Table A.3: Firm Financial Constraints and Plant Emission Responses: Pooled Regressions with Constraint Dummies (continued)

Panel B: Spillovers to Non-California Plants (Diversified vs Undiversified Firms)

	Dependent Variable: Log(1+Emissions)						
	Composite	KZ	HP	WW	Size	Payout	Rating
DivFirm x After x Const.	0.295***	0.180**	0.193**	0.100	0.179	0.312***	0.211***
	(0.076)	(0.081)	(0.089)	(0.145)	(0.130)	(0.076)	(0.071)
DivFirm x After	-0.004	0.125***	0.054	0.060*	0.060	0.039	0.058
	(0.036)	(0.042)	(0.040)	(0.034)	(0.037)	(0.039)	(0.044)
DivFirm x Constrained	-0.502**	0.060	-0.489**	-0.110	-0.488*	-0.237*	-0.370**
	(0.233)	(0.176)	(0.197)	(0.194)	(0.257)	(0.136)	(0.174)
After x Constrained	-0.356***	-0.248***	-0.272***	-0.291***	-0.304***	-0.288***	-0.223***
	(0.054)	(0.071)	(0.057)	(0.059)	(0.061)	(0.061)	(0.055)
DivFirm	0.183**	0.047	0.166**	0.149*	0.156**	0.147	0.166**
	(0.076)	(0.068)	(0.077)	(0.085)	(0.074)	(0.104)	(0.082)
PP&E	-0.510***	-0.575***	-0.538***	-0.468**	-0.532***	-0.576***	-0.479***
	(0.180)	(0.187)	(0.179)	(0.187)	(0.181)	(0.180)	(0.180)
R&D stock	-0.000	-0.004	0.017	0.007	0.004	-0.004	0.006
	(0.036)	(0.036)	(0.036)	(0.034)	(0.039)	(0.036)	(0.036)
Firm x Plant FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	10,672	10,111	10,666	10,189	10,618	10,455	10,672
Adjusted R^2	0.774	0.774	0.773	0.771	0.773	0.772	0.772

Table A.4: Plant Ownership Responses to California Cap-and-Trade Rule

The table presents results from linear probability estimations of plant closure and opening decisions. The dependent variables are indicators for whether a firm closes (opens) a plant or not. The first two columns compare California and non-California plants based on geographically diversified firms. The last two columns study spillovers to non-California Plants comparing geographically diversified and non-diversified firms. The treatment indicator CalPlant equals to 1 if the plant is located in California and 0 otherwise. The After indicator is equal to 1 if the time period is 2013 or onward and 0 otherwise. The firm-level dummy DivFirm is an indicator for whether a firm owns plants both in California and in other states during a given year or not. The firm-level dummy Constrained is an indicator for whether a firm is financially constrained according to our composite measure. Control variables include PP&E, R&D Stock as well as firm x plant and year fixed effects. Coefficients report marginal effects and standard errors are adjusted for clustering at the plant level. ***, **, and * indicate significance at the 1%, 5% and 10% level, respectively. The table also reports the number of observations and adjusted R^2.

	California vs Non-California Plants (Geographically Diversified Firms)		Spillovers to Non-California Plants (Diversified vs Undiversified Firms)	
	Close	Open	Close	Open
CalPlant x After x Const.	0.076	-0.054		
	(0.053)	(0.060)		
CalPlant x After	0.016	0.028		
	(0.028)	(0.045)		
DivFirm x After x Const.			-0.112***	0.133***
			(0.023)	(0.032)
DivFirm x After			0.016	-0.005
			(0.016)	(0.020)
DivFirm x Constrained			0.040	-0.235***
			(0.037)	(0.066)
DivFirm			0.006	-0.038
			(0.018)	(0.029)
After x Constrained	-0.037*	0.056*	0.071***	-0.079***
	(0.020)	(0.029)	(0.013)	(0.017)
PP&E	0.036	-0.193	0.056	-0.239***
	(0.074)	(0.137)	(0.054)	(0.073)
R&D stock	0.232	2.994**	1.160***	0.118***
	(0.599)	(1.500)	(0.328)	(0.009)
Firm x Plant FE	Yes	Yes	Yes	Yes
Year FE	Yes	Yes	Yes	Yes
Observations	2,485	2,692	8,138	9,261
Adjusted R^2	0.361	0.185	0.324	0.253

Measuring the Impact of Climate Policy Risk

Michael Barnett *

University of Chicago

PRELIMINARY AND INCOMPLETE

May 25, 2019

Abstract

I estimate the impact of the risks associated with climate policy action on oil production, the spot price of oil, and asset prices. Using an event study analysis I show that events corresponding to an upward shift in the likelihood of significant future climate policy action lead to an increase in the returns of sector portfolios that are more exposed to climate policy risk, while a downward shift leads to increased returns. I then construct a climate policy index measure to estimate the impact of climate policy events on oil production and oil prices over the times series. In reduced form estimation, I find that in the recent, policy-relevant time period increases in the likelihood of future climate policy action increase oil production and decrease oil prices, as well as oil firm values. Finally, using a structural VAR framework I find evidence that in the recent, policy-relevant time period climate policy likelihood shocks that increase the likelihood of significant future climate policy action not only increase oil production and decrease the price of oil, but do so in a dynamic and persistent manner.

Keywords: Asset Pricing, Oil Prices, Climate Change, Climate Policy, Stranded Assets

---

*Michael Barnett. Email: mdbarnett@uchicago.edu. Address: University of Chicago, 1126 E. 59th Street - Saieh Hall, Chicago, IL 60637. I am very grateful for advice and suggestions of my advisors and committee Lars Peter Hansen, Pietro Veronesi, Michael Greenstone and Bryan Kelly. I want to thank Buz Brock, Paymon Khorrami, Ralph Koijen, Stefan Nagel, Alan Sanstad, Willem Van Vliet, and Amir Yaron. I'd also like to think Pietro Veronesi and Yoshio Nozawa for providing code for the CME data. Parts of this originally compromised a chapter from my dissertation "A Run on Oil: Climate Policy, Stranded Assets, and Asset Prices." I am grateful for the financial support of the National Science Foundation, the Fama-Miller Center, the Energy Policy Institute at the University of Chicago (EPIC), the Stevanovich Center for Financial Mathematics, and the University of Chicago.

1 Introduction

Climate change, and its potential physical consequences, has become one of the most significant issues currently facing governments, institutions, and individuals around the world. These concerns have led policymakers to propose significant policy actions in an attempt stave off the possible long-term economic damages of climate change. Yet, these policy actions have generated additional risk that comes not only from the uncertainty about when or if such climate policies will ever be implemented, but also from the far-reaching implications of such policies for the profitability of significant sectors of the economy and the significant possibility of important natural resources becoming stranded or unusable. Understanding the effects of these climate policy risks is therefore essential for understanding the full economic and financial consequences of climate change.

This paper focuses on empirically estimating how changes in climate policy risk affects economic and financial outcomes. The empirical analysis is motivated by a number of key predictions from a simplified version of the general equilibrium, production-based asset pricing model of Barnett (2019). The model incorporates climate change, as well as climate policy that carries significant risk due to the fact that it restricts the use of oil and has an unknown arrival time. The key predictions of this model are as follows: the risk associated with climate policy leads to dynamically increasingly oil production to avoid having reserves become stranded; the risk of stranded assets from climate policy and the run on oil that causes lead to dynamic reductions in the price of oil; the value of oil firms decreases substantially due to the lower expected usable reserves that stranded assets risk implies and because of the shift in oil production and price that this risk generates.

I test these model predictions through variety of empirical estimation techniques. I first do this using an event-study analysis of climate policy events that shift the likelihood of future climate policy actions taking place. For events that imply a downward shift in the likelihood of future climate policy occurring, such as the 2016 US presidential election or the US Supreme Court decision to put a stay on the Clean Power Plan, the model would predict these events should increase the value of firms with high climate policy risk exposure, such as oil firms, and also increase the price of oil. The opposite should hold for events that increase the likelihood of future climate policy actions, such as the announcements of the Clean Power Plan and the UN's Paris Climate Accord. I estimate the effect of shifts in the likelihood of future climate policy due to climate policy events by regressing sectors' cumulative abnormal returns after the event on their exposure to climate policy risk, proxied for by exposure to oil price shocks as motivated by the model prediction. I find sectors with the highest climate policy risk exposure experienced the largest increases in cumulative abnormal returns for events that decreased the likelihood of future climate policy action and the largest decreases in cumulative abnormal returns for events that increased the likelihood of future climate policy actions, consistent with the model predictions.

Finally, I construct a climate policy event index from realized climate policy, energy sector, and climate-related events to estimate the dynamic impact of changes in climate policy shocks. In estimated reduced-form regressions, I find that increases in the likelihood of major climate policy measured by my index lead to increased global and regional oil production. I also find that positive climate policy shocks lead to increasingly negative returns for the US oil sector and the spot price of oil. Finally, I estimate a structural VAR for the global oil market that includes the climate policy index, and calculate impulse response functions for a shock to climate policy. The results suggest that increases in the likelihood of significant climate change policy leads to long-term and permanent increases in crude oil production and a statistically significant decreases in the oil spot price, consistent with the dynamic predictions of my model. For each index-based empirical test, the statistical and economic significance are greater during the more recent, policy-focused time period (1996-2017) than for the entire available time sample (1973-2017), further validating the temperature dependence of outcomes implied by the model and the dynamic effect of climate policy risk the model predicts. Finally I extend the VAR estimates by using returns from a type of factor-mimicking portfolio for the impacts climate policy event interacted with the climate policy index. Using this measure of shocks to the likelihood of future climate policy action, designed to better capture the magnitude and dynamics of these shocks, I find dynamic and persistent impacts on oil prices and oil production consistent the previous results that are even more significant.

2 Examples of Climate Policy Events

To highlight the effect I am interested in identifying, I outline a number of important climate policy events that have a signaled a shift in the expectation of future significant climate policy action with implications for the use of fossil fuels. In the US, a number of such policy events have occurred. The Energy Policy Conservation Act set the first fuel economy goals in the US, leading to the development of unleaded gas and catalytic converters. The Clean Air Act gave air pollution and vehicle emissions standards for the future while providing technical and financial assistance to state and local governments. The Diesel Emissions Reduction Act, part of the Energy Policy Act which was enacted in 2005, established diesel engine emissions standards for the future which have led to innovations in diesel engine technology from companies like Cummins. The passing of the Energy Independence and Security Act established the Corporate Average Fuel Economy standards regulating vehicle emissions in the future which motivated the development of hybrid and electric vehicles like the Prius, Leaf, and Tesla. The critical features of these policy events for my analysis are that they established expectations for future outcomes, with clear implications for the usability of fossil fuels and fossil fuel dependent goods, while containing uncertainty about how and if such policy actions would remain in place or be implementable in the future because of their long-term nature.

Renewable Portfolio Standards (RPS) are another type of policy action with similar implications that have been used in the US and internationally. RPS policies set required standards for future power and electricity production. These standards typically require an increased fraction of power and electricity to be produced by renewable and green sources, and thus a sizeable reduction coming from fossil fuels. Over half the states in the US, the European Union, Germany, Japan, China, the UK, and numerous other countries and regions have implemented or tried to implement Renewable Portfolio Standard-type policies. Implementation or proposal of such policies again have two significant effects that are central to my analysis. First, these policy events shift expectations about the likelihood of strict future policy requirements and the future use of fossil fuels as signification restriction of use of fossil fuels will be required to meet such targets. Second, with this type of long-term policy action there is no certainty that these targets will be achieved or enforced by future policymakers because, even when these targets are put in to law and carry significant financial penalties for being violated, they may not be economically feasible or in line with future political consensus.

A prime example of the risk implications I focus on in my analysis can be seen with the implementation of the Clean Power Plan and the related climate policy events connected to this regulation. The Clean Power Plan was established by US President Barack Obama in conjunction with the 2015 Paris Climate Accord to reduce greenhouse gas emissions and by setting future renewable portfolio standards requiring increases in the fraction of energy and electricity produced from low-emission and renewable resources while phasing out high-emissions sources like coal and oil. While this policy clearly has significant implications for restriction of the use of fossil fuels in the future, numerous policy actions since then have put in doubt whether the policy will ever be implemented. On February 9, 2016 the US Supreme ruled to put a stay on the policy regulation until a lower court could determine whether the EPA was legally allowed to implement thus type of policy action. Furthermore, the election of Donald Trump to be the President of the United States came with policy promises such as repealing the Clean Power Plan. Thus the Clean Power Plan, as well the follow-up policy events of the US Supreme Court ruling and election of Donald Trump, are events that shift the likelihood of future climate policy restrictions given the forward-looking ramifications of these events.

The aforementioned Paris Climate Accord is another key policy example with global implications. This agreement is the result of the UN Framework Convention on Climate Change to limit change in the global mean temperature (GMT). The agreement was seen as a significant step toward limiting climate change, and likely limiting the use of fossil fuels in the future, but also carries substantial uncertainty about when and if the necessary policy actions to achieve this target will occur. There is no centralized enforcement mechanism holding countries accountable for meeting contribution targets, and furthermore countries contributions are self-determined and self-reported. In addition, some countries have already proposed or suggested they would propose withdrawing from the agreement, such as the US and Brazil in conjunction with the elections of Donald Trump and Jair Bolsonaro, while others have yet to set any concrete plans to reach the proposed goal.

These policy events demonstrate the potential risks of climate policy action tied to their implementation and future impact that I study in this paper. Through the use of event-study analysis around individual climate policy related events such as these, and through reduced form and structural vector autoregression estimates using the time series of these types of climate policy related events, I empirically measure the impact of climate policy risk on oil production, the spot price of oil, and asset prices. By focusing on these three outcomes, I can capture the real and financial implications of these risks, as well as the interaction of these effects. Furthermore, I am able to study both cross-sectional and dynamic implications related to the risk of climate policy actions these events embody.

3 Related Literature

This paper builds on and contributes to various areas of work in economics, finance, and climate change. The first area is the literature focusing on the interaction between economics and climate change. Stern (2007), Nordhaus (2014), Golosov et al. (2014), Acemoglu et al. (2016), Pindyck and Wang (2013), Hambel et al. (2015), and Cai et al. (2015) are recent examples of theoretical frameworks that examine the social cost of carbon and optimal carbon taxation, directed technological change, and other key climate-economic elements using theoretical models. Kelly and Kolstad (1999), Crost and Traeger (2011), Lemoine and Traeger (2012), Anderson et al. (2016), Brock and Hansen (2017), and Barnett et al. (2018) focus on elements of uncertainty related to risk, ambiguitiy, and misspecification related to climate change and climate models. Deschenes and Greenstone (2007), Dell et al. (2012), Hsiang et al. (2017), and Burke et al. (2018) empirically estimate climate damages in different economic sectors and regions and the impact of climate change on economic growth. McGlade and Ekins (2015) and the Grantham Research Institute study the potential magnitude of stranded assets for fossil fuels based on proposed temperature ceiling policies using least-cost analysis and implications for a potential "carbon bubble," or possible overvaluation of oil firms from not accounting for stranded assets risk, respectively. The Green Paradox, a theory proposed by Sinn (2007) and recently extended by Kotlikoff et al. (2016), suggests the possibility that climate policy intended to mitigate climate change on the demand side may cause firms to alter the timing of their fossil fuel production in a possibly harmful way. This paper builds on elements from these areas and extends this work by motivating critical features of the climate economic models and the potential for stranded assets from climate policy risk to motivate the empirical estimation that examines the dynamic implications of climate policy risk for real and financial outcomes in the oil sector

This paper also contributes to important areas in the asset pricing literature. The interaction between government and asset prices has been studied by Santa-Clara and Valkanov (2003), Pastor and Veronesi (2012), Belo et al. (2013), and Kelly et al. (2016), which focus election and political uncertainty for asset prices, exploiting differential exposures to these types of risk. Sialm (2006) and Koijen et al. (2016) are two examples closely related to this paper that study the impacts of policy risk related to taxes and healthcare on asset prices, whilePástor and Veronesi (2009) provides an important example of the asset pricing impacts that arise from a shift in the production function, in their case due to learning about and adopting a new technology, similar to the impact of the policy events in my model. There is a growing recent literature about climate change and asset prices that my paper adds to. Examples here include Giglio et al. (2015), Bansal et al. (2016), Dietz et al. (2017), Hong et al. (2016), Barnett (2017), and Engle et al. (2019). These papers explore the term structure of long-run discount rates and their interaction with climate change mitigation efforts, the impact of climate change and long-run risk on the social cost of carbon and asset prices, the elasticity of climate damages, the reaction of stock prices in the food sector to climate change, the cross-sectional and time series implications of climate change and climate model uncertainty on economic and asset pricing outcomes, and the ability to hedge climate risk, respectively. My paper provides an important contribution here by empirically studying the impacts of climate policy risk on production and asset prices.

