Supply, Demand, and Risk Premiums in Electricity Markets

Kris Jacobs Yu Leo Li Craig Pirrong

University of Houston

March 1, 2017

Abstract

We model the impact of supply and demand variables on the price of electricity futures in a no-arbitrage model, using daily data between 2003 and 2014. The model allows for unspanned economic risk, which is captured by the supply and demand variables but not identified by the futures price. The model provides a satisfactory fit as well as a consistent framework to study the interactions between the electricity futures and the demand and supply variables. We characterize the risk premium implied by the model and identify the risk premium components associated with demand and supply. The unspanned risk premium associated with supply is highly time-varying and constitutes the most important component of the total risk premium embedded in electricity futures. This risk premium becomes negative after 2010.

JEL Classification: G12, G13, Q02

Keywords: Electricity Futures; Economic Determinants; Supply; Demand; Risk Premium; Unspanned Risk.

1 Introduction

Because of the non-storability of the electricity, researchers have long been interested in what are the impacts of economic variables on the electricity price dynamic as well as on the risk premiums.

This paper uses a no-arbitrage model to quantify the impact of economic variables on the electricity prices. Our model features in the unspanned economic risk, which is captured by the supply and demand variables but not identified by the futures price. We use this model to disentangle the risk premium that associated with the electricity price and the risk premium that associated with the economic variables.

Our analysis reveals several new findings. We first document that economic variables do have useful information about the risk premium and this information is not captured by the electricity futures. Particularly, we find that the supply variable strongly predicts future returns. Second, we find that while the economic variables contain additional information about the risk premium, the futures curve contains most information about the futures pricing. Specifically, we show that using the first two principal components can achieve a decent fit about the term structure of electricity futures. Third, by analyzing the risk premium implied by our model, we find that our estimated risk premiums are highly time-varying and closely related with the property of the electricity spot prices. Consistent with the theoretical model proposed by Bessembinder and Lemmon (2002), our estimated spot premium is positively correlated with the spot price volatility and negatively correlated with the spot price skewness. Finally, we find that the unspanned risk premium associated with the supply variable contributes most to the total risk premium embedded in electricity futures. This part becomes to be negative after 2010.

Our paper is related with three strands of literature. We first connect to the studies which uses the reduced form models to study the electricity price modeling and derivative pricing (see, for example Lucia and Schwartz (2002), Cartea and Figueroa (2005), Deng and Oren (2006), Geman and Roncoroni (2006), Benth et al. (2007), Bierbrauer et al.

(2007), Benth et al. (2008), Geman (2009)). This strand of literature focuses on how to build statistical models to not only capture the properties of electricity spot price but also yields to convenient derivative pricing formulas. Our model also belongs to this type but we augment it with the economic variables. We find that the economic variables are important in explaining the risk premium associated with the electricity futures. We also connect to the literature which use the structural approach to price the electricity derivatives (Pirrong and Jermakyan (2008), Cartea and Villaplana (2008), Pirrong (2011)). These papers use a bottom-up approach by first specifying the dynamic for the economic variables and then deriving the spot price as a function of those variables. This approach exploits the information contained in the economic variable and is consistent with the economic theory. We contribute to this strand of literature by showing while the economic variables are important in explaining the risk premium, latent factors could improve the model fits significantly. Finally, we connect to the literature which studies the risk premium of electricity futures from the equilibrium perspective (Bessembinder and Lemmon (2002), Longstaff and Wang (2004), Dong and Liu (2007), Douglas and Popova (2008), Bunn and Chen (2013)). We find that the implied risk premium in our models are consistent with the economic theory.

The remainder of the paper proceeds as follows. Section 2 outlines the spanned and unspanned models and Section 3 discusses model estimation. Section 4 discusses the data. Section 5 presents the empirical results, and Section 6 concludes.

2 Model Specification

The economy is described by a N by 1 state vector X_t, which contains demand and supply variables as well as latent factors. X_t is assumed to follow a Gaussian VAR under the P measure,

$$X_{t+1} = K_{0X}^P + K_{1X}^P X_t + \Sigma \epsilon_{t+1}^P \tag{1}$$

where $\epsilon^P_{t+1} \sim N(0, I_N)$, Σ is a N by N lower triangular matrix, and I_N is a N by N identity matrix. The spot price s_t is an affine function of the state vector X_t. We now discuss four models that are nested in this general specification.

2.1 The Unspanned Model

The unspanned model assumes that only part of the information in X_t is spanned by the futures prices. Consider the N_L by 1 latent pricing vector L_t which fully determines futures prices. In the unspanned model, L_t spans only part of X_t. The other part of X_t which is not reflected in L_t is the so-called unspanned part, which we denote by UM_t. Thus we have $X_t = L_t \cup UM_t$ and the P dynamic can be re-written as

$$\begin{bmatrix} L_{t+1} \\ UM_{t+1} \end{bmatrix} = K^P_{0X} + K^P_{1X} \begin{bmatrix} L_t \\ UM_t \end{bmatrix} + \Sigma \epsilon^P_{t+1} \tag{2}$$

Under the the Q dynamic, the dynamics for the pricing factor L_t are given by

$$L_{t+1} = K^Q_{0L} + K^Q_{1L} L_t + \Sigma_1 \epsilon^Q_{t+1} \tag{3}$$

where $\epsilon^Q_t \sim N(0, I_{N_L})$, Σ_1 is the N_L by N_L upper-left submatrix of Σ, and I_{N_L} is a N_L by N_L identity matrix.

The stochastic discount factor (SDF) is given by

$$e^{\Lambda'_\epsilon \epsilon_{t+1}} = e^{(\Lambda_0 + \Lambda_1 X_t)' \epsilon_{t+1}} \tag{4}$$

Our focus is on the interaction between demand/supply variables and electricity futures prices. Let M_t denote the vector of demand and supply variables, which consists of a spanned and unspanned part, i.e. $M_t = SM_t + UM_t$, where SM_t is the spanned part and

UM_t is the unspanned part. It is possible to rotate UM_t to M_t and write the final version of the unspanned model as

$$X_t = \begin{bmatrix} L_t \\ M_t \end{bmatrix} \tag{5}$$

$$\begin{bmatrix} L_{t+1} \\ M_{t+1} \end{bmatrix} = X_{t+1} = K_{0X}^P + K_{1X}^P X_t + \Sigma \epsilon_{t+1}^P \tag{6}$$

$$L_{t+1} = K_{0L}^Q + K_{1L}^Q L_t + \Sigma_1 \epsilon_{t+1}^Q \tag{7}$$

$$s_t = \rho_{0X} + \rho_{1X} X_t \tag{8}$$

$$e^{\Lambda_\epsilon' \epsilon_{t+1}} = e^{(\Lambda_0 + \Lambda_1 X_t)' \epsilon_{t+1}} \tag{9}$$

2.2 The Spanned Model with Latent Factors

In contrast to the unspanned model, spanned models with latent factors assume that the information in the economy can be fully spanned by the information in futures prices, i.e.

$$X_t \equiv L_t \tag{10}$$

This approach nests traditional reduced-form models which use only latent factors to price futures. In general these models can be written as

$$L_{t+1} = K_{0L}^{Q} + K_{1L}^{Q} L_t + \Sigma_L \epsilon_{t+1}^{Q} \tag{11}$$

$$L_{t+1} = K_{0L}^{P} + K_{1L}^{P} L_t + \Sigma_L \epsilon_{t+1}^{P} \tag{12}$$

$$s_t = \rho_{0L} + \rho_{1L} L_t \tag{13}$$

$$e^{\Lambda_\epsilon' \epsilon_{t+1}} = e^{(\Lambda_0 + \Lambda_1 L_t)' \epsilon_{t+1}} \tag{14}$$

where $\epsilon_t^P \sim N(0, I_{N_L})$, $\epsilon_t^Q \sim N(0, I_{N_L})$, and Σ_L is the N_L by N_L lower triangular matrix.

In our empirical work we implement this general representation of a spanned model with two latent factors. Existing latent factor models can be obtained by imposing restrictions on the dynamics of the state variables in equations 11-14. To provide additional perspective, we also estimate a simple example of an existing spanned model, the two-factor model of Lucia and Schwartz (2002). Because this is a model for log prices and it is formulated in continuous time, the comparison with the unspanned model is made somewhat harder, but this is compensated by the fact that the model is widely known and used. This model is specified as follows:

$$ln(S_t) = X_t + \epsilon_t \tag{15}$$

$$\begin{aligned} dX_t &= -\kappa X_t dt + \sigma_X dz_X^P \\ dX_t &= (-\kappa X_t - \lambda_X) dt + \sigma_X dz_X^Q \end{aligned} \tag{16}$$

$$\begin{aligned} d\epsilon_t &= \mu_\epsilon dt + \sigma_\epsilon dz_\epsilon^P \\ d\epsilon_t &= (\mu_\epsilon - \lambda_\epsilon)dt + \sigma_\epsilon dz_\epsilon^Q \end{aligned} \tag{17}$$

$$\begin{aligned} dz_X dz_\epsilon &= \rho_{x\epsilon} dt \\ dz_X^Q dz_\epsilon^Q &= \rho_{x\epsilon} dt \end{aligned} \tag{18}$$

where $dz_X^P, dz_X^Q, dz_\epsilon^P, dz_\epsilon^Q$ are all standard Brownian motions.

2.3 The Spanned Model with Demand and Supply Variables

Alternatively, we can exclusively use demand and supply variables to price futures. The resulting model is related to the framework of Pirrong and Jermakyan (2008), in which all the information in the demand and supply variables is spanned by futures, i.e.

$$X_t \equiv M_t \tag{19}$$

This type of model can be written in general as

$$M_{t+1} = K_{0M}^Q + K_{1M}^Q M_t + \Sigma_M \epsilon_{t+1}^Q \tag{20}$$

$$M_{t+1} = K_{0M}^P + K_{1M}^P M_t + \Sigma_M \epsilon_{t+1}^P \tag{21}$$

$$s_t = \rho_{0M} + \rho_{1M} M_t \tag{22}$$

$$e^{\Lambda'_\epsilon \epsilon_{t+1}} = e^{(\Lambda_0 + \Lambda_1 M_t)' \epsilon_{t+1}} \tag{23}$$

where $\epsilon_t^P \sim N(0, I_{N_M})$, $\epsilon_t^Q \sim N(0, I_{N_M})$, and Σ_M is the N_M by N_M lower triangular matrix. N_M is the number of variables/factors in M_t.

2.4 Futures Pricing

In all models discussed in this section, futures prices can be derived analytically. Take the unspanned model for example. Denoting the log price of the futures contract with maturity j at time t as $f_t^j = log(F_t^j)$, we have

$$f_t^j = A_j + B_j X_t \tag{24}$$

where

$$A_j = A_{j-1} + B_{j-1} K_{0L}^Q + \frac{1}{2} B_{j-1} \Sigma_L \Sigma'_L B'_{j-1} \tag{25}$$

$$B_j = B_{j-1}[I_L + K_{1L}^Q] \tag{26}$$

with $A_0 = \rho_{0X}$ and $B_0 = \rho_{1X}$. The pricing formulas for the spanned models have a similar structure, but involve different state variables and parameter matrices.

3 Model Estimation

Motivated by the existing literature, we choose one demand variable, one supply variable, and two latent factors in the unspanned model. To benchmark the performance of the unspanned model, we also estimate a spanned model with two latent factors, a spanned model with one supply and one demand variable and the Lucia-Schwartz two factor model. We now discuss the estimation procedure for each of these models.