This paper contributes to the expansive work on oil prices and extraction. Hotelling (1931) and Dasgupta and Heal (1974) provide the foundations for the literature on natural resource extraction, while Hamilton (2005), Hamilton (2008), Kilian (2008), Kilian and Park (2009), Kilian (2009), and Baumeister and Hamilton (2019) provide estimates for the link between oil prices and economics and financial shocks. Carlson et al. (2007), Casassus et al. (2009), Kogan et al. (2009), David (2015), Ready (2015), and Bornstein et al. (2017) provide model frameworks to explain varying stylized facts of oil prices and identify important model mechanisms required to match those outcomes. This paper provides estimates for measuring the impact of climate policy risk on oil production and oil prices, a previously unexplored channel in this literature.

Finally, this paper is connected to the narrative approach empirical work of Romer and Romer (2010), Ramey (2011), and Mertens and Ravn (2014). These papers use narrative records to construct an index of relevant policy shocks and then use this index as a proxy for these shocks in their empirical estimation strategy. The climate policy index I construct is similar to this strategy of constructing a policy shock proxy from the text of narrative records by hand-collecting key events from lists of significant climate policy related events that I sign based on the expected impact on the likelihood of future policy. Furthermore, I extend this analysis by interacting my narrative-style climate policy index with a type of factor-mimicking portfolio to capture magnitude and dynamic implications of the climate policy shocks my index is designed to capture.

4 Modelling the Impact of Dynamic Climate Policy Risk

To motivate the empirical tests in my analysis, consider the implications from a simplified version of the model from Barnett (2019). I provide details for this simplified model in the appendix, and direct the reader to Barnett (2019) for full details and a more rich theoretical analysis of climate policy risk. The model consists of households, final output and oil good production, and climate and climate policy components. The solution to this model which includes the risk of climate policy that could strand oil reserves and has a climate-dependent arrival rate λ_t is given by the following proposition:

Proposition 1. *With dynamic climate policy risk where the arrival rate of policy is given by the climate-dependent function λ_t and where oil use is restricted after the policy shock, the optimal extraction, spot price of oil, and oil firm value are given by*

$$O_t = a_n \frac{v}{v_R - v_{CC}}, \quad P_{O,t} = \nu \tilde{Y}_t O_t^{-1}, \quad S_t^O = a_O \frac{v_R R}{v} \tilde{Y}_t$$

Note v is the solution to the HJB equation characterizing the planner's problem (given in the appendix), v_R is the marginal value of oil reserves and $-v_{CC}$ is the marginal cost of climate change. The constants a_n, a_O are functions of the model parameters only (also given in the appendix), ν is the oil input demand share and $\tilde{Y}_t$ is final output net of climate damages.

Without the dynamic climate policy risk, this model behaves like a standard Hotelling-type model where the effects of climate change are internalized through an optimal tax. As oil reserves decrease, the marginal value of holding reserves increases in order to produce in future periods. The internalized climate impact would further diminish oil production because of the climate damages resulting from carbon emissions that increase atmospheric temperature. However, dynamic climate policy risk importantly alters the key forces that drive these results. Directly, temperature impacts climate damages and the likelihood of climate policy occurring, but now in a way that the temperature-related adjustment to extraction and the subjective discount rate adjustment depend on the state of climate change. Indirectly, temperature has greater influence on the marginal value of reserves because the value function is no longer separable.

Though the production and asset pricing impacts associated with oil reserves and temperature can only be determined numerically, we know the impact of climate policy will matter because of the expressions for oil extraction, the spot price of oil, and the oil firm value given above. The key impact of the climate-linked policy risk, through both the direct and indirect effects, is that it creates a dynamic, climate-related feedback mechanism in the model. The risk of a policy shock that restricts the use of oil leads to an increased level of oil extraction. This is true whether λ is constant or temperature dependent. However, as the arrival rate is state dependent, increased oil extraction caused by the stranded asset risk leads to increased climate change. This further exacerbates the stranded assets risk, by increasing the likelihood of a policy shock, and thus provides motivation for oil firms to further increase their oil extraction. The link between climate change and climate policy that strands oil reserves generates a feedback loop that can lead to a dynamic increase in oil production, not simply a level shift up. This can occur even as oil reserves are decreasing and temperature is increasing.

The expression for the spot price of oil clearly demonstrate the inverse relationship between oil production and the spot price of oil. As a result, the forces in play generating a run on oil production will lead to a drop in oil spot prices because of the significant supply of oil in the market. This drop is a dynamic effect, amplified over time through the dynamic feedback loop that alters the optimal path of oil extraction over time.

Finally, the price of the oil firm incorporates two features. The first is that the firm price includes the damage-scaled final output. Therefore, forces related to the impact of climate damages and the impact of the run on oil production for the damage-scaled final output that impact the final output firm and green energy firm prices still matter here. However, the second feature is that the price of the oil firm is also scaled by the marginal value of reserves v_R. Thus, the oil firm has an additional force impacting its price. We know from the macroeconomic outcomes that the risk of stranded assets from a climate policy shock will cause the marginal value of oil reserves to decrease over time as reserves diminish and climate change increases. Therefore, we expect that the price of the oil firm will be lower than without the temperature-dependent risk of climate policy, due to the reduced value of holding oil reserves in this setting. Moreover, we also expect that the price will decrease dynamically due to the increasing likelihood of policy occurring that is driving the run on oil extraction.

5 Empirical Analysis

The solution to the model outlined above provides a number of important predictions that I will now examine empirically. The first prediction of the model is that the dynamic risk of climate policy generates a run on oil, meaning the production of oil dynamically increases with increases in climate change which correspond to increases in the likelihood of a climate policy shock. The model also predicts that the temperature-dependent climate policy risk depresses the spot price of oil due to the increased oil production. Empirically, these outcomes correspond to an observed increase in the likelihood of future climate policy leading to an increase in current and future oil production and a decrease in current and future oil spot prices. The model further predicts that the value of oil firms decreases due to the risk of stranded assets leading to expectations that not all oil reserves held by firms can be used and because the run on oil depresses oil prices. Lastly, the model predicts that the value of the final output firm will increase due to decreased oil prices and policy expectations. These final two predictions can be re-stated as empirically observed increases in the likelihood of future climate policy occurring should lead to oil firms and firms with the highest climate policy risk exposure experiencing decreased returns and non-oil firms and firms with the lowest climate policy risk exposure experiencing increased returns.

I examine these predictions using the following empirical exercises. The first exercise examines the impact of individual events that shift the likelihood of future climate policy action occurring. Estimating cross-sectional regressions for returns of US sector portfolios on a proxy for climate policy exposure shows sectors with greater climate policy exposure experienced larger increases in returns from climate policy events that decreased the likelihood of future climate policy action and larger decreases in returns from climate policy events that increased the likelihood of future climate policy action. Thus, the post-event outcomes provide evidence that the shift in expectations for future climate policy did impact asset prices as my model predicts.

I then extend this event-type analysis of the model predictions by exploring the impact of the time series of climate policy related events. I construct an index of the time series of climate policy events by aggregating lists of key climate- and energy-related events from non-partisan, informational websites. With this time-series index of relevant events, I first test the model predictions using reduced-form regressions of the impact of climate policy shocks on oil production in different regions, returns of US oil sector firms, and returns of the oil spot price. The second approach focuses on the dynamic impact of climate policy shocks by estimating a vector autoregression that incorporates the climate policy events index into a standard global oil market model to examine the impact of climate shocks on current and future oil production and oil prices. Results from each of these exercises again appear to be consistent with the model predictions. Finally, to better capture the dynamics and magnitudes of the effect of the climate policy events in my index I augment the index. I do this by constructing a type of factor-mimicking portfolio, based on the climate policy risk exposure measures used in the event study analysis, and weight the climate policy index by the returns of the portfolio. Repeating the VAR analysis with this return weighted index, I find amplified effects, consistent in direction with the previous estimates, on production and the price of oil for shocks to the likelihood of future climate policy actions.

5.1 Data Sources

The data I use for oil production and oil prices comes from the US Energy Information Administration (EIA). I use global atmospheric temperature as the variable for climate change, which is available from NOAA and NASA. Data on returns for the oil sector and the market come from Ken French's website and the Compustat/CRSP merged database available from WRDS. Macroeconomic variables such as GDP growth, deflators, and other indicators come from FRED, the BEA, Lutz Kilian's website, and James Hamilton's website. Additional oil spot price data come from FRED and the CME's End of Day database, made available to me through the University of Chicago Booth School of Business Fama-Miller Center. Finally, I construct a proxy for changes in the probability of climate policy occurring in my model by compiling a time series of significant climate, climate policy, and energy events (major fossil fuel and alternative energy events, IPCC meetings, US presidential election results, and lists of major climate policies and US energy policies) from non-partisan government, academic, and non-profit informational websites (ProCon.org, IPCC website, and Wikipedia.org). Table 8 shows the list of events since 1997, though the full list extends back to match the full range of dates available from the EIA (1973). As indicated in the table, events can either be positive or negative policy shocks in terms of an increased or decreased likelihood, respectively, of a shift to the production function. The variable contains values of 0 for no event, 1 for a positive event, or -1 for a negative event at the daily level that are then aggregated to a monthly count for the empirical analysis. Section 5.3 provides further details on the index.

5.2 Climate Policy Event Study Analysis

The first empirical exercise I do estimates the impact of individual events that in expectation would shift the likelihood of future climate policy action occurring on stock returns for different sectors in the US economy based on their estimated exposure to climate policy. I focus on events which signal changes in expectations for future policy outcomes as these events more directly test the mechanism related to the likelihood of future climate policy action that my analysis is focused on. Within the context of the model, these events can be compared to comparative statics in the model where λ_t is shifted up or down. Moreover, the more unexpected the shock to the value of λ_t, the more cleanly we can identify the impact as a comparative static shock rather than a more prolonged response to an expected outcome. Such shocks have clear implications in the model: a shock that increases (decreases) the likelihood of future climate policy action should lead to increased (decreased) oil production, negative (positive) realized returns for the oil sector, and negative (positive) realized returns for the spot price of oil due to increased (decreased) stranded assets risk.

To formalize this exercise using an event-study analysis, I estimate the impact of the unexpected shift in climate policy risk expectations on cumulative abnormal returns after the a climate policy event by exploiting the cross-sectional variation in climate policy risk exposure across different sectors. To estimate this cross-sectional regression, I use daily returns for the 49 sector portfolios provided on Ken French's website. I derive abnormal returns as unexplained differences with respect to the market portfolio, or the CAPM model. I estimate the following regression for each sector i using daily returns for the year leading up to the event date to estimate abnormal returns:

$$R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \epsilon_{i,t}$$

The residual for this regression $\epsilon_{i,t}$ is then the abnormal return. I aggregate these residuals for each sector in order to get the cumulative abnormal returns:

$$CAR_{i,t} = \sum_{0}^{t} \epsilon_{i,t} = \sum_{0}^{t}(R_{i,t} - \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}))$$

Next, I derive a measure of climate policy risk exposure. The model predicts that changes in climate-policy expectations influence oil prices and oil production, as well as firm values. Therefore, the model predictions suggest a sector's exposure to climate policy risk can be proxied for by the sector's exposure to oil price innovations or oil production innovations. Given that oil prices are available at a daily frequency, are in direct units of comparison, and are closely linked to oil production, I use exposure to oil price returns as the proxy for exposure to climate policy risk. I estimate this exposure as the beta for oil price returns from the following regression for each sector i over the full available time series of oil prices:

$$R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,OilPrice}R_{OilPrice,t} + \varepsilon_{i,t}$$

To estimate the impact on returns of climate policy risk exposure after the event, I run the following cross-sectional regression:

$$CAR_{i,event} = \delta_0 + \delta_1 \frac{\beta_{i,OilPrice}}{\sigma(\beta_{i,OilPrice})} + e_i$$

Because I normalize the climate policy risk beta by the cross-sectional standard deviation of the beta estimates ($\sigma(\beta_{i,OilPrice})$), the coefficient δ_1 can be interpreted as the percent change in cumulative abnormal returns due to a one-standard deviation increase in the climate policy risk beta resulting from the change in climate policy expectations from the event outcome.

Table 1 provides estimates for a number of recent climate policy-related events. The estimates are the cumulative abnormal return response one day and 4 weeks after the policy events for value- and equal-weighted sector portfolios with t-stats for the heteroskedastic-robust standard errors and the z-stats for the bootstrapped standard errors for the two-stage estimation to account for the inclusion of a generated regressor. Events in the table are those recent events where the impact on returns was statistically significant. Other events tested (such as other US presidential elections since 1996 and the Kyoto Protocol) provided null results.

In each case with significant estimates, the results are consistent with the model predictions. For events that increased the likelihood of future climate policy (the publication of the Clean Power Plan and date of the Paris Climate Agreement) there is a negative CAR response for sectors with higher climate policy exposure. For the events that decreased the likelihood of future climate policy action (the Trump Presidential elections, the announcement date of the US plan to withdraw from the Paris Climate Agreement, and the US Supreme Court ruling to put a hold on implementing the Clean Power Plan) there is a positive CAR response for sectors with higher climate policy exposure.