3.1 The Unspanned Model

Estimating a unspanned model is in general not straightforward for at least two reasons. First, the risk-neutral state variables are latent. Second, the unspanned model has a large number of parameters. For example, a model with two latent factors and two economic variables has 36 free parameters. Because of these two concerns, traditional methods like the Kalman filter may not yield good results. We instead follow the estimation method proposed by Joslin et al. (2014).

We start by estimating the state variables. Joslin et al. (2014) show that under realistic assumptions, one can use the principal components (PCs) of futures prices as the state variables. Next, we estimate the physical dynamic of the state variables in equation 2, i.e. K_{0L}^{P}, K_{1L}^{P}, and Σ. Because the state variables and the economic variables are both observed, the coefficients under the P dynamic can simply be estimated using vector autoregression. Finally, we estimate the Q dynamic of the state variables in equation 3 with the Kalman filter. Given observable futures prices and state variables, the measurement equations are

$$f_t^j = \hat{f}_t^j + \eta_j, \quad j = 1/30, 1, 2, ..., 12 \tag{27}$$

where f_t^j is the log futures price of contract j at time t and $\eta_j \sim N(0, \sigma_{1j})$. The fitted log futures price is given by

$$\hat{f}_t^j = A_j + B_j PC_t \tag{28}$$

with

$$A_j = A_{j-1} + B_{j-1} K_{0L}^{Q} + \frac{1}{2} B_{j-1} \Sigma_L \Sigma_L' B_{j-1}' \tag{29}$$

$$B_j = B_{j-1}[I_L + K_{1L}^{Q}], \quad j = 1/30, 1, ..., 12 \tag{30}$$

where $A_0 = \rho_{0L}$, $B_0 = \rho_{1L}$, and Σ_L is the 2 by 2 upper-left submatrix of Σ.

The transition equation is

$$PC_{t+1} = K_{0L}^{Q} + K_{1L}^{Q} PC_t + \Sigma_L \epsilon_{t+1}^{P} \tag{31}$$

The parameters to be estimated are ρ_{0L}, ρ_{1L}, K_{0L}^{Q}, K_{1L}^{Q}.

3.2 The Spanned Model with Latent Factors

Because the spanned model with latent factors has the same Q dynamic as unspanned model, we once again use the technique in Joslin et al. (2014). Specifically, we use the first two PCs to estimate both the Q and P dynamic as in equation 11 and 12. The Q dynamic is estimated by Kalman filter with the measurement and transition equations 27 and 31. The P dynamic is estimated by vector autoregressive techniques.

3.3 The Spanned Model with Demand and Supply Variables

In the spanned model with demand and supply variables, the economic variables are used to price futures. The model specification is provided in equation 20 to 23. We use the extended Kalman filter to estimate this model. The measurement equations are

$$\begin{aligned} f_t^j &= \hat{f}_t^j + \eta_j, \quad j = 1/30, 1, 2, ..., 12 \\ M_t &= \hat{M}_t + \eta_2 \end{aligned} \tag{32}$$

where $\eta_j \sim N(0, \sigma_{1j})$, $\eta_2 \sim N(0, \sigma_{2j})$ and η_j and η_2 are assumed to be independent.

The fitted futures price is given by

$$\hat{f}_t^j = A_j + B_j \hat{M}_t \tag{33}$$

with

$$A_j = A_{j-1} + B_{j-1}K_{0M}^{Q} + \frac{1}{2}B_{j-1}\Sigma_M\Sigma_M'B_{j-1}' \tag{34}$$

$$B_j = B_{j-1}[I_L + K_{1M}^{Q}], \quad j = 1/30, 1, ..., 12 \tag{35}$$

with $A_0 = \rho_{0M}$ and $B_0 = \rho_{1M}$. The transition equation is

$$\hat{M}_{t+1} = K_{0M}^{P} + K_{1M}^{P}\hat{M}_t + \Sigma_M\epsilon_{t+1}^{P} \tag{36}$$

The parameters to be estimated are ρ_{0M}, ρ_{1M}, K_{0M}^{Q}, K_{1M}^{Q}, K_{0M}^{P}, K_{1M}^{P}, and Σ.

3.4 The Lucia-Schwartz Model

We again use the extended Kalman filter to estimate the Lucia-Schwartz model. The model specification is in equation 15 through 18. The measurement equations are given by

$$f_t^j = \hat{f}_t^j + \eta_j, \quad j = 1/30, 1, 2, ..., 12 \tag{37}$$

where

$$\begin{aligned}\hat{f}_t^j &= e^{-\kappa}X_0 + \epsilon_0 + \mu_\epsilon^* - (1-e^{\kappa})\frac{\lambda_X}{\kappa} + \frac{1}{2}((1-e^{-2\kappa})\frac{\sigma_X^2}{2\kappa} + \sigma_\epsilon^2 \\ &+ 2(1-e^{-\kappa})\frac{\rho_{x\epsilon}\sigma_X\sigma_\epsilon}{\kappa}), \quad j = 1/30, 1, 2, ..., 12\end{aligned} \tag{38}$$

and $\eta_j \sim N(0, \sigma_j)$. The transition equation is given by

$$\begin{aligned}dX_t &= -\kappa X_t dt + \sigma_X dz_X^P \\ d\epsilon_t &= \mu_\epsilon dt + \sigma_\epsilon dz_\epsilon^P\end{aligned} \tag{39}$$

with $dz_X dz_\epsilon = \rho_{x\epsilon}dt$ and .

The parameters to be estimated are κ, μ_ϵ^*, μ_ϵ, λ_X, $\rho_{x\epsilon}$, σ_X, and σ_ϵ.

4 Data

As a proxy for s_t, we obtain the day-ahead peak contract for the PJM Western Hub market from the PJM website.[1] We model this price as a short-term futures contract maturing in one day. We obtain the PJM Western Hub real-time peak calendar-month 2.5 MW futures contracts from the CME. We include futures contracts with maturities of $1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11$ and 12 months. On each day, our sample therefore includes 13 futures prices. The sample period is from May 1, 2003 to May 30, 2014. All futures prices are deseasonalized by regressing on monthly dummies. Descriptive statistics for the raw futures prices and the deseasonalized prices are reported in Table 1. The time series for the one-month and 12-month raw and deseasonalized futures prices are plotted in Figure 1.

The economic data include demand and supply variables. We use the natural gas price as the supply variable. We obtain daily day-ahead natural gas middle prices for Columbia Gas and Texas Eastern Pipeline zone M-3 from Bloomberg. We use either the load or the temperature as the demand variable. We estimate the models using either the average daily load or the maximum load for the PJM Western Hub market. We obtain the load data from the PJM website.[2]

We obtain the temperature data from the National Climatic Data Center (NCDC). We first get the daily average, maximum and minimum temperature for Washington, D.C. and Pittsburgh from NCDC. Then we calculate the average temperature of the two cities and use it as a temperature proxy for the PJM Western Hub market. We also compute the cooling degree days (CDD) and heating degree days (HDD) according to the weather derivatives literature (see for example Alaton et al. (2002) and Jewson and Brix (2005)).

[1] See `https://dataminer.pjm.com/dataminerui/pages/public/lmp.jsf`.

[2] See `http://www.pjm.com/markets-and-operations/ops-analysis/`.

By combining the supply variable with the different demand proxies, we obtain seven different combinations of supply and demand variables. The supply variable is the natural gas price. The demand variable is either the maximum load, the average load, CDD, HDD, the maximum temperature, the minimum temperature, or the average temperature. All demand and supply variables are deseasonalized by regressing them on monthly dummies, and then standardized to have zero mean and unit variance. Descriptive statistics are reported in Table 2. In the empirical section below, we focus on results obtained using maximum load and CDD. The empirical results for other demand variables are qualitatively and quantitatively similar and are reported in the online appendix. The time series of the raw and deseasonalized natural gas price are plotted in Figure 2. The time series of raw and deseasonalized maximum load and CDD are plotted in Figure 3.

5 Empirical Results

We first conduct a principal component analysis of the electricity futures. Subsequently we empirically test the spanning conditions, we estimate the risk-neutral and physical dynamics, and we discuss model fit. Finally we analyze the model-implied spot premiums.

5.1 Principal Components of Electricity Futures

To provide insight into the factor structure of electricity futures, we conduct a principal component analysis. Figure 4 shows the loadings of the futures prices on the first three principal components (PC) and the fraction of total variance explained by each PC. The first two PCs explain 98% of total variation. This suggests that the electricity futures can be summarized by few factors and motives us to specify the economy under the Q measure using two factors.

5.2 Do Electricity Futures Span the Demand and Supply Variables?

We investigate the spanning properties of the demand and supply variables. We start with the natural gas price. Subsequently we analyze the load and temperature variables.

To examine if the natural gas price (PX) is spanned by the electricity futures, we project PX onto the first five principal components (PC^{1-5}) of the electricity futures.

$$PX_t = \alpha + \gamma_{pc} PC_t^{1-5} + unPX_t \tag{40}$$

The results are reported in column (1) of Table 3. The adjusted R^2 is 0.530, indicating that the PCs explain approximately half of the variation in the natural gas price. By definition the residuals from this regression, which we denote by unPX, represent the unspanned part of the natural gas price.

To determine if the residuals are spanned or unspanned, we need to verify if they contain relevant information about the futures. We use unPX to forecast changes in the first two PCs to investigate if unPX helps to predict the PCs.

$$\Delta PC_{t \to t+1}^{1-2} = \alpha + \beta_{pc} PC_t^{1-5} + \beta_{um} unPX_t + \epsilon_t \tag{41}$$

If the residuals are unspanned by the natural gas price, the loading on unPX in this forecasting regression should be statistically significant and the adjusted R^2 should increase when we add the residuals to the regression. Column (2) of Table 4 indicates that this is indeed the case. The loading on unPX is significant and positive for both PC^1 and PC^2. For PC^1, the loading is 0.026. Moreover, after including PX, the adjusted R^2 substantially increases compared to column (1), which only includes the PCs as regressors. The adjusted R^2 also substantially increases for PC^2 in Panel B. Higher adjusted R^2s mean that unPX contains useful information for predicting futures prices, suggesting that the natural gas price is indeed unspanned by the electricity futures.

Columns (2) and (3) in Table 3 report on the forecasting regression with the demand variables, the maximum load and the cooling degree days variable CDD.

$$Load_t^{max} = \alpha + \gamma_{pc} PC_t^{1-5} + unLoad_t^{max} \tag{42}$$

$$CDD_t = \alpha + \gamma_{pc} PC_t^{1-5} + unCDD_t \tag{43}$$

Table 3 shows that the principal components explain a larger fraction of the CDD variable compared to the Load variable. The adjusted R^2s for both demand variables in Table 3 are lower than for the natural gas variable. Results for other demand variables yield similar results and are reported in Table A.2 in the online appendix.