Focusing on a couple of key events, figures 1-4 provide the one-day and four-week cumulative abnormal return responses for the value- and equal-weighted sector portfolios, sorted by climate policy risk exposure beta, and the estimated climate policy risk exposure impact slope coefficient δ_1 (with the t-statistic and z-statistic for the estimate) for two events that I want to highlight briefly. The first event was the 2016 US presidential election. This event was a surprise shift down in the likelihood of future climate policy action given President Trump's campaign statements about supporting the coal and oil sectors, withdrawing from the Paris Climate Agreement, removing emissions regulations policies such as the Clean Power Plan, and doubting the impact of human behavior on the climate. After the election, for the value-weighted (equal-weighted) portfolios a one-standard deviation increase

Table 1: Event Study Analysis of Significant Climate Policy-Related Events

	1-Day, VW	4-Weeks, VW	1-Day, EW	4-Weeks, EW
Clean Power Plan	**−1.08**	**−0.77**	**−1.01**	**−2.08**
T-Stat: Robust SE	−4.62	−1.14	−2.64	−1.80
Z-Stat: Bootstrap SE	−4.56	−0.98	−2.63	−1.51
Paris Climate Accord	**−0.68**	**−0.39**	**−0.79**	**−1.03**
T-Stat: Robust SE	−2.49	−0.35	−5.30	−2.48
Z-Stat: Bootstrap SE	−2.49	−0.32	−4.50	−1.55
USSC Hold on CPP	**0.55**	**4.49**	**0.42**	**6.82**
T-Stat: Robust SE	2.38	2.17	0.48	8.84
Z-Stat: Bootstrap SE	1.89	2.12	0.46	6.96
Trump 2016 Election	**1.24**	**2.06**	**1.11**	**2.81**
T-Stat: Robust SE	3.05	1.82	1.96	1.99
Z-Stat: Bootstrap SE	2.47	1.57	1.93	1.90
US Paris Withdrawal	**0.71**	**0.30**	**0.71**	**−0.07**
T-Stat: Robust SE	4.25	0.43	5.52	−0.22
Z-Stat: Bootstrap SE	3.91	0.41	4.42	−0.16

Table 1 shows the relationship between the cumulative abnormal returns of sectors after a given climate policy related event and their standardized exposure to climate policy risk. The events are major recent events which had significant responses for cumulative abnormal. The regression specification is given by $R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,OilPrice}R_{OilPrice,t} + \varepsilon_{i,t}$. Cumulative abnormal returns are normalized to zero at the election date, and are estimated with respect to the value-weighted excess market return. Estimates are for value- and equal-weighted sector portfolio cumulative abnormal returns one day and 4 weeks after the election. I provide the t-state for the coefficient for heteroskedasticity-robust standard errors and the z-state for bootstrapped standard errors of the two-stage estimation procedure to account for the use of a generated regressor. See text for full definition of variables.

in climate policy risk beta would have resulted in a 1.24% (1.11%) increase in cumulative abnormal returns after one day and a 2.06% (2.81%) increase in cumulative abnormal return after four weeks. For the value-weighted portfolio estimates, both estimates are statistically significant using heteroskedasticity-robust standard errors (t-statistic), while only the one-day cumulative abnormal return response is statistically significant for the bootstrapped standard errors (z-statistic). For the equal-weighted portfolio estimates, both estimates are statistically significant using heteroskedasticity-robust standard errors (t-statistic) and bootstrapped standard errors (z-statistic).

The second event was the 2016 US Supreme Court decision to put a stay on the Clean Power Plan. This event was also a surprise decrease in the likelihood of future climate policy action given the lower courts had yet to rule on the constitutionality of the policy. After the court decision, for the value-weighted (equal-weighted) portfolios a one-standard deviation increase in climate policy risk

Figure 1: Election Impact on Returns by Climate Policy Exposure - Value-Weighted

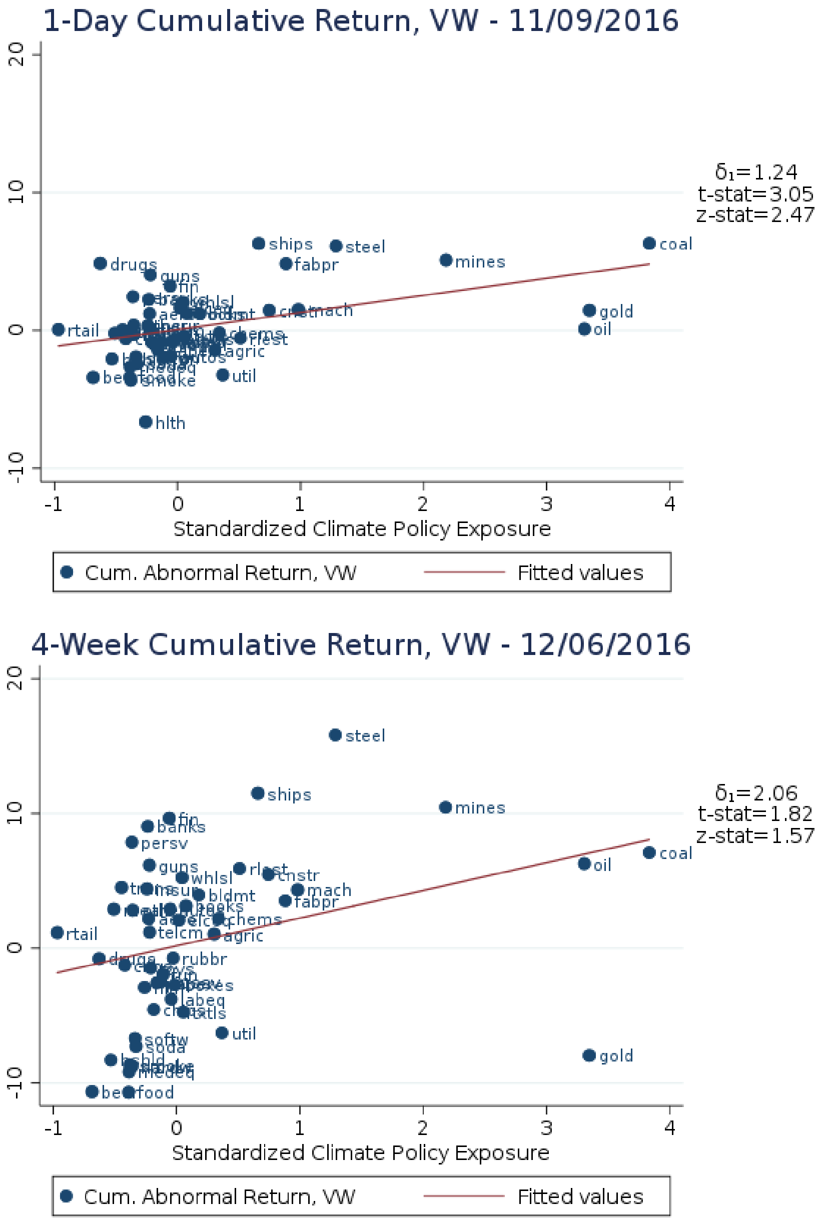


Figure 1 shows the relationship between the cumulative abnormal returns of sectors after the 2016 US presidential election and their standardized exposure to climate policy risk. The regression specification is given by $R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,OilPrice}R_{OilPrice,t} + \varepsilon_{i,t}$. Cumulative abnormal returns are normalized to zero at the election date, and are estimated with respect to the value-weighted excess market return. The top panel are estimates for cumulative abnormal returns one day after the election, and the bottom panel are estimates for cumulative abnormal returns four weeks after the election. See text for full definition of variables.

beta would have resulted in a 0.55% (0.42%) increase in cumulative abnormal returns after one day and a 4.49% (6.82%) increase in cumulative abnormal return after four weeks. For the value-weighted portfolio estimates, both estimates are statistically significant using heteroskedasticity-robust stan-

Figure 2: Election Impact on Returns by Climate Policy Exposure - Equal-Weighted

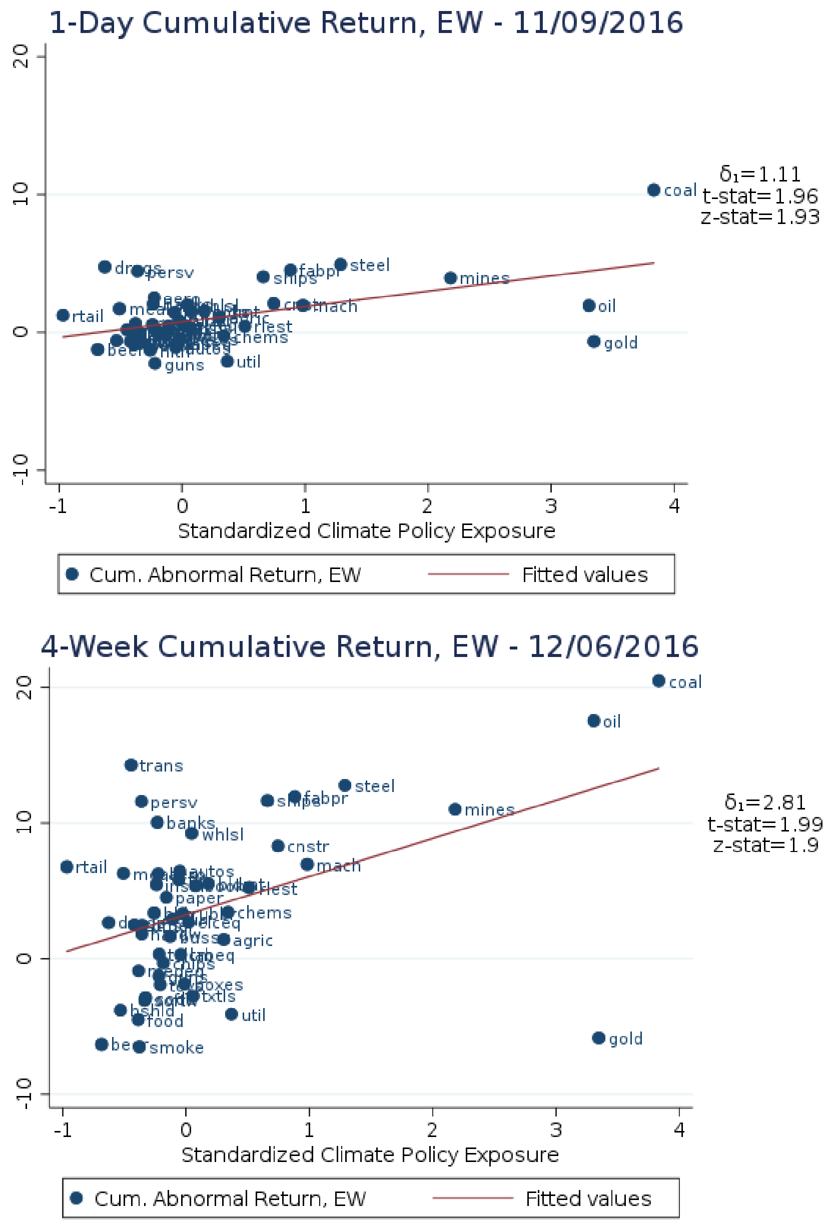


Figure 2 shows the relationship between the cumulative abnormal returns of sectors after the 2016 US presidential election and their standardized exposure to climate policy risk. The regression specification is given by $R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,OilPrice} R_{OilPrice,t} + \varepsilon_{i,t}$. Cumulative abnormal returns are normalized to zero at the election date, and are estimated with respect to the value-weighted excess market return. The top panel are estimates for cumulative abnormal returns one day after the election, and the bottom panel are estimates for cumulative abnormal returns four weeks after the election. See text for full definition of variables.

dard errors (t-statistic) and the bootstrapped standard errors (z-statistic). For the equal-weighted portfolio estimates, only the 4-week estimated impacts are statistically significant, for both the heteroskedasticity-robust standard errors (t-statistic) and bootstrapped standard errors (z-statistic).

Figure 3: Court Impact on Returns by Climate Policy Exposure - Value-Weighted

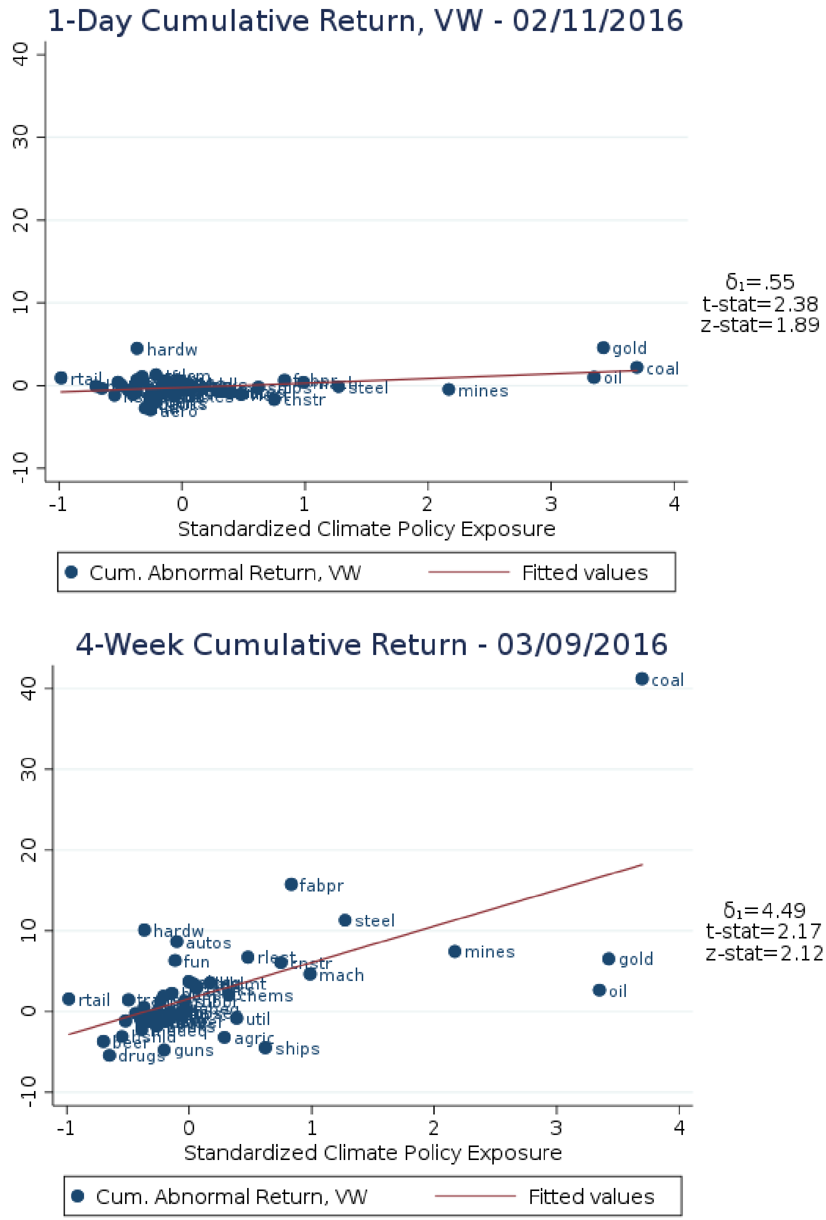


Figure 3 shows the relationship between the cumulative abnormal returns of sectors after the 2016 US Supreme Court decision to put a stay on the Clean Power Plan and their standardized exposure to climate policy risk. The regression specification is given by $R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,OilPrice}R_{OilPrice,t} + \varepsilon_{i,t}$. Cumulative abnormal returns are normalized to zero at the election date, and are estimated with respect to the value-weighted excess market return. The top panel are estimates for cumulative abnormal returns one day after the election, and the bottom panel are estimates for cumulative abnormal returns four weeks after the election. See text for full definition of variables.