In Table 4, we also add the demand variables to the regressors to forecast changes in the first two PCs.

$$\Delta PC_{t\rightarrow t+1}^{1-2} = \alpha + \beta_{pc} PC_t^{1-5} + \beta_{um}[unPX_t, unLoad_t^{max}] + \epsilon_t \tag{44}$$

$$\Delta PC_{t\rightarrow t+1}^{1-2} = \alpha + \beta_{pc} PC_t^{1-5} + \beta_{um}[unPX_t, unCDD_t] + \epsilon_t \tag{45}$$

Column (3) in Table 4 shows that the coefficient of $unLoad^{max}$ is statistically significant in both Panel A and Panel B. However, the adjusted R^2 increases only slightly, from 0.126 to 0.131 for PC^1 and from 0.251 to 0.261 for PC^2. These results are due to the fact that the maximum load is volatile and difficult to predict. Changes in the maximum load reflect electricity demand shocks and contain important information about futures prices. The tables in the online appendix show that similar conclusions obtain when using average load.

The results in Table 4 show that CDD is unspanned by the electricity futures. The loadings are significant, and the adjusted R^2s substantially increase. The unspanned part of the cooling degree days variable thus contains useful information about electricity futures. The results for the other temperature variables are reported in Table A.2 and are overall very similar. Table A.3 in the online appendix reports results when lagged changes of the PCs are included as additional controls. The results confirm that the natural gas variable and the temperature variables contain important unspanned components.

5.3 Estimating the Risk-Neutral Dynamics

Now that we have identified the spanned and the unspanned part of the demand and supply variables, we use the spanned part to estimate the models' risk neutral dynamics. Table 5 reports these estimates. The upper left entry of the K_1^Q matrix is essentially one and highly statistically significant. This parameter captures the persistence of the model implied spot price under the risk neutral measure, and suggests that under this measure the spot price is a unit root process. Note that this estimate not only reflects the persistence of the spot price under the physical measure, but also its risk premium.

The loading of the spot price on PC^2 is negative and statistically significant. This loading measures the impact of the slope of the futures curve on the spot price. A larger PC^2 indicates a flatter slope, thus the negative sign means that a flatter slope predicts a lower spot. The interaction between the slope and the spot is ignored in the existing literature because different pricing factors are assumed to be independent.

The bottom left entry of K_1^Q is insignificant, indicating that the level of the futures curve does not affect its slope. Finally, the bottom right entry of K_1^Q indicates that the futures slope is mean reverting.

5.4 Model Fit

We report on the fit of the unspanned model and compare its performance to that of the benchmark models. The first benchmark is the spanned model with demand and supply variables, in which the demand and supply variables are fully spanned by the futures. The second benchmark is the Lucia and Schwartz model, which is a latent factor model without demand and supply variables. Note that by design the fit of the spanned model with latent factors is identical to that of the unspanned model. For each model, Table 6 reports the root mean squared error (RMSE) and the relative root mean squared error (RRMSE) for each futures contract as well as the overall RMSE and RRMSE. Figure 5 graphically illustrates the fit of the unspanned model for the one-month and twelve-month contracts.

The unspanned model has the smallest RMSE and RRMSE. The overall RMSE (RRMSE) is 0.055 (0.047) for the unspanned model compared to 0.106 (0.095) for the spanned model and 0.061 (0.052) for the Lucia and Schwartz model. The spanned model with demand and supply variables has the largest RMSE and RRMSE. This is not surprising because the assumption of full spanning constrains the pricing factors, which results in a worse fit. The Lucia and Schwartz model also provides a good fit, which is again not surprising because of its use of latent factors.

The main conclusion from Table 6 and Figure 5 is that the unspanned model provides a good fit. It improves on the fit of the latent model. This highlights the additional information in the demand and supply variables, but it is important not to over-emphasize this finding. The unspanned model contains more pricing factors and it nests the two other models. It is therefore to be expected that the in-sample fit is better. Our main conclusion is therefore rather that the model provides a satisfactory fit. While a latent factor model also provides a good fit, the unspanned model has an additional advantage because it allows us to study the role of the demand and supply variables in the pricing of the electricity futures.

5.5 Estimating the Physical Dynamics

We discuss the estimated P-dynamics, which capture the interaction between the electricity prices and the demand and supply variables under the physical measure.

5.5.1 The Electricity Price and the Natural Gas Price

Natural gas is the marginal fuel for electricity production, and consequently the natural gas price is closely tied to electricity prices. Table 7 confirms that the natural gas price affects the electricity price, consistent with existing findings in Cartea and Villaplana (2008) and Pirrong and Jermakyan (2008). Regardless of the demand variable, the loading of the spot price s_t on PX_t is approximately 0.089 and highly statistically significant. The positive sign reflects the economic relation between production cost and electricity price. This effect is economically meaningful. A one standard deviation increment in the natural gas price increases the electricity futures price by $e^{0.089} \approx 1.093$ dollars.

The natural gas price not only affects the spot price but also the slope of the electricity futures curve. Table 7 shows that the loading of PC^2_{t+1} on PX_t is positive and significant. A smaller PC^2 indicates a steeper futures slope, so a higher natural gas price predicts a flatter slope. This reflects the short-term impact of the natural gas price. The short-term impact of the natural gas price on the electricity price is larger than the long-term impact, resulting in a flatter slope.

The unspanned model also indicates that the electricity price affects the natural gas price. In Table 7 the loadings of PX_{t+1} on s_t are positive and significant. These positive signs indicate that higher electricity spot prices lead to higher natural gas prices. The high electricity price results from high demand for electricity, resulting in higher usage of natural gas and higher prices.

5.5.2 The Electricity Price and Demand

The load variable measures the demand for electricity, which should affect electricity prices. Panel A of Table 7 reports results for the maximum load variable. The impact of $Load_t^{max}$ on s_{t+1} and PC_{t+1}^2 is economically and statistically significant. Higher maximum load predicts higher spot prices and a flatter slope.

The electricity price also affects the maximum load. The impact of PC_t^2 on $Load_{t+1}^{max}$ is positive and significant. A higher PC^2 means a higher spot, which in turn predicts a higher electricity demand.