I highlight these two examples as they help link the observed outcomes to the model. First, these two events are arguably two of the more unanticipated outcomes listed, meaning the estimates are more likely to capture the full impact of the event whereas the other events being more anticipated

Figure 4: Court Impact on Returns by Climate Policy Exposure - Equal-Weighted

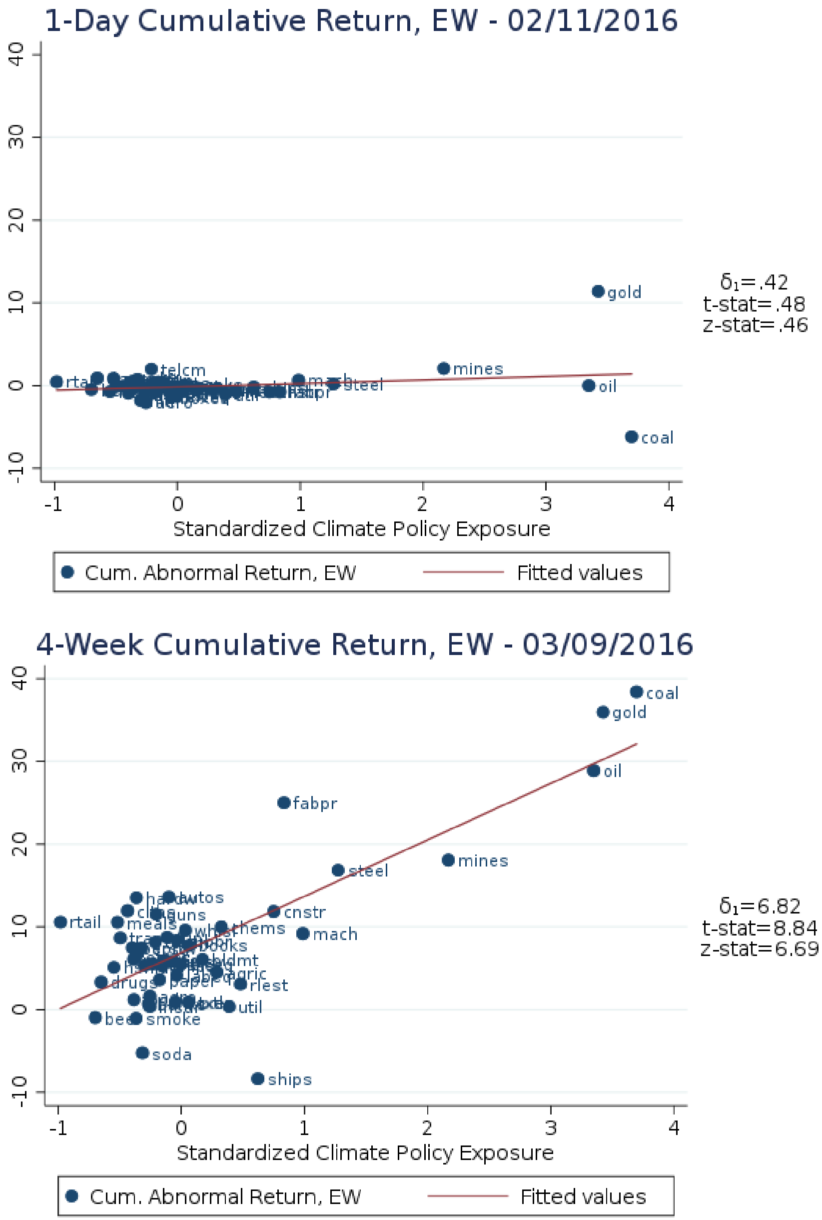


Figure 4 shows the relationship between the cumulative abnormal returns of sectors after the 2016 US Supreme Court decision to put a stay on the Clean Power Plan and their standardized exposure to climate policy risk. The regression specification is given by $R_{i,t} = \alpha_i + \beta_{i,Mkt}(R_{Mkt,t} - R_{f,t}) + \beta_{i,OilPrice}R_{OilPrice,t} + \varepsilon_{i,t}$. Cumulative abnormal returns are normalized to zero at the election date, and are estimated with respect to the value-weighted excess market return. The top panel are estimates for cumulative abnormal returns one day after the election, and the bottom panel are estimates for cumulative abnormal returns four weeks after the election. See text for full definition of variables.

mean that the estimates are likely lower bounds on the estimated event impacts. This corresponds with the fact that these events had the largest estimated impacts on cumulative returns. Second, these events also saw dynamic effects that played out for up to four weeks that were statistically significant.

While the use of the climate policy risk exposure measure based on oil prices and production is one way that helps link these outcomes to my specific model, the dynamic responses, corresponding to the dynamic responses related to production and pricing impacts my model predicts, provide further formal evidence consistent with my model mechanism in the direction of the impact and dynamics.

5.3 Climate Policy Events Index

I now extend the empirical analysis to the time series of climate policy related shocks. To identify the impact of climate policy shocks to oil production and oil sector returns, I first estimate reduced-form regressions focused on the link between changes in the probability of climate policy occurring, as measured by climate- and climate-policy-related events, and oil production decisions, oil sector returns for US firms, and oil spot price returns. The climate policy shocks measure is labeled as $ClimPol$, the index variable tracking different climate related events discussed previously. For example, events in $ClimPol$ include the establishment of the Paris Accord in 2015, as well as the election of Donald Trump as the President of the United States in 2016. The Paris Accord is considered a positive shock to the arrival rate of a significant climate policy action and so a positive one in the index and the election of Trump is considered a negative shock and so a negative one in the index. I identify these events at a daily level, and then aggregate them up to monthly values for my analysis.

The goal of this exercise is to identify whether events related to changes in the likelihood of future climate policy action lead to changes in production and prices consistent with the model predictions. The model predicts that a positive shock to the arrival rate should cause an increase in oil production and negative oil firm returns and oil price returns, whereas a negative shock to the arrival rate should lead to a decrease in oil production and increase in oil firm returns and oil price returns.

The reduced-form regression approach provides a estimate of how climate policy driven demand shocks influence economic and financial outcomes. To determine whether the empirical outcomes from this simple analysis are consistent with the model, I focus on the signs and statistical significance of the estimates, and compare those with the qualitative results of the model. Furthermore, I will estimate each regression on the full time sample (1973-2017) and on a shorter, more recent subsample (1996-2017). The recent sub-sample I refer to as the policy-relevant sample. I choose 1996 as the starting year for this sample as it is near the time when major climate policy begins to take place, such as the Kyoto Protocol, which was an early global climate agreement similar to the recent Paris Climate Accord. The model would predict that impacts estimated in the policy relevant subsample should be higher as temperature has increased and the likelihood of climate policy occurring is higher.

5.3.1 Oil Production

I begin by focusing on the impact of climate policy on oil production. To estimate the effect of climate policy shocks, I estimate the following regression:

$$Y_t = \alpha + \beta_Y Y_{t-1} + \phi_{ClimPol} ClimPol_t + \epsilon_t$$

where Y_t is crude oil production, Y_{t-1} is the one-period lag of crude oil production, and $ClimPol$ is the index for the climate events mentioned above. I exploit the time series and panel dimensions of the data by estimating this regression using information from the changes in the dependent variable across time for different regions of interest. Tables 2 and 3 show the results for oil production across four different regions using the full time sample (1973-2017), and the bottom table shows the results for oil production across four different regions using the more recent, policy-focused time sample (1996-2017).

Table 2: Climate Policy Impact on Oil Production (1973-2017)

	OPEC	US	Non-OPEC	World
ClimPol	-0.022	0.026	0.099	0.069
S.E.	(0.034)	(0.063)	(0.012)	(0.072)
# Obs.	536	536	536	536
R^2	0.995	0.979	0.991	0.988

Table 3: Climate Policy Impact on Oil Production (1996-2017)

	OPEC	US	Non-OPEC	World
ClimPol	0.035	0.041	0.141	0.176
S.E.	(0.055)	(0.015)	(0.047)	(0.074)
# Obs.	260	260	260	260
R^2	0.980	0.972	0.988	0.987

Tables 2 and 3 show the impact of climate policy events as measured by the $ClimPol$ index on oil production for the Non-OPEC, OPEC, US, and World regions. The top table are estimates using the full time sample of data (1973-2017), and the bottom table are estimates using the policy-relevant time subsample (1996-2017). The regression specification is given by $Y_t = \alpha + \beta_Y Y_{t-1} + \phi_{ClimPol} ClimPol_t + \epsilon_t$. I omit the constant and lag variable coefficients from the table. See text for full definition of variables.

Controlling for the lag value, meant to capture relevant market conditions and current market effects or trends, oil production increases for the US, non-OPEC countries, and globally for an event that increases the likelihood of climate related policy. The effect is larger and more statistically significant in the recent policy-focused time period estimates. Climate policy events have essentially no impact on the production of the OPEC-countries region, though the estimate goes from negative to positive when comparing the full time sample to the policy-relevant time sample. These results are in line with two predictions of the model. First, they are consistent with the result of a run on oil occurring for an increased likelihood in climate policy occurring, at least for all regions that are not exclusively the OPEC region. Second, the increased magnitude for the most recent time period is consistent with the prediction that an increased likelihood of climate policy, tied to higher temperatures, should generate larger impacts on production.

5.3.2 Oil Sector and Oil Price Returns

Next, I test the model implication for the impact of climate policy on oil sector returns and oil price returns. To do this, I estimate whether shocks to climate policy predict negative changes to oil sector returns by estimating the following regression:

$$r_{i,t+1,t+h} = a_i + b_i X_t + c_i ClimPol_t + \varepsilon_{i,t}$$

where $r_{i,t+1,t+h}$ are 1-, 6-, 12-, 18-, and 24-month ahead cumulative returns, i is for cumulative excess returns for the oil sector portfolio or cumulative returns for the WTI oil price, $ClimPol_t$ is the climate policy dummy, and X_t includes non-contemporaneous controls for the market portfolio, economic productivity, spot price and sector returns, oil production innovations, and log OECD industrial production innovations. Tables 4 through 7 show the results across the five different cumulative return scenarios using the full time sample (1973-2017), and the results across the five different cumulative return scenarios using the more recent, policy-focused time sample (1996-2017).

After including controls to capture relevant market and macroeconomic conditions and trends, I find that shocks to climate policy have a negative impact on oil sector and oil price returns, as can be seen across the different horizons and time samples used for estimation. The coefficients are negative or insignificant for all horizons of cumulative returns. The effect is more negative and is statistically significant for the estimates based on the more recent time period sample at the 18- and 24-month horizons. The magnitude of the impact and predictability are also increasing with the horizon of the cumulative returns, as the coefficients become more negative and the R^2's become larger. These results are in line with three more predictions of the model. First, the results are consistent with the prediction that shocks to climate policy that lead to an oil run also depress oil sector firm values and oil prices. Second, the impact has a dynamic effect on outcomes as the negative returns persist

Table 4: Climate Policy Impact on Oil Sector Returns (1973-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol	0.000	0.005	-0.009	-0.035	-0.027
S.E.	(0.005)	(0.013)	(0.020)	(0.026)	(0.036)
# Obs.	534	529	523	517	511
R^2	0.004	0.016	0.014	0.028	0.013

Table 5: Climate Policy Impact on Oil Sector Returns (1996-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol	-0.002	-0.002	-0.034	-0.072	-0.085
S.E.	(0.007)	(0.015)	(0.024)	(0.032)	(0.038)
# Obs.	260	260	260	260	260
R^2	0.021	0.017	0.016	0.045	0.041

Tables 4 and 5 show the impact of climate policy events as measured by the *ClimPol* index on returns for the value-weighted US Oil sector portfolio. The regression specification is given by $r_{i,t+1,t+h} = a_i + b_i X_t + c_i ClimPol_t + \varepsilon_{i,t}$. $r_{i,t+1,t+h}$ is the k-month cumulative return for the value-weighted US Oil sector portfolio. X_t is a vector of control variables that includes the lagged values for the value-weighted market portfolio return, the value-weighted oil sector portfolio return, lagged oil production innovations, lagged oil spot price returns, and lagged real economic activity innovations. The top panel are estimates using the full time sample of data (1973-2017), and the bottom panel are estimates using the policy-relevant time subsample (1996-2017). I omit the constant and control coefficients from the table. See text for full definition of variables.

and increase in magnitude over the longer cumulative return horizons explored. Finally, as was seen with oil production, the increased magnitude over the most recent time period as compared to the full time period estimates is consistent with an increased impact of climate policy as temperature increases and the likelihood of significant climate policy occurring increases.

5.3.3 Climate Policy*Temperature Interaction Estimates

To strengthen the validity of the climate policy index analysis and further connect the estimated results to the model, I augment these regressions by using a climate policy index and temperature interaction term. The model specifies that the likelihood of climate policy is tied to increases in temperature, and therefore increases in temperature should amplify the impact of climate policy risk. The use of the climate policy index interacted with temperature directly tests this link, while also still testing the impact that policy and climate have on oil production, oil sector returns, and oil

Table 6: Climate Policy Impact on Oil Price Returns (1973-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol	-0.002	0.042	0.050	0.018	0.017
S.E.	(0.008)	(0.026)	(0.042)	(0.045)	(0.047)
# Obs.	534	529	523	517	511
R^2	0.045	0.021	0.022	0.016	0.021

Table 7: Climate Policy Impact on Oil Price Returns (1996-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol	-0.004	0.021	-0.001	-0.076	-0.099
S.E.	(0.011)	(0.033)	(0.045)	(0.048)	(0.050)
# Obs.	260	260	260	260	260
R^2	0.013	0.020	0.029	0.043	0.047

Tables 6 and 7 show the impact of climate policy events as measured by the *ClimPol* index on returns for the WTI spot price of oil. The regression specification is given by $r_{i,t+1,t+h} = a_i + b_i ClimPol_t + c_i X_t + \varepsilon_{i,t}$. $r_{i,t+1,t+h}$ is the k-month cumulative return for the WTI spot price of oil. X_t is a vector of control variables that includes the lagged values for the value-weighted market portfolio return, the value-weighted oil sector portfolio return, lagged oil production innovations, lagged oil spot price returns, and lagged real economic activity innovations. The top panel are estimates using the full time sample of data (1973-2017), and the bottom panel are estimates using the policy-relevant time subsample (1996-2017). I omit the constant and control coefficients from the table. See text for full definition of variables.

price returns. Though I have previously proposed that the increased effects seen when comparing the subsample of recent, policy relevant data are related to increases in temperature and increases in policy concern, these interaction estimates help verify whether or not this is the case.

I continue to include the same lag values and controls as before in the regression equations. The key difference is that the dependent variable of interest is now an interaction variable of the climate policy index and the one year moving average of global mean temperature, $ClimPol * Temp$. More precisely, the regression specifications for the production, oil return, and oil price returns estimates are respectively given by

$$Y_t = \alpha + \beta_Y Y_{t-1} + \phi_{ClimPol*Temp} ClimPol_t * Temp_t + \epsilon_t$$
$$r^e_{t+1,t+h} = a + bX_t + c_{ClimPol*Temp} ClimPol_t * Temp_t + \varepsilon_t$$
$$r^{spot}_{t+1,t+h} = a + bX_t + c_{ClimPol*Temp} ClimPol_t * Temp_t + \varepsilon_t$$

Table 8: Climate Policy Impact on Oil Production (1973-2017)

	OPEC	US	Non-OPEC	World
ClimPol*Temp	0.061	0.053	0.200	0.236
S.E.	(0.095)	(0.023)	(0.070)	(0.117)
# Obs.	536	536	536	536
R^2	0.995	0.979	0.991	0.988

Table 9: Climate Policy Impact on Oil Production (1996-2017)

	OPEC	US	Non-OPEC	World
ClimPol*Temp	0.035	0.062	0.210	0.264
S.E.	(0.055)	(0.024)	(0.075)	(0.116)
# Obs.	260	260	260	260
R^2	0.980	0.973	0.988	0.987

Tables 8 and 9 show the impact of climate policy events as measured by the *ClimPol* index on oil production for the Non-OPEC, OPEC, US, and World regions. The top table are estimates using the full time sample of data (1973-2017), and the bottom table are estimates using the policy-relevant time subsample (1996-2017). The regression specification is given by $Y_t = \alpha + \beta_Y Y_{t-1} + \phi_{ClimPol*Temp} ClimPol_t * Temp_t + \epsilon_t$. I omit the constant and lag variable coefficients from the table. See text for full definition of variables.

Tables 8-13 provide the estimated coefficients and I outline briefly here the estimation results. For oil production, an increase in the interaction term leads generally to an increase in oil production as it did before. Also similar to before, the impacts are increasing in magnitude and significance for the more recent subsample of data. However, the interaction term for climate policy and temperature is now positive and statistically significant for the US, Non-OPEC, and World regions in the 1973-2017 sample of date. Thus we see an enhanced effect by accounting for temperature within the impact of the climate policy risk, which is in line with the models prediction of increased effects from increased temperature and the proposed justification for increased effects seen in the more recent, policy-relevant and higher temperature subsample of data.

The estimated effects of the interaction term for oil sector and oil price returns also line up with the previous results. An increase to the interaction term has a negative impact on oil sector and oil price returns. The effect is more negative and is statistically significant for the estimates based on the more recent time period sample at the 18- and 24-month horizons, and the magnitude of the impact and predictability are also increasing with the horizon of the cumulative returns. However, again with

Table 10: Climate Policy Impact on Oil Sector Returns (1973-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol*Temp	-0.003	-0.005	-0.058	-0.119	-0.131
S.E.	(0.010)	(0.023)	(0.037)	(0.048)	(0.060)
# Obs.	534	529	523	517	511
R^2	0.004	0.016	0.019	0.037	0.023

Table 11: Climate Policy Impact on Oil Sector Returns (1996-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol*Temp	-0.003	-0.015	-0.077	-0.153	-0.177
S.E.	(0.011)	(0.025)	(0.040)	(0.051)	(0.063)
# Obs.	260	260	260	260	260
R^2	0.021	0.019	0.023	0.059	0.052

Tables 10 and 11 show the impact of climate policy events as measured by the $ClimPol$ index on returns for the value-weighted US Oil sector portfolio. The regression specification is given by $r^e_{t+1,t+h} = a + bX_t + c_{ClimPol*Temp}ClimPol_t * Temp_t + \varepsilon_t$. $r_{i,t+1,t+h}$ is the k-month cumulative return for the value-weighted US Oil sector portfolio. X_t is a vector of control variables that includes the lagged values for the value-weighted market portfolio return, the value-weighted oil sector portfolio return, lagged oil production innovations, lagged oil spot price returns, and lagged real economic activity innovations. The top panel are estimates using the full time sample of data (1973-2017), and the bottom panel are estimates using the policy-relevant time subsample (1996-2017). I omit the constant and control coefficients from the table. See text for full definition of variables.

these estimates there are key main differences from the previous results. First, the estimated impact of the interaction term on the oil sector returns is now monotonically increasing in magnitude and statistical significance, and are statistically significant for the 18- and 24-month cumulative return horizons for the 1973-2017 data sample estimates. And though the impact on oil price returns is not statistically significant for 1973-2017 data sample estimates, the impacts are now all negative and monotonically increasing in magnitude and significance for longer horizons. This results again validates that there is an enhanced effect by accounting for temperature within the impact of the climate policy risk as the model implies, as now even the full sample of data estimates are significant, and further confirms the impacts on oil production and prices estimated in the previous regressions that are consistent with the model implications for climate policy risk.

Table 12: Climate Policy Impact on Oil Price Returns (1973-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol*Temp	-0.009	0.040	0.010	-0.071	-0.083
S.E.	(0.017)	(0.052)	(0.075)	(0.081)	(0.088)
# Obs.	534	529	523	517	511
R^2	0.051	0.018	0.018	0.018	0.023

Table 13: Climate Policy Impact on Oil Price Returns (1996-2017)

	$r_{oil,t+1}$	$r_{oil,t+6}$	$r_{oil,t+12}$	$r_{oil,t+18}$	$r_{oil,t+24}$
ClimPol*Temp	-0.009	0.028	-0.013	-0.145	-0.176
S.E.	(0.017)	(0.028)	(0.078)	(0.079)	(0.081)
# Obs.	260	260	260	260	260
R^2	0.017	0.026	0.030	0.046	0.048

Tables 12 and 13 show the impact of climate policy events as measured by the $ClimPol$ index on returns for the WTI spot price of oil. The regression specification is given by $r^{spot}_{t+1,t+h} = a + bX_t + c_{ClimPol*Temp}ClimPol_t * Temp_t + \varepsilon_t$ is the k-month cumulative return for the WTI spot price of oil. X_t is a vector of control variables that includes the lagged values for the value-weighted market portfolio return, the value-weighted oil sector portfolio return, lagged oil production innovations, lagged oil spot price returns, and lagged real economic activity innovations. The top panel are estimates using the full time sample of data (1973-2017), and the bottom panel are estimates using the policy-relevant time subsample (1996-2017). I omit the constant and control coefficients from the table. See text for full definition of variables.

5.4 Vector Autoregression Analysis

To further the empirical estimation of the dynamic effects of the risk of climate policy shocks on oil sector quantity and price outcomes, I estimate a structural vector autoregression (VAR) for the global oil market. Augmenting the global oil market VARs proposed and used by Kilian and Park (2009), Baumeister and Hamilton (2017), and others, I estimate:

$$y_t = \nu + \sum_j A_j y_{t-j} + u_t$$

where the vector of endogenous state variable vector y_t is defined by

$$y_t = [ClimPol_t, \Delta prod_t, rea_t, \Delta p^{oil}_t]'$$

ClimPol is the climate policy index measure I mentioned previously. $\Delta prod$ is the percent change in global oil production available from the EIA. REA is a measure of real economic activity given by innovations in the log OECD industrial production index suggested in recent work by James Hamilton. r_t^{mkt} is log differences in the real West Texas Intermediate (WTI) monthly closing price for crude oil.

I use a Cholesky decomposition of the estimated variance-covariance matrix for identification of the structural shocks. This identification strategy imposes a recursive interpretation of the impact of the shocks. The general representation and interpretation of this identification is as follows:

$$u_t = B\left[\epsilon_{\text{climate policy}}, \quad \epsilon_{\text{oil supply}}, \quad \epsilon_{\text{aggregate demand}}, \quad \epsilon_{\text{oil-specific demand}}\right]'$$

where B is the lower triangular matrix derived from the Cholesky decomposition of the estimated variance covariance matrix $\hat{\Sigma}$, i.e., $BB' = E_t[u_t' u_t] = \hat{\Sigma}$. I outline the specific interpretation and identification of each shock in what follows.

$ClimPol_t$, the focus of this exercise, captures changes in the likelihood of future climate policy that restricts oil use, that is changes in λ_t from the model. Although long-run temperature directly maps to the likelihood of significant climate policy action in the model, in practice this link is less precise. Figure 9 in the appendix, which shows the US temperature anomaly time series over the annual *ClimPol* index measure, demonstrates this relationship. The time series for the two variables are positively correlated, but the correlation is obviously not one. For this reason, I use the more direct measure of *ClimPol* to capture changes in the likelihood of future climate policy that restricts the use of oil.

This ordering assumes the likelihood of significant climate policy is contemporaneously predetermined with respect to oil sector shocks and the oil sector is contemporaneously influenced by shifts to the likelihood of future climate policy. This assumption is intuitive and maintains consistency with the model in that the likelihood of significant climate policy responds only with a lag to oil sector shocks as a result of emissions from oil production impacting the climate policy arrival rate. Recent climate science work by Matthews et al. (2009), Ricke and Caldeira (2014), and Zickfeld and Herrington (2015) has shown that impacts on temperature from carbon emissions can take many years or even decades to fully realize, which further validates this restriction. The order for the remaining variables follows the setting of Kilian (2009). Thus, this interpretation of the structure fits this setting as well: 1.) a vertical short-run supply curve and downward sloping demand curve; and 2.) oil demand and supply shocks imply immediate changes in the real oil price.

To further highlight the consistency of the VAR model with the theoretical model, consider the following. The supply shocks and real aggregate demand shocks can be interpreted as the shocks to oil reserves and shocks to capital in the model. As in the model, supply and policy shocks are important determinants of oil supply or production, along with temperature shocks, which are correlated with climate policy likelihood. Combined with capital shocks, and the preferences of agents, these shocks then determine oil prices. These effects together then pin down asset prices in the model. The link

Figure 5: ClimPol Shock IRF - 1973-2017 vs. 1996-2017 Time Samples

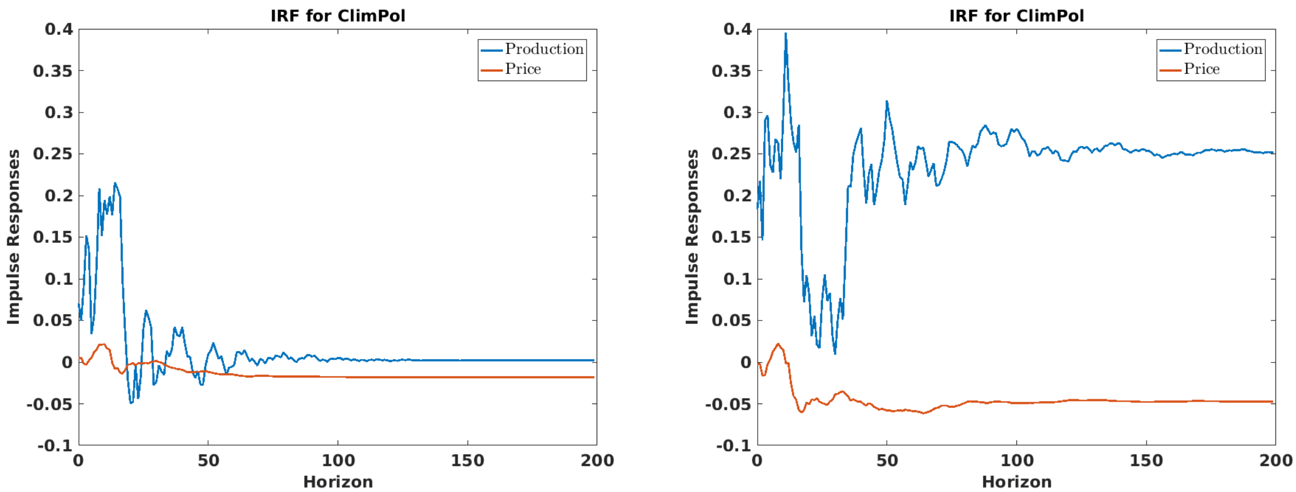


Figure 5 shows estimated impulse response functions for global oil production and the WTI spot price of oil for a shock to the *ClimPol* index. The left panel are estimates using the full time sample of data (1973-2017), and the right panel are estimates using the policy-relevant time subsample (1996-2017). See text for the full VAR specification used and definition of variables. The VAR is estimated using 24 lags.

between the climate and the economy in the model comes through how emissions from oil produced impact temperature and then how temperature influences climate damages and the climate policy likelihood, which feed into the determination of the economic and financial outcomes of interest. Thus, the variables included in the VAR and the ordering of the variables in the recursive decomposition is consistent with the theoretical model framework.

From the VAR estimates and the recursive identification structure, I derive impulse response functions (IRFs), or the cumulative responses to a given structural variable shock, which are the results I use to examine the validity of the model mechanism. To understand how the IRFs generated from the VAR estimation can help validate the model, consider first the expected IRFs from alternative model settings. In the setting without any anticipated risk of a policy shock, a shift in climate policy corresponding to an increased carbon tax would lead to a decrease in oil production and an increase in the spot price of oil, and no change in outcomes if the event did not directly change the carbon tax. In the policy setting where the arrival rate is climate-independent and constant, a shock to the climate policy variable should lead to an increase in oil production and decrease in the spot price of oil, but this effect would not persist because of the lack of temperature dependence and the prevailing Hotelling-type forces. However, in the dynamic climate policy risk setting with a climate-dependent arrival rate, a shock to the likelihood of significant climate policy occurring leads to an increase in oil production and a decrease in the oil spot price. Furthermore, the impacts of a shock to the likelihood of significant climate policy occurring should produce impacts that are persistent and potentially increasing in magnitude dynamically for these outcomes, two defining features of what I have termed

Figure 6: ClimPol Shock IRF - 1996-2017 Time Sample w/ C.I.s

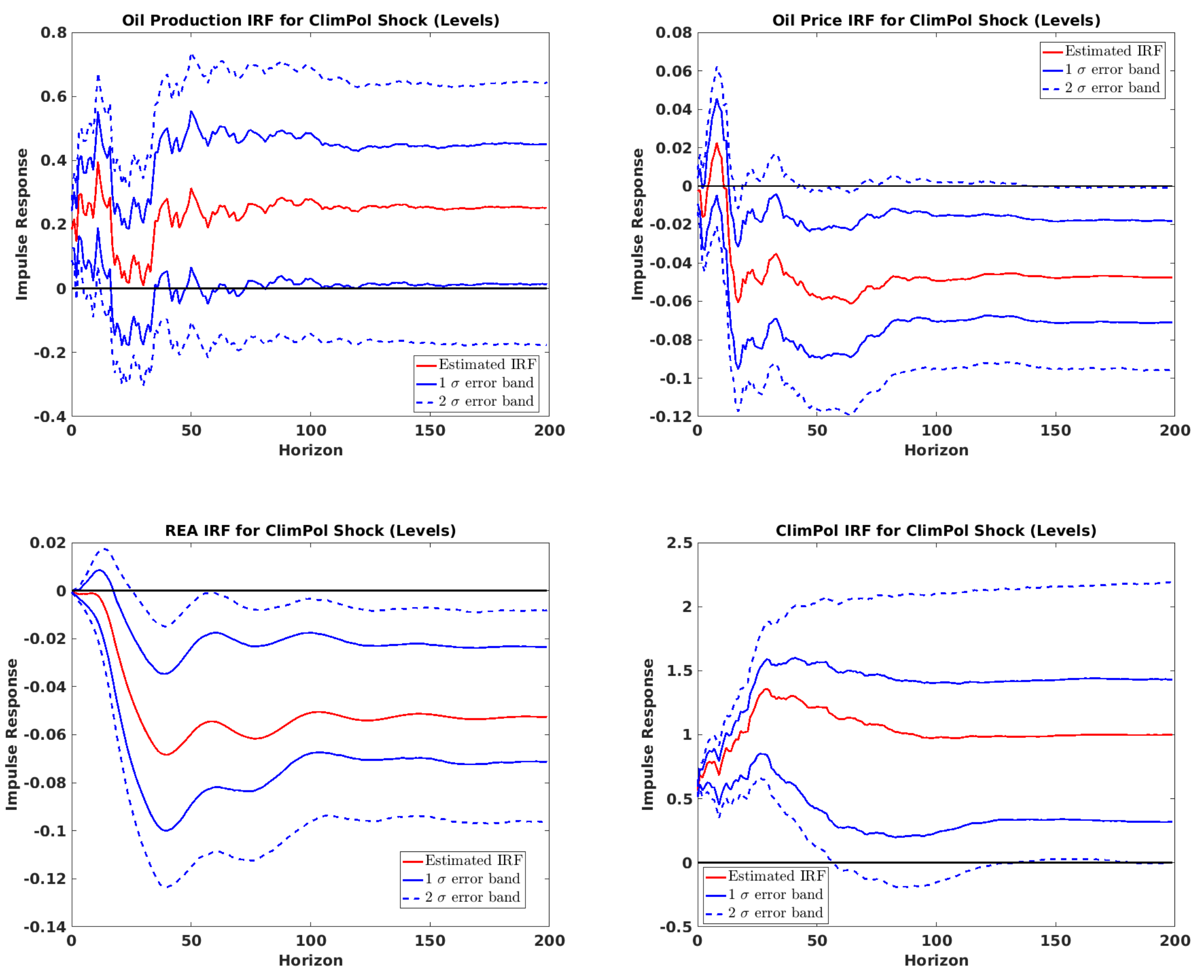


Figure 6 shows the estimated impulse response functions for global oil production, the WTI spot price of oil, real economic activity, and the *ClimPol* climate policy index measure for a shock to the likelihood of climate policy measured by the *ClimPol* index. The red line is the estimated IRF, the solid blue lines represent the on-standard deviation error bands, and the blue dashed lines represent the two standard deviation error bands. Error bands are estimated using boostrapping. The estimates use the policy-relevant time subsample (1996-2017). See text for the full VAR specification used and definition of variables. The VAR is estimated using 24 lags.

a run on oil. Therefore, one can "test" the validity of the model proposed in this paper by looking at the sign and dynamics of the IRFs for oil production and oil spot prices resulting from a shock to the likelihood of significant climate policy occurring.