Temperature is another proxy for electricity demand. We investigated five different temperature variables. Panel B of Table 7 reports results using the cooling degree days (CDD) variable. Results for the other temperature variables are reported in Table A.4 in the online appendix.

The results are consistent with economic intuition. First, CDD positively affects the expected spot price. This reflects the fact that higher temperature in general leads to higher electricity demand. Second, because a higher PC_{t+1}^2 implies a flatter futures curve, the positive impact of CDD on PC_{t+1}^2 implies that the CDD negatively affects the slope of the futures curve. Third, CDD does not affect the natural gas price. This may be surprising but natural gas is a marginal fuel and mostly used on low temperature days.

5.6 The Spot Premium

We analyze the spot premium implied by the unspanned model and compare it with spot premiums in alternative models.

5.6.1 Estimating the Spot Premium

The spot premium measures the compensation required investors for investing in electricity futures. It is defined as the expected return in excess of the one-period basis:

$$\pi_{s,t} = E_t^P[r_{s,t+1}] - y_t^{(1)} = E_t^P[ln(S_{t+1}) - ln(S_t)] - y_t^{(1)} = E_t^P[s_{t+1} - s_t] - y_t^{(1)} \tag{46}$$

A more detailed discussion is provided in the appendix.

Figure 6 demonstrates plots the time series of the spot premium for the unspanned model and several spanned models. Table 8 reports the average daily spot premium for each year in the sample. The spot premium is time-varying in all models except the Lucia and Schwartz model. See the appendix for a discussion of the spot premium in the Lucia and Schwartz model.

Comparing the unspanned model with the alternatives, there are a number of noteworthy differences. First, there is a very large positive spike in the spot premium of the unspanned model towards the end of the sample, and a smaller one at the start of the sample. The other models do not exhibit these spikes in the spot premium. In the spanned model with demand and supply variables, there are two downward spikes in those time periods. The spanned model with demand and supply variables also shows a downward spike towards the end of the sample. Second, in the unspanned model the spot premium turns negative starting in 2010. This is also the case in the spanned model with latent factors but not in the spanned model with demand and supply variables. Third, the term premium is substantially less volatile in the unspanned model compared to the spanned model with demand and supply variables. Finally, while the spot premium in the unspanned model and the spanned model with latent factors co-move to a large extent, Figure 6 clearly illustrates that the properties of the two time series are quite different.

We conclude that ignoring the risk premium associated with the demand and supply variables leads to different estimates of the overall risk premium. Consistent with Cartea and Villaplana (2008) and Pirrong and Jermakyan (2008), our results therefore suggest that the demand and supply variables contain important information about electricity futures.

However, in contrast with Cartea and Villaplana (2008) and Pirrong and Jermakyan (2008), we conclude that the risk premiums are not entirely due to the demand and supply variables. For example, by imposing the restriction of full spanning, the spanned model with demand and supply variables implies positive spot premiums after 2010, while the implied spot premium in the unspanned model is negative. Figure 6 demonstrates that the spanned model with demand and supply variables does not capture the peak in the spot premium in the winter of 2013, when cold weather dramatically increased futures prices.

In summary, our results show that it is important to consider both latent factors and demand and supply variables when modeling risk premiums in electricity futures.

5.6.2 Decomposing the Spot Premium

To further investigate the dynamics of the spot premium, we decompose the spot premium into four components: a constant, a component associated with the natural gas price, a component associated with the load or the temperature, and a component associated with electricity itself. The appendix provides details on this decomposition.

Table 9 and Figure 7 present the results for the model with PX and $Load^{max}$. Table 10 and Figure 8 present the results for the model with PX and CDD. Several conclusions obtain. First, the component associated with electricity is time-varying. It can be both positive and negative, indicating that investors sometimes require compensation to bear this risk while at times paying to hedge this risk. Second, the negative spot premium from 2010 onward is largely due to the risk associated with the natural gas price. The third column in

Table 9 and Table 10 shows that the spot premium associated with the natural gas price is large and becomes negative after 2010. During this period, the natural gas price is highly volatile, and therefore investors want to hold the futures contract to hedge the fuel price risks. Consequently the futures price is bid up, which generates a negative spot return. Third, load and temperature have relatively minor effects on the spot premium.

5.6.3 Spot Premiums By Season

Table 11 reports the daily average of spot premiums of unspanned model, spanned model, macro model, and Lucia and Schwartz model in each season. The spot premium is defined as follows.

$$\text{Spot Premium}_t = E_t^P[\text{Spot Price}_{t+1}] - E_t^Q[\text{Spot Price}_{t+1}] \tag{47}$$

calculated as the difference between the expected spot price under P and the expected spot price under Q. We calculate the daily spot premium associated with the day-ahead price and report the simple average in each season.

Table 11 shows summary statistics of the spot premium in unspanned model of each season. It first shows that spot premiums are negative in every season (except for the Lucia-Schwartz model). Because spot premium equals to P price minus Q price, a negative premium suggests electricity consumers would like to pay to hedge the risk associated with the time-varying electricity prices. Furthermore, the standard deviation, minimum, and maximum all shows spot premiums are more volatile in Winter/Summer than in Spring/Fall. This is consistent with the fact that demands are more volatile in peak seasons. At last, compared with different models, models with macro variables (unspanned model and macro model) can generate a wider range for spot premiums than models without macro variables (latent model and Lucia and Schwartz model) .

5.6.4 Spot Premiums and Properties of Spot Prices

The model proposed by Bessembinder and Lemmon (2002) implies that the electricity futures price is related with statistical properties of spot prices. Specifically, they predict the forward premium is negative correlated with the variance of spot prices and positively correlated with the skewness of spot prices. To explore this relationship in our model, we regress the spot premium against the variance and skewness of spot price prices. The regression specification is as follows.

$$\text{Spot Premium}_t = \alpha + \beta_{\text{Variance}} \times \text{Variance}_t + \beta_{\text{Skewness}} \times \text{Skewness}_t + \epsilon_t \tag{48}$$

where Spot Premium_t is the average daily spot premium of unspanned model in month t. The spot premiums are calculated with different economic variable pairs. We use variance of natural gas price as a proxy for the daily power price. Variance_t is the variance of daily natural prices in month t, and Skewness_t is the skewness of natural spot prices in month t. Because the spot premium equals to the negative of the forward premium, we expect a positive β_{Variance} and a negative β_{Skewness}.

Table 12 reports the estimated coefficients of equation 48. We find our estimated coefficients are consistent with the theory. We find no matter how we computed spot premiums, the spot premium is positively related with the variance and negatively related with the skewness. Thus, our estimated spot premiums are consistent with Bessembinder and Lemmon (2002).

5.6.5 Spot Premiums during the Polar Vertex

Figure 11 plots the spot premiums of 4 different models during the 2014 polar vertex period. It shows how macro variables affect the spot premium. Similar to the results of Table 11, we find models with macro variables could reflect the dramatic change of the risk premiums, but models without macro variables could not.

5.7 Forward Premiums

The forward premium is defined as the difference between the P forward rate and the Q forward rate. It represents the expected forward return which can be locked down today. The detailed definition is given as follows.

$$\text{Foward Premium}_t^{i\to i+j} = \text{Implied Foward Rate in P}_t^{i\to i+j} - \text{Implied Foward Rate in Q}_t^{i\to i+j} \tag{49}$$

$$\text{Implied Foward Rate in P}_t^{i\to i+j} = (E^P[f_t^{i+j}]/E^P[f_t^{i+j}])^{\frac{1}{j}} - 1$$

$$\text{Implied Foward Rate in Q}_t^{i\to i+j} = (E^Q[f_t^{i+j}]/E^Q[f_t^{i+j}])^{\frac{1}{j}} - 1$$

where f_t^i is the expected price of futures maturing in i at time t.

Figure 9 plots three forward premiums of unspanned model. The blue line denotes the forward premium of 3 month to 1 month. The red line is 6 month to 1 month and the yellow line is the 9 month to 1 month. The figure first shows forward premiums can be either positive or negative, suggesting power suppliers both pay or demand premium for bearing different power risks. Moreover, the figure shows the short term forward premium is of larger magnitude and more volatile than the long term forward premium, suggesting investors care about the transitory shock in the short term. At last, the figure shows different forward premiums share similar patterns. All of them peak at the end of the sample period when the polar vertex occurs.

We also compare the forward premium of unspanned model with the spanned or latent model. Figure 10 plots the 3 to 1 forward premium in those two models as well as their differences. While the general patterns are quite similar, residuals show that unspanned model can capture the premium change due to macro variables (e.g. at the end of the sample period) while spanned model cannot capture that fact.

6 Concluding Remarks

This papers uses an unspanned model to conduct a extensive study about the electricity futures pricing. We find that the latent variables could achieve a decent fit about the futures price but the economic variables contain important information about the risk premium of electricity futures. We find both the spot premium and the forward premium are highly time-varying and are largely affected by the supply side variable. Our results show that economic variables are important not only because they affect the futures price but also because they contain important information about the risk premium.

References

Peter Alaton, Boualem Djehiche, and David Stillberger. On modelling and pricing weather derivatives. *Applied Mathematical Finance*, 9(1):1–20, 2002.

Fred Espen Benth, Jan Kallsen, and Thilo Meyer-Brandis. A non-gaussian ornstein–uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, 14(2):153–169, 2007.

Fred Espen Benth, Álvaro Cartea, and Rüdiger Kiesel. Pricing forward contracts in power markets by the certainty equivalence principle: explaining the sign of the market risk premium. *Journal of Banking & Finance*, 32(10):2006–2021, 2008.

Hendrik Bessembinder and Michael L Lemmon. Equilibrium pricing and optimal hedging in electricity forward markets. *The Journal of Finance*, 57(3):1347–1382, 2002.

Michael Bierbrauer, Christian Menn, Svetlozar T Rachev, and Stefan Trück. Spot and derivative pricing in the eex power market. *Journal of Banking & Finance*, 31(11): 3462–3485, 2007.

Derek W Bunn and Dipeng Chen. The forward premium in electricity futures. *Journal of Empirical Finance*, 23:173–186, 2013.

Alvaro Cartea and Marcelo G Figueroa. Pricing in electricity markets: a mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4):313–335, 2005.

Álvaro Cartea and Pablo Villaplana. Spot price modeling and the valuation of electricity forward contracts: the role of demand and capacity. *Journal of Banking & Finance*, 32 (12):2502–2519, 2008.

Shi-Jie Deng and Shmuel S Oren. Electricity derivatives and risk management. *Energy*, 31 (6):940–953, 2006.

Lingxiu Dong and Hong Liu. Equilibrium forward contracts on nonstorable commodities in the presence of market power. *Operations Research*, 55(1):128–145, 2007.

Stratford Douglas and Julia Popova. Storage and the electricity forward premium. *Energy Economics*, 30(4):1712–1727, 2008.

Helyette Geman. *Commodities and commodity derivatives: modeling and pricing for agriculturals, metals and energy*. John Wiley & Sons, 2009.

Hélyette Geman and Andrea Roncoroni. Understanding the fine structure of electricity prices. *The Journal of Business*, 79(3):1225–1261, 2006.

Davidson Heath. Unspanned macroeconomic risks in oil futures. *Available at SSRN 2506146*, 2016.

Stephen Jewson and Anders Brix. *Weather derivative valuation: the meteorological, statistical, financial and mathematical foundations*. Cambridge University Press, 2005.