Figure 5 shows the cumulative level impulse response functions of oil production and oil prices for a shock to the likelihood of significant climate policy occurring. The left plot is for the VAR estimated using the full time sample (1973-2017) and the right plot for the VAR estimated using the more recent, policy-focused time sample (1996-2017). Figure 6 shows the individual impulse response functions with bootstrapped standard errors for the policy focused time sample (1996-2017), where the solid blue lines are for the one-standard deviation confidence interval, the dashed blue lines are for the two-standard deviation confidence interval, and the red line is for the estimated IRF.

The impulse response functions further confirm the results seen previously in the reduced-form estimates. For the impulse response functions generated from the full time sample VAR estimates, the responses are substantially muted and quite close to zero. However, focusing on the impulse responses generated from the VAR estimates using the policy-focused time sample, we see results that correspond to the model predictions. A shock to climate policy leads to an increase in oil production and a decrease in the spot price of oil. Moreover, these impacts grow in magnitude and persist over time. Though only the impact on the spot price of oil is statistically significant for the two variables of interest, as seen in Figure 6, the direction and dynamics of these results taken together are consistent with the key predictions of the model and the other empirical exercises.

5.5 Return-Weighted Climate Policy Index

Finally, I provide an extension of the VAR analysis which uses return-weighted climate policy index measures. The purpose of this extension is to account for the magnitude and dynamics of the effects of the climate policy events on the oil price and production of oil by incorporating the forward-looking information of relevant asset price returns into my climate policy index. This extension ties the asset pricing and production implications into a single analysis to provide a more complete test of the model implications. The results of this extension are not only consistent with the results given above, but in fact identify larger and more significant impacts by exploiting the informational value of asset prices and accounting for policy magnitudes and dynamic implications.

A natural asset price measure of the magnitude of the impact of climate policy events follows from the event study analysis done previously. Each event in the event study analysis required estimating a climate policy risk exposure measure for the cross-section of sector portfolios. These climate policy risk exposure measures provide a weighting scheme for a type of factor-mimicking portfolio that captures not only the realization of climate policy events, but the also the magnitude of the impacts of these events. Furthermore, this method will also capture the dynamic implications of the events through the return response as we saw with the event study analysis. Thus this extension should allow me to determine not only the effect of a climate policy event realization, but also the magnitude and dynamics of the effects of the climate policy events on the oil price and production of oil.

I implement this method as follows. First, I focus on a particular event to estimate the climate policy risk exposures from exposure to oil price innovations for each sector. Given these estimate exposure measure, I rescale the exposures to sum to one to provide a portfolio weighting scheme. These event-study estimated weights are then used to calculate a factor-mimicking portfolio from the sector portfolios. I focus on the equal weighted portfolios for simplicity. With this factor-mimicking portfolio, I can implement the analysis in two ways. The first is to incorporate the factor mimicking portfolio returns directly into the VAR in place of the climate policy event index as a continuous measure of responses to climate policy events. The second is to use an interaction term of the climate policy event index dummy used originally with the factor-mimicking portfolio that I have constructed. The interaction term will highlight directly the events as given by the climate policy event index, while also incorporating magnitudes through the return measure. This analysis ties the asset pricing and production implications into a single analysis to test the full model implications. I will first focus on the weights generated from the Supreme Court hold on the Clean Power Plan, and for robustness will test portfolio weightings based on different dynamic response times and different events.

Figure 7 shows the cumulative level impulse response functions of oil production and oil prices for a shock to the likelihood of significant climate policy occurring using these alternative climate policy indices. All the plots are for the VAR estimated using the more recent, policy-focused time sample (1996-2017). Each plot include includes the the estimated IRF (red line), the one-standard deviation bootstrapped confidence interval (solid blue line), and the two-standard deviation bootstrapped confidence interval (dashed blue line).

The top two plots show the results for the factor mimicking portfolio returns used as the climate policy risk likelihood measure. Here the results are quite similar to the original ClimPol IRFs. A shock to climate policy leads to an increase in oil production and a decrease in the spot price of oil. Moreover, these impacts grow in magnitude and persist over time. As before, the direction and dynamics of these results are consistent with the key predictions of the model. While again only the impact on the spot price of oil is statistically significant, the significance is even greater in this setting.

The bottom two plots show the results for the factor mimicking portfolio returns interacted with the original ClimPol index used as the climate policy risk likelihood measure. Again, a shock to climate policy leads to an increase in oil production and a decrease in the spot price of oil, and these impacts grow in magnitude and persist over time. There are a number of key differences in this setting. First, the magnitude of the IRFs is greater for the impact on both oil prices and oil production. Second, the impact on the spot price of oil and oil production is statistically significant. This setting shows that previous results that do not account for magnitude likely understate the measured impact from before, as well as show the value of using asset prices to study the impact of climate policy risk.

Figure 7: ClimPol Shock IRF - 1996-2017 Time Sample w/ C.I.s

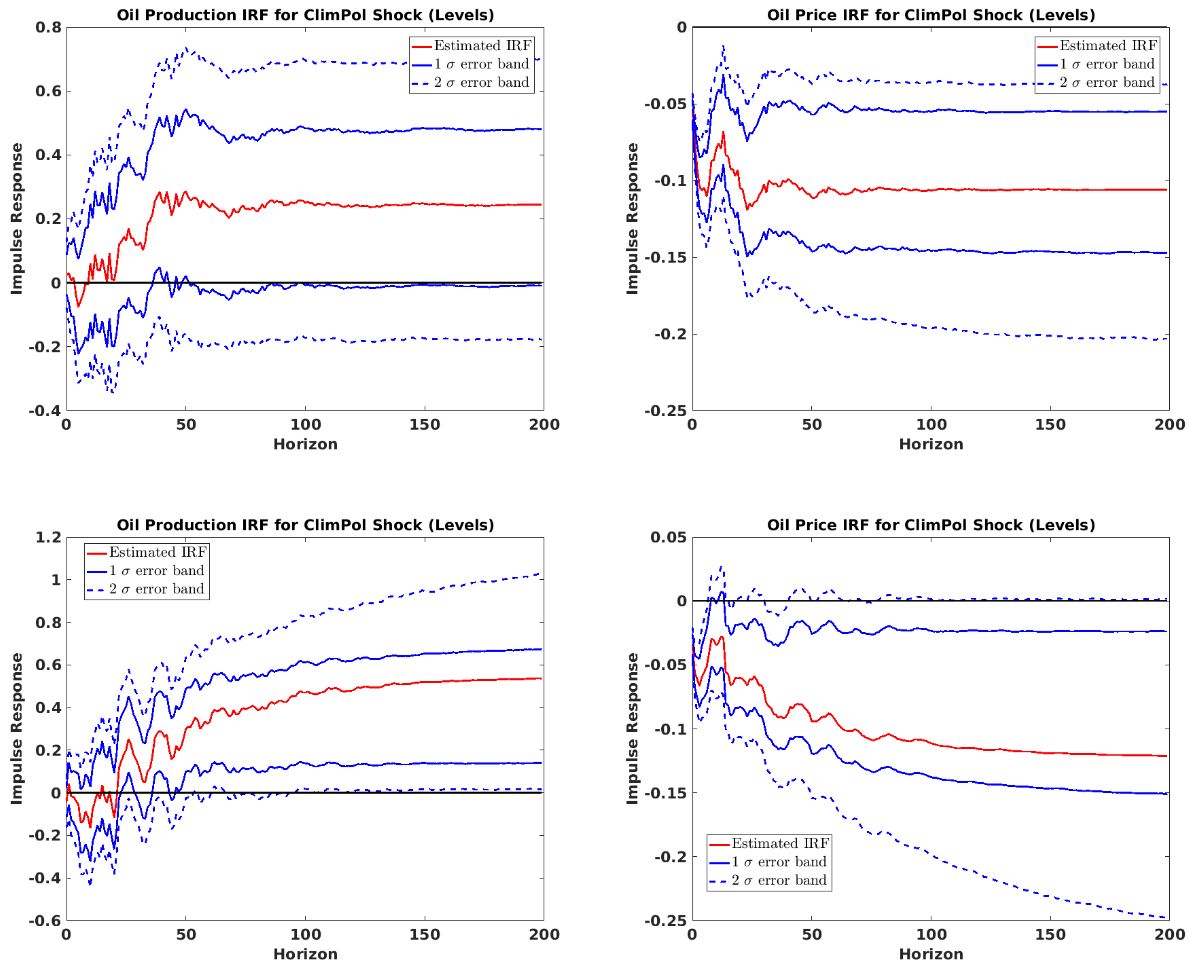


Figure 7 shows the estimated impulse response functions for global oil production, the WTI spot price of oil, real economic activity, and variations of the *ClimPol* climate policy index measure for a shock to the likelihood of climate policy. The first variation uses the returns of the factor mimicking portfolio created from sector portfolios weighted by their normalized climate policy risk exposure estimated value. The second variation uses the returns for this same factor mimicking portfolio, but is interacted with the original ClimPol index. The red line is the estimated IRF, the solid blue lines represent the on-standard deviation error bands, and the blue dashed lines represent the two standard deviation error bands. Error bands are estimated using boostrapping. The estimates use the policy-relevant time subsample (1996-2017). See text for the full VAR specification used and definition of variables. The VAR is estimated using 24 lags.

6 Conclusions

In this paper I have empirically estimated the impact of climate policy risk on real and financial outcomes. The empirical analysis examines the predicted effects of the risk of future climate policy action that could strand fossils fuels that come from a simplified version of the model developed by Barnett (2019). These predictions are that an increased likelihood of future climate policy regulation leads to a dynamic run-up in oil production, a dynamically decreasing spot price of oil due to the increasing supply generated by the run on oil, and significant decreases in the value of oil firms, or firms with high exposure to climate policy risk, because of the decreased expected level of usable oil reserves and the decreased expected value of those usable oil reserves.

I test these model predictions through variety of empirical estimation techniques. I first do an event-study analysis of climate policy events that shift the likelihood of future climate policy actions taking place. I estimate the effect of shifts in the likelihood of future climate policy due to climate policy events by regressing sectors' cumulative abnormal returns after the event on their exposure to climate policy risk, proxied for by exposure to oil price shocks as motivated by the model prediction. I find sectors with the highest climate policy risk exposure experienced the largest increases in cumulative abnormal returns for events that decreased the likelihood of future climate policy action and the largest decreases in cumulative abnormal returns for events that increased the likelihood of future climate policy actions, consistent with the model predictions.

I then construct a "climate policy" event index from realized climate policy, energy sector, and climate-related events to estimate the dynamic impact of changes in climate policy shocks. In estimated reduced-form regressions, I find that increases in the likelihood of major climate policy measured by my index lead to increased global and regional oil production. I also find that positive climate policy shocks lead to increasingly negative returns for the US oil sector and the spot price of oil. Finally, I estimate a structural VAR for the global oil market that includes the climate policy index, and calculate impulse response functions for a shock to climate policy. The results suggest that increases in the likelihood of significant climate change policy leads to long-term and permanent increases in crude oil production and a statistically significant decreases in the oil spot price, consistent with the dynamic predictions of my model. For each index-based empirical test, the statistical and economic significance are greater during the more recent, policy-focused time period (1996-2017) than for the entire available time sample (1973-2017), further validating the temperature dependence of outcomes implied by the model and the dynamic effect of climate policy risk the model predicts.

I leave for future work a number of interesting empirical tests related to my current analysis. In particular, examining the term structure implications of shocks to the likelihood of future climate policy action as observed oil options and oil futures is a particularly interesting area of continued work that would provide insight into the long-run implications of climate policy risk. Exploring the impact of climate policy risk at the firm and country level would provide valuable understanding of the cross-sectional differences in the impacts of this type of risk. Finally, focusing on alternative financial instruments, alternative measures of climate policy risk, and additional real outcomes would provide further depth into the implications of the climate policy risk I focus on here. Each of these would offer additional insight to help us understand the impacts of climate change and the dynamic risk associated with climate policy.

Appendix A Model Details and Derivations

A.1 Households

Households have recursive preferences of the Duffie-Epstein-Zin type, given by

$$h(C,V) \;=\; \rho(1-\xi)V(\log C - \frac{1}{1-\xi}\log((1-\xi)V))$$

where ρ is the subjective discount rate, ξ is risk aversion, V is the value function or continuation value, and C_t is consumption. These preferences allow for the separation of risk aversion and the elasticity of intertemporal substitution (EIS) and incorporate agents' concerns about the resolution of future uncertainty into the decision-making process. Given this preference structure, the household maximizes discounted lifetime utility subject to their budget constraint:

$$\begin{aligned} V &= \max_{C_t} E[\int_0^{\infty} \rho(1-\xi)V(\log C_t - \frac{1}{1-\xi}\log((1-\xi)V))dt] \\ \text{subject to} & \\ C_t &\leq \Pi_t + \mathcal{T}_t \end{aligned}$$

where Π_t are profits from the household-owned firms and $\mathcal{T}_t$ are taxes rebated to households.

A.2 Production

A.2.1 Final Output

The final output firm produces the consumption good using a Cobb-Douglas technology:

$$Y_t = AK_t^{1-\nu_t}O_t^{\nu_t}$$

where A is total factor productivity (TFP), K_t is capital, O_t is oil energy, and ν is the factor input share of oil. Climate policy governs the value of ν_t and I assume, for simplicity, there are only two possible values of ν_t, ν in the pre-policy state and 0 in the post-policy state. When $\nu_t = \nu$ it represents loose climate policy and a high demand for oil in final good production, and when $\nu_t = 0$ it represents strict climate policy where final good production can only be done without oil. The process for policy change is governed by a state-dependent arrival rate characterized in section A.3.

The final output sector is perfectly competitive, and so firms in this sector maximize discounted, expected lifetime profits by optimally choosing investment, labor, and energy inputs subject to state variable evolution, market clearing, and taking prices as given:

$$
\begin{aligned}
V_C &= \max_{O,I} E \int \pi_t(\tilde{Y}_t - P_{I,t}I_t - P_{O,t}O_t)ds \\
\text{subject to} & \\
dK_t &= K_t(\ln B + \delta_1 \ln I_t - \delta_2 \ln K_t)dt + \sigma_K K_t dB_K
\end{aligned}
$$

Note that $\tilde{Y}_t$, used in the firm problem above, is final output after accounting for climate change damages, which I explain in section A.3. The stochastic discount factor (SDF), π_t, provides the necessary discounting across time and states of nature in order to derive firm values, which I derive and elaborate on in section A.6.1. The evolution of the capital stock is subject to a specific case of the adjustment costs used by Jermann (1998) and others, highlighted in recent work by Anderson and Brock (2017), which are empirically indistinguishable from other common forms used in the literature for observable outcomes in the data and allows for tractability when solving the model.

A.2.2 Oil Input

The oil firm produces using the linear technology

$$O_t = N_t$$

where O_t is the oil used for final output production and N_t is oil extracted. Oil firms maximize discounted expected lifetime profits by choosing extraction subject to evolution of state variables and market clearing:

$$
\begin{aligned}
V_O &= \max_{N} E \int \pi_t(P_{O,t}N_t)ds \\
\text{subject to} & \\
dR_t &= (-N_t)dt + \sigma_R R_t dB_R \\
dT_t &= \varphi N_t dt + \sigma_T dB_T
\end{aligned}
$$

Again π_t is the SDF used for discounting firm profits. T_t is atmospheric temperature, discussed in section A.3. The evolution of reserves is determined by extraction of oil and shocks to supply. There are no explicit costs of extraction. However, because oil firms take into account the shadow value of holding reserves, this implicit cost limits extraction done at any given time. I assume a competitive oil sector so that oil firms take prices as given.

A.3 Climate and Climate Policy

Atmospheric temperature in excess of pre-industrial levels evolves as

$$dT_t = \varphi N_t dt + \sigma_T dB_T$$

where φ is the carbon-climate response (CCR) to emissions from oil. This climate process is a stochastic version of the relationship estimated by Matthews et al. (2009), Matthews et al. (2012), and MacDougall and Friedlingstein (2015). I use the "Matthews approximation" in place of more complex climate dynamics for tractability and because of the longer-run nature of the approximation that is designed for climate change fluctuations that I am focusing on in my analysis.

The damage function $D(T_t)$, which captures how climate change impacts economic outcomes, multiplicatively scales final good output. Furthermore, the damage function has the properties $D(T_t) \in [0,1] \forall T_t$, $D(0) = 1$, $D(\infty) = 0$, and $\frac{dD}{dT} < 0$. The functional forms for the damage function and consumable final output are given by

$$D(T_t) = exp(-\eta T_t) \quad \text{and} \quad \tilde{Y}_t = D(T_t) Y_t$$

Change in policy is modeled by a permanent jump in the energy input share of oil, ν_t, which occurs according to a Poisson jump process. This critical feature of the model generates the key mechanism for the results in the model related to the impact of climate policy, which includes an uncertain arrival time and the risk of stranded assets, on oil production, oil prices, and the oil firm value. The arrival rate of the shock to ν_t, or climate policy shock, is given by

$$\lambda_t = \lambda(T_t) = \psi(1 - exp(-\varpi T_t^p))$$

The arrival rate is importantly dependent on the endogenously evolving level of climate change due to emissions generated by oil use. One can interpret this as an increasing likelihood of significant climate policy being enacted as climate change becomes more pronounced either by increases in temperature, or, as functional form is quite similar to the damage function, as observed climate damages increase.

A.4 Planner Outcomes

A.4.1 Climate and Climate Policy

The planner's problem can be solved by the pre-policy First Order Conditions (FOC)

$$\begin{aligned} N &= \frac{\rho(1-\xi)V(1-C_1)^{-1}\nu}{V_R - \varphi_T V_T} \\ I &= \frac{V_K K\delta_1}{\rho(1-\xi)V + V_K K\delta_1} exp(-\eta T)AK^{1-\nu}(N)^{\nu} \end{aligned}$$

the post-policy First Order Conditions (FOC)

$$I = \frac{V_K K\delta_1}{\rho(1-\xi)V + V_K K\delta_1} exp(-\eta T)AK$$

and by guessing and verifying that the pre- and post-policy value functions are given by

$$V_{pre} = K^{c_1}v(R,T) \qquad V_{post} = \tilde{c}_0 K^{c_1} exp(c_3 T)$$

where the coefficients of the value functions are given by

$$\begin{aligned} \tilde{c}_0 &= \frac{1}{1-\xi} exp(\frac{1}{\rho}\{\rho(1-\xi)\log(A(1-C_1)) + c_3^2\frac{1}{2}\sigma_T^2 \\ &\quad +c_1(\log B + \delta_1 \log(AC_1)) + \frac{\sigma_K^2}{2}c_1(c_1-1)\}) \\ c_1 &= \frac{\rho(1-\xi)(1-\nu)}{\rho-(1-\nu)\delta_1+\delta_2} \\ c_3 &= -\frac{\eta(1-\xi)(\rho+\delta_2)}{\rho-\delta_1(1-\nu)+\delta_2} \\ C_1 &= \frac{c_1\delta_1}{\rho(1-\xi)+c_1\delta_1} \end{aligned}$$

The remaining differential equation v solves

$$\begin{aligned} 0 &= \rho(1-\xi)v(\log(exp(-\eta T)A(N)^{\nu}(1-C_1)) \\ &\quad -\frac{1}{(1-\xi)}\log((1-\xi)v)) + v_R(-N) + \varphi N v_T \\ &\quad +vc_3(\ln B + \delta_1 \ln(exp(-\eta T)AC_1)) \\ &\quad +\frac{1}{2}\sigma_R^2 R^2 v_{RR} + \frac{1}{2}\sigma_T^2 v_{TT} + \frac{1}{2}\sigma_K^2 v c_1(c_1-1) + \lambda(T)[\tilde{c}_0 exp(c_3 T) - v] \end{aligned}$$

A.5 Decentralized Outcomes

A.5.1 Household

The household optimization problem is given by

$$\begin{aligned} V &= \max_{C} E\int \rho(1-\xi)V(\log C - \frac{1}{1-\xi}\log(1-\xi)V)dt \\ s.t. \quad & W_t \geq \int \pi C \end{aligned}$$

The SDF is given by

$$\pi = exp(\int h_V)\rho(1-\xi)VC^{-1}$$

and

$$\begin{aligned} h_C &= \rho(1-\xi)VC^{-1} \\ h_J &= \rho(1-\xi)\log C - \rho\log((1-\xi)V) - \rho \end{aligned}$$

A.5.2 Final Output

The final output firm's profit maximization problem is given by

$$\begin{aligned} V_F &= \max_{I,O} E\int \pi(exp(-\eta T)AK^{1-\nu}O^{\nu} - P_I I - P_O O)ds \\ s.t. \quad & dK = K(\ln B + \delta_1 \ln I - \delta_2 \ln K) \end{aligned}$$

The FOC are given by

$$\begin{aligned} P_I &= \lambda_K K\delta_1 \pi^{-1} I^{-1} \\ P_O &= \nu exp(-\eta T)AK^{1-\nu}O^{\nu-1} \end{aligned}$$

Taking the SDF and value function as given, by definition the Langrangian multiplier λ_K is given by the discounted marginal value of another unit of capital, i.e., $\lambda_K = exp(\int h_V)V_K$, and so

$$\frac{c_1\delta_1}{\rho(1-\xi)+c_1\delta_1}\tilde{Y} = I$$

given $h_C = \rho(1-\xi)VC^{-1}$, $V = K^{c_1}v(R,T)$, and $P_I = 1$.

A.5.3 Oil Firm and Optimal Tax

From the oil firm's profit maximization problem, which includes a tax on the oil extraction piece of output only as that is the only piece contributing to emissions, we see

$$\begin{aligned} V_O &= \max_n E \int \pi(P_O R((1-\tau_{opt})n))ds \\ s.t. &\quad dR/R = -ndt + \sigma_R dB \\ &\quad dT = \varphi_T nRdt + \sigma_T dB \end{aligned}$$

The FOC for extraction is given by

$$P_O = \lambda_R R(1-\tau)^{-1}\pi^{-1}$$

Taking P_O, the SDF, and the value function as given previously, and by definition the Langrangian multiplier λ_R is the discounted marginal value of another unit of oil, i.e., $\lambda_R = exp(\int h_V)V_R$. Plugging in these expressions we find

$$n = \frac{\rho(1-\xi)vY(C)^{-1}\nu(1-\tau)}{v_R}$$

Note that the Social Planner's FOC derived from the HJB equation are given by

$$n = \frac{\rho(1-\xi)v(1-C_1)^{-1}\nu}{v_R - \varphi_T v_T}$$

Equating the SP and decentralized FOCs provides a system of equations from which the optimal tax can be derived as

$$(1-\tau_{opt}) = \frac{v_R}{v_R - \varphi_T v_T}$$

A.6 Asset Pricing Outcomes

A.6.1 The Stochastic Discount Factor (SDF)

Note that the intertemporal marginal rate of substitution (IMRS) or stochastic discount factor (SDF) following Duffie and Skiadas (1994) is

$$\pi_t = exp(\int_0^t h_J(C,V)ds)h_C(C,V)$$

As shown by Duffie and Skiadas (1994), Ito's Lemma then gives $\frac{d\pi_t}{\pi_t} = h_J dt + \frac{\mathcal{D}h_C}{h_C}$ and the risk prices are the loadings on the Brownian motions and jump process for the SDF evolution:

$$\begin{aligned}
\sigma_{\pi,K} &= (1-\nu-c_1)\sigma_K \\
\sigma_{\pi,R} &= \{\nu\frac{O_R}{O} - \frac{v_R}{v}\}\sigma_R R \\
\sigma_{\pi,T} &= \{\nu\frac{O_T}{O} - \frac{v_T}{v} - \eta\}\sigma_T \\
\Theta_\pi &= \{1 - \frac{v_{post}\tilde{Y}_{post}^{-1}}{v_{pre}\tilde{Y}_{pre}^{-1}}\}
\end{aligned}$$

A.6.2 Firm Prices

To derive firm prices, I apply the envelope theorem to the social planner's Lagrangian. This follows the methodology used by Papanikolaou (2011) for example. Note for the final output firm we have

$$\begin{aligned}
\pi_t S_t^C &= E_t \int_t^\infty \pi_s (exp(-\eta T) A K^{1-\nu} O^\nu - P_O O - I) ds \\
&= E_t \int_t^\infty \pi_s (exp(-\eta T) A K^{1-\nu} O^\nu (1-\nu) - i^* K) ds \\
&\Longrightarrow \\
S_t^C &= E_t \int_t^\infty exp(\int_t^s h_V) \frac{h_{C,s}}{h_{C,t}} (exp(-\eta T) A K^{1-\nu} O^\nu (1-\nu) - i^* K) ds
\end{aligned}$$

For the oil firm, plugging in the socially optimal choices, we have

$$\begin{aligned}
\pi_t S_t^O &= E_t \int_t^\infty \pi_s (P_O R((1-\tau)n - i_R) ds \\
&\Longrightarrow \\
S_t^O &= E_t \int_t^\infty exp(\int_t^s h_V) \frac{h_{C,s}}{h_{C,t}} P_O R(n^* - i_R^*) ds
\end{aligned}$$

Note the Lagrangian for the social planner's problem is given by

$$\mathcal{L} = E_t \int_t^\infty \{h(C,V) - \pi_s(C - exp(-\eta T) A K^{1-\nu} O^\nu + iK) - P_O \pi_s (O - nR)\} ds$$

Therefore, by application of the envelope theorem we know that

$$\frac{\partial \mathcal{L}}{\partial K} = \frac{\partial V}{\partial K} \quad , \quad \frac{\partial \mathcal{L}}{\partial R} = \frac{\partial V}{\partial R}$$

Furthermore, we also know that

$$\frac{\partial K_s}{\partial K_t} K_t = K_s \quad , \quad \frac{\partial R_s}{\partial R_t} R_t = R_s$$

Calculating derivatives of Lagrangian and comparing I find that

$$
\begin{aligned}
S_t^C &= \frac{1}{h_C}\frac{\partial V}{\partial K}K = c_1\frac{(1-C_1)}{\rho(1-\xi)}\tilde{Y}_t \\
S_t^O &= \frac{1}{h_C}\frac{\partial V}{\partial R}R = \frac{(1-C_1)}{\rho(1-\xi)}\frac{v_R R}{v}\tilde{Y}_t
\end{aligned}
$$

Therefore, the firm prices for each sector in the model are given by

$$
S_t^C = c_1\frac{(1-C_1)}{\rho(1-\xi)}\tilde{Y}_t, \quad S_t^O = \frac{(1-C_1)}{\rho(1-\xi)}\frac{v_R R}{v}\tilde{Y}_t
$$

Appendix B Climate Policy Index Details

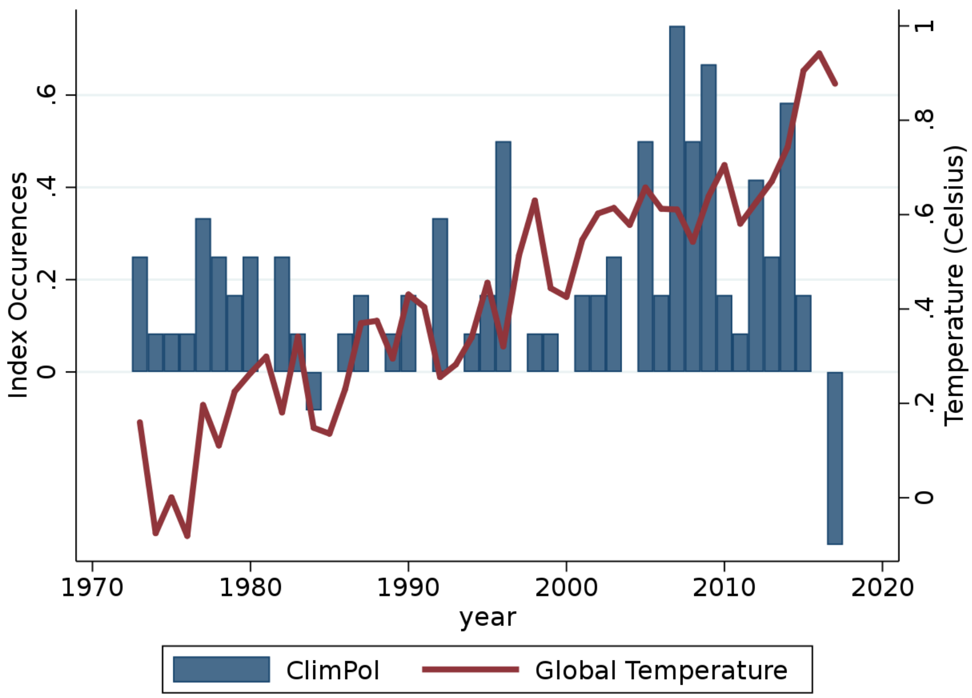
Figure 8: Climate Policy Index List, 1996-Present

Date	Event	Shock Sign	Source
5-Jun-96	Solar Two Plant Demonstrates Low Cost Method of Storing Solar Energy	+	ProCon.org
18-Jul-96	COP 2, Geneva, Switzerland	+	IPCC
9-Oct-96	Hydrogen Future Act of 1996 Is Passed to Further Expand Hydrogen Power Development	+	ProCon.org
29-Oct-96	European Union adopts target of a maximum 2 °C rise in average global temperature	+	Wikipedia
5-Nov-96	Bill Clinton Elected POTUS	+	U.S. Presidential Elections
5-Dec-96	EV1 Electric Car Is Made Available to the Public For Lease; Lease Program and EV1 Later Dismantled by GM	+	ProCon.org
25-Jun-97	US Senate passes Byrd–Hagel Resolution rejecting Kyoto	-	Wikipedia
11-Dec-97	COP 3, The Kyoto Protocol on Climate Change	+	Wikipedia/IPCC
14-Nov-98	COP 4, Buenos Aires, Argentina	+	IPCC
5-Nov-99	COP 5, Bonn, Germany	+	IPCC
7-Nov-00	George W. Bush Elected POTUS	-	U.S. Presidential Elections
25-Nov-00	COP 6, The Hague, Netherlands	+	IPCC
28-Mar-01	President George W. Bush withdraws from the Kyoto negotiations	-	Wikipedia
27-Jul-01	COP 6, Bonn, Germany	+	IPCC
29-Sep-01	IPCC Third assessment report	+	IPCC
10-Nov-01	COP 7, Marrakech, Morocco	+	IPCC
13-May-02	Farm Security and Rural Investment Act	+	Wikipedia
1-Nov-02	COP 8, New Delhi, India	+	IPCC
6-Feb-03	President Bush Unveils the Hydrogen Fuel Initiative	+	ProCon.org
27-Feb-03	Plans Announced to Build Worlds First Zero Emissions Coal Power Plant	+	ProCon.org
12-Dec-03	COP 9, Milan, Italy	+	IPCC
2-Nov-04	George W. Bush Elected POTUS	-	U.S. Presidential Elections
17-Dec-04	COP 10, Buenos Aires, Argentina	+	IPCC
1-Jan-05	EU Emissions Trading Scheme is launched, the first such scheme	+	Wikipedia
16-Feb-05	Kyoto Protocol comes into force (not including the US or Australia)	+	Wikipedia
8-Jul-05	31st G8 summit discusses climate change, relatively little progress made	+	Wikipedia
8-Aug-05	Energy Policy Act	+	Wikipedia
9-Nov-05	US House Prevents Drilling for Oil in the Arctic National Wildlife Refuge	+	ProCon.org
9-Dec-05	COP 11/CMP 1, Montreal, Canada	+	Wikipedia/IPCC
30-Oct-06	The Stern Review is published	+	Wikipedia
17-Nov-06	COP 12/CMP 2, Nairobi, Kenya	+	IPCC
16-Feb-07	February 2007 Washington Declaration	+	IPCC
7-Jun-07	33rd G8 summit	+	IPCC
31-Jul-07	2007 UN General Assembly plenary debate	+	IPCC
3-Aug-07	September 2007 Washington conference	+	IPCC
31-Aug-07	2007 Vienna Climate Change Talks and Agreement	+	IPCC
24-Sep-07	September 2007 United Nations High-Level-Event	+	IPCC
17-Nov-07	IPCC Fourth assessment report	+	IPCC/ProCon.org
17-Dec-07	COP 13/CMP 3, Bali, Indonesia	+	IPCC
19-Dec-07	Energy Independence and Security Act	+	Wikipedia
30-Jan-08	First Commercial Cellulosic Ethanol Plant Goes Into Production	+	ProCon.org
22-May-08	Food, Conservation, and Energy Act	+	Wikipedia
7-Oct-08	National Biofuel Action Plan Unveiled	+	ProCon.org
4-Nov-08	Barack Obama Elected POTUS	+	U.S. Presidential Elections

12-Dec-08	COP 14/CMP 4, Poznań, Poland	+	IPCC
22-Dec-08	Worst Coal Ash Spill in US History in Kingston, Tennessee	+	ProCon.org
17-Feb-09	ARRA (2009) Contains Funding for Renewable Energy	+	ProCon.org/Wikipedia
22-Apr-09	First Framework for Wind Energy Development on the US Outer Continental Shelf Announced	+	ProCon.org
5-May-09	President Obama Issues Presidential Directive to USDA to Expand Access to Biofuels	+	ProCon.org
27-May-09	US Announces Funding in Recovery Act Funding for Solar and Geothermal Energy Development	+	ProCon.org
26-Jun-09	US House of Representatives passes the American Clean Energy and Security Act (Waxman)	+	Wikipedia
22-Sep-09	September 2009 United Nations Secretary General's Summit on Climate Change	+	IPCC
27-Oct-09	US Invests $3.4 Billion to Modernize Energy Grid	+	ProCon.org
18-Dec-09	COP 15/CMP 5, Copenhagen, Denmark	+	IPCC
20-Apr-10	BP Oil Rig Explodes & Causes Largest Oil Spill in US History	+	ProCon.org
10-Dec-10	COP 16/CMP 6, Cancún, Mexico	+	IPCC
11-Mar-11	Earthquake off Coast of Japan Damages Six Powerplants at Fukushima	+	ProCon.org
1-Sep-11	Solar Power Company Solyndra Declares Bankruptcy	-	ProCon.org
9-Dec-11	COP 17/CMP 7, Durban, South Africa	+	IPCC
9-Feb-12	US Nuclear Regulatory Commission (NRC) Approves New Nuclear Power Plants	+	ProCon.org
27-Mar-12	EPA Announces First Clean Air Act Standard for Carbon Pollution from New Power Plants	+	ProCon.org
17-Apr-12	EPA Issues First Ever Clean Air Rules for Natural Gas Produced by Fracking	+	ProCon.org
6-Nov-12	Barack Obama Elected POTUS	+	U.S. Presidential Elections
7-Dec-12	COP 18/CMP 8, Doha, Qatar	+	IPCC
25-Jun-13	President Obama Releases His Climate Action Plan	+	ProCon.org
20-Sep-13	EPA Issues New Proposed Rule to Cut Greenhouse Gas Emissions from Power Plants	+	ProCon.org
23-Nov-13	COP 19/CMP 9, Warsaw, Poland	+	IPCC
13-Feb-14	Ivanpah, the World's Largest Concentrated Solar Power Generation Plant, Goes Online	+	ProCon.org
9-May-14	President Obama Announces Solar Power Commitments and Executive Actions	+	ProCon.org
2-Jun-14	EPA Proposes First Ever Rules to Reduce Carbon Emissions from Existing Power Plants	+	ProCon.org
22-Sep-14	Rockefellers and over 800 Global Investors Announce Fossil Fuel Divestment	+	ProCon.org
23-Sep-14	Climate Summit 2014	+	IPCC
1-Nov-14	IPCC Fifth assessment report	+	IPCC
12-Dec-14	COP 20/CMP 10, Lima, Peru	+	IPCC
3-Aug-15	President Obama Announces Clean Power Plan (finalized Oct. 23, 2015; Active December, 22, 2015)	+	ProCon.org
12-Dec-15	COP 21/CMP 11, Paris, France	+	Wikipedia/IPCC
8-Nov-16	Donald Trump Elected POTUS	-	U.S. Presidential Elections
18-Nov-16	COP 22/CMP 12/CMA 1, Marrakech, Morocco	+	IPCC
28-Mar-17	President Trump Signs Executive Order to Begin Reversal of President Obama' Clean Power Plan	-	ProCon.org
1-Jun-17	President Donald Trump withdraws the United States from the Paris Agreement	-	Wikipedia
31-Jul-17	Two Nuclear Power Reactors in South Carolina Abandoned Before Construction Completed	-	ProCon.org
22-Dec-17	Tax Bill Opens Arctic National Wildlife Refuge for Oil Drilling	-	ProCon.org
9-May-18	Solar Power to Be Required on All New California Homes by 2020	+	ProCon.org

*Events come from ProCon.org Fossil Fuel and Alternative Energy timeline, IPCC/UNFCC Meetings, U.S. Presidential Election outcomes, and Wikipedia.org Selective historical timeline of significant climate change political events and List of United States energy acts

Figure 9: ClimPol Index and Annual Global Mean Temperature



References

Acemoglu, D., U. Akcigit, D. Hanley, and W. Kerr (2016). Transition to clean technology. *Journal of Political Economy 124*(1), 52–104.

Anderson, E. W. and W. A. Brock (2017). Logarithmic depreciation. *SSRN Workin Paper 1*(1), 1.

Anderson, E. W., W. A. Brock, and A. Sanstad (2016). Robust consumption and energy decisions.

Bansal, R., M. Ochoa, and D. Kiku (2016). Climate change and growth risks. Technical report, National Bureau of Economic Research.

Barnett, M. (2017). Climate change and uncertainty: An asset pricing perspective. *Working Paper*.

Barnett, M. (2019). A run on oil: Climate policy, stranded asets, and asset prices. *Working Paper*.

Barnett, M., W. Brock, and L. P. Hansen (2018). Pricing uncertainty induced by climate change. *Working Paper. In-principle acceptance of RFS Climate Finance registered report, completion in progress.*.

Baumeister, C. and J. D. Hamilton (2019). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. *American Economic Review 109*(5), 1873–1910.

Baumeister, C. J. and J. D. Hamilton (2017). Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks. Technical report, National Bureau of Economic Research.

Belo, F., V. D. Gala, and J. Li (2013). Government spending, political cycles, and the cross section of stock returns. *Journal of Financial Economics 107*(2), 305–324.

Bornstein, G., P. Krusell, and S. Rebelo (2017). Lags, costs, and shocks: An equilibrium model of the oil industry. Technical report, National Bureau of Economic Research.

Brock, W. A. and L. P. Hansen (2017). Wrestling with uncertainty in climate economic models.

Burke, M., W. M. Davis, and N. S. Diffenbaugh (2018). Large potential reduction in economic damages under un mitigation targets. *Nature 557*(7706), 549.

Cai, Y., K. L. Judd, and T. S. Lontzek (2015). The Social Cost of Carbon with Economic and Climate Risks. *Working Paper*.

Carlson, M., Z. Khokher, and S. Titman (2007). Equilibrium exhaustible resource price dynamics. *The Journal of Finance 62*(4), 1663–1703.

Casassus, J., P. Collin-Dufresne, and B. R. Routledge (2009). Equilibrium commodity prices with irreversible investment and non-linear technologies.

Crost, B. and C. P. Traeger (2011). Risk and aversion in the integrated assessment of climate change.

Dasgupta, P. and G. Heal (1974). The optimal depletion of exhaustible resources. *The review of economic studies 41*, 3–28.

David, A. (2015). Exploration activity, long run decisions, and the risk premium in energy futures.

Dell, M., B. F. Jones, and B. A. Olken (2012). Temperature shocks and economic growth: Evidence from the last half century. *American Economic Journal: Macroeconomics 4*(3), 66–95.

Deschenes, O. and M. Greenstone (2007). The economic impacts of climate change: evidence from agricultural output and random fluctuations in weather. *The American Economic Review 97*(1), 354–385.

Dietz, S., C. Gollier, and L. Kessler (2017). The climate beta. *Journal of Environmental Economics and Management*.

Duffie, D. and C. Skiadas (1994). Continuous-time security pricing: A utility gradient approach. *Journal of Mathematical Economics 23*(2), 107–131.

Engle, R. F., S. Giglio, H. Lee, B. T. Kelly, and J. Stroebel (2019). Hedging climate change news. *Available at SSRN 3317570*.

Giglio, S., M. Maggiori, J. Stroebel, and A. Weber (2015). Climate change and long-run discount rates: Evidence from real estate. Technical report, National Bureau of Economic Research.

Golosov, M., J. Hassler, P. Krusell, and A. Tsyvinski (2014). Optimal Taxes on Fossil Fuel in General Equilibrium. *Econometrica 82*(1), 41–88.

Hambel, C., H. Kraft, and E. Schwartz (2015). Optimal carbon abatement in a stochastic equilibrium model with climate change. Technical report, National Bureau of Economic Research.

Hamilton, J. D. (2005). Oil and the macroeconomy. *The New Palgrave Dictionary of Economics 2*.

Hamilton, J. D. (2008). Understanding crude oil prices. Technical report, National Bureau of Economic Research.

Hong, H., F. W. Li, and J. Xu (2016). Climate risks and market efficiency. Technical report, National Bureau of Economic Research.

Hotelling, H. (1931). The economics of exhaustible resources. *Journal of political Economy 39*(2), 137–175.

Hsiang, S., R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, et al. (2017). Estimating economic damage from climate change in the united states. *Science 356*(6345), 1362–1369.

Jermann, U. J. (1998). Asset pricing in production economies. *Journal of Monetary Economics 41*(2), 257–275.

Kelly, B., L. Pástor, and P. Veronesi (2016). The price of political uncertainty: Theory and evidence from the option market. *The Journal of Finance.*

Kelly, D. L. and C. D. Kolstad (1999). Bayesian learning, growth, and pollution. *Journal of economic dynamics and control 23*(4), 491–518.

Kilian, L. (2008). Exogenous oil supply shocks: how big are they and how much do they matter for the us economy? *The Review of Economics and Statistics 90*(2), 216–240.

Kilian, L. (2009). Not all oil price shocks are alike: Disentangling demand and supply shocks in the crude oil market. *American Economic Review 99*(3), 1053–69.

Kilian, L. and C. Park (2009). The impact of oil price shocks on the us stock market. *International Economic Review 50*(4), 1267–1287.

Kogan, L., D. Livdan, and A. Yaron (2009). Oil futures prices in a production economy with investment constraints. *The Journal of Finance 64*(3), 1345–1375.

Koijen, R. S., T. J. Philipson, and H. Uhlig (2016). Financial health economics. *Econometrica 84*(1), 195–242.

Kotlikoff, L. J., A. Polbin, and A. Zubarev (2016). Will the paris accord accelerate climate change? Technical report, National Bureau of Economic Research.

Lemoine, D. M. and C. P. Traeger (2012). Tipping points and ambiguity in the economics of climate change. Technical report, National Bureau of Economic Research.

MacDougall, A. H. and P. Friedlingstein (2015). The origin and limits of the near proportionality between climate warming and cumulative co2 emissions. *Journal of Climate 28*(10), 4217–4230.

Matthews, H. D., N. P. Gillett, P. A. Stott, and K. Zickfeld (2009). The proportionality of global warming to cumulative carbon emissions. *Nature 459*(7248), 829–832.

Matthews, H. D., S. Solomon, and R. Pierrehumbert (2012). Cumulative carbon as a policy framework for achieving climate stabilization. *Phil. Trans. R. Soc. A 370*(1974), 4365–4379.

McGlade, C. and P. Ekins (2015). The geographical distribution of fossil fuels unused when limiting global warming to 2 degrees c. *Nature 517*(7533), 187–190.

Mertens, K. and M. O. Ravn (2014). A reconciliation of svar and narrative estimates of tax multipliers. *Journal of Monetary Economics 68*, S1–S19.

Nordhaus, W. (2014). *A question of balance: Weighing the options on global warming policies*. Yale University Press.

Papanikolaou, D. (2011). Investment shocks and asset prices. *Journal of Political Economy 119*(4), 639–685.

Pástor, L. and P. Veronesi (2009). Technological revolutions and stock prices. *American Economic Review 99*(4), 1451–83.

Pastor, L. and P. Veronesi (2012). Uncertainty about government policy and stock prices. *The Journal of Finance 67*(4), 1219–1264.

Pindyck, R. S. and N. Wang (2013). The Economic and Policy Consequences of Catastrophes. *American Economic Journal: Economic Policy 5*(4), 306–339.

Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *The Quarterly Journal of Economics 126*(1), 1–50.

Ready, R. (2015). Oil Consumption, Economic Growth, and Oil Futures: A Fundamental Alternative to Financialization. *Working Paper*.

Ricke, K. L. and K. Caldeira (2014). Maximum warming occurs about one decade after a carbon dioxide emission. *Environmental Research Letters 9*(12), 124002.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Santa-Clara, P. and R. Valkanov (2003). The presidential puzzle: Political cycles and the stock market. *The Journal of Finance 58*(5), 1841–1872.

Sialm, C. (2006). Stochastic taxation and asset pricing in dynamic general equilibrium. *Journal of Economic Dynamics and Control 30*(3), 511–540.

Sinn, H.-W. (2007). Public policies against global warming. Technical report, National Bureau of Economic Research.

Stern, N. (2007). *The Stern Review: The economics of climate change*. Cambridge University Press.

Zickfeld, K. and T. Herrington (2015). The time lag between a carbon dioxide emission and maximum warming increases with the size of the emission. *Environmental Research Letters 10*(3), 031001.