Scott Joslin, Marcel Priebsch, and Kenneth J Singleton. Risk premiums in dynamic term structure models with unspanned macro risks. *The Journal of Finance*, 69(3):1197–1233, 2014.

Francis A Longstaff and Ashley W Wang. Electricity forward prices: a high-frequency empirical analysis. *The Journal of Finance*, 59(4):1877–1900, 2004.

Julio J Lucia and Eduardo S Schwartz. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 5(1):5–50, 2002.

Craig Pirrong. *Commodity price dynamics: A structural approach*. Cambridge University Press, 2011.

Craig Pirrong and Martin Jermakyan. The price of power: The valuation of power and weather derivatives. *Journal of Banking & Finance*, 32(12):2520–2529, 2008.

Marta Szymanowska, Frans Roon, Theo Nijman, and Rob Goorbergh. An anatomy of commodity futures risk premia. *The Journal of Finance*, 69(1):453–482, 2014.

Table 1
Descriptive Statistics for Electricity Futures Prices

We report descriptive statistics for PJM Western Hub day-ahead prices and PJM Western Hub real-time peak calendar-month 2.5 MW futures prices. We report the number of observations, mean, standard deviation, minimum, maximum, skewness, kurtosis and autocorrelation for each series. Panel A reports the statistics for the raw data. Panel B reports the statistics for the deseasonalized data. $f_{1/30}$ denotes the day-ahead PJM Western Hub real-time peak price. f_1 to f_{12} denote the futures price with maturity 1 month to 12 months, respectively. The sample period is from May 1, 2003 to May 30, 2014.

Panel A: Electricity Futures

	N. Obs.	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis	Autocorrelation
$F_{1/30}$	2722	57.914	29.641	20.684	683.162	6.957	109.400	0.639
F_1	2722	57.958	18.729	31.350	164.750	1.760	7.385	0.776
F_2	2722	58.771	19.387	33.060	154.330	1.740	6.994	0.781
F_3	2722	59.325	19.249	35.170	140.000	1.549	5.568	0.782
F_4	2722	59.613	18.990	34.170	139.230	1.433	4.932	0.779
F_5	2722	59.683	18.216	35.040	128.290	1.197	3.942	0.771
F_6	2722	59.823	17.895	34.110	128.420	1.076	3.465	0.771
F_7	2722	60.186	18.109	34.320	128.420	1.105	3.810	0.774
F_8	2722	60.747	18.490	35.360	127.080	1.054	3.621	0.774
F_9	2722	61.046	18.558	35.320	116.250	0.967	3.159	0.774
F_{10}	2722	61.040	18.038	35.000	114.330	0.900	2.929	0.770
F_{11}	2722	60.975	17.648	34.300	119.500	0.894	2.960	0.766
F_{12}	2722	60.950	17.871	38.260	138.490	1.042	3.753	0.768

Panel B: Deseasonalized Electricity Futures

	N. Obs.	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis	Autocorrelation
$F_{1/30}$	2722	1.072	0.515	0.104	11.613	6.463	98.056	0.637
F_1	2722	1.036	0.301	0.607	2.815	1.514	5.321	0.778
F_2	2722	1.036	0.302	0.614	2.233	1.509	5.113	0.784
F_3	2722	1.037	0.303	0.638	2.249	1.437	4.680	0.780
F_4	2722	1.038	0.305	0.643	2.110	1.374	4.425	0.777
F_5	2722	1.037	0.298	0.652	2.126	1.208	3.768	0.773
F_6	2722	1.036	0.292	0.651	2.097	1.092	3.345	0.772
F_7	2722	1.035	0.287	0.666	2.114	1.013	3.170	0.776
F_8	2722	1.035	0.285	0.672	2.013	0.908	2.813	0.773
F_9	2722	1.035	0.283	0.657	1.903	0.855	2.639	0.777
F_{10}	2722	1.034	0.279	0.658	1.871	0.800	2.471	0.777
F_{11}	2722	1.034	0.276	0.656	2.032	0.792	2.498	0.776
F_{12}	2722	1.033	0.276	0.656	1.930	0.849	2.704	0.775

Table 2
Descriptive Statistics for Demand and Supply Variables

We report descriptive statistics for the demand and supply variables. We report the number of observations, mean, standard deviation, minimum, maximum, skewness, kurtosis and autocorrelation for each series. Panel A reports the statistics for the raw data. Panel B reports the statistics for the deseasonalized data. PX denotes the day-ahead natural gas price, $Load^{avg}$ denotes the daily average load, $Load^{max}$ denotes the daily maximum load, CDD denotes cooling degree days, HDD denotes heating degree days, T^{max} denotes the daily maximum temperature, T^{min} denotes the daily minimum temperature, and T^{avg} denotes the daily average temperature. The sample period is from May 1, 2003 to May 30, 2014.

Panel A: Demand and Supply Variables

	N. Obs.	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis	Autocorrelation
PX	2722	6.355	3.903	1.945	79.846	7.772	121.595	0.618
$Load^{avg}$	2722	33420.580	4443.382	19951.250	51998.380	0.821	3.229	0.712
$Load^{max}$	2722	39009.350	6202.780	24824.000	61646.000	0.970	3.455	0.701
CDD	2722	77.768	71.555	0.000	242.000	0.303	1.608	0.733
HDD	2722	18.831	37.106	0.000	222.500	2.175	7.325	0.675
T^{max}	2722	177.725	103.988	−116.000	381.000	−0.376	2.084	0.718
T^{min}	2722	70.150	93.627	−199.000	244.500	−0.243	2.090	0.721
T^{avg}	2722	123.937	97.092	−157.500	307.000	−0.318	2.076	0.735

Panel B: Deseasonalized Demand and Supply Variables

	N. Obs.	Mean	Std.Dev	Minimum	Maximum	Skewness	Kurtosis	Autocorrelation
PX	2722	0.000	3.779	−6.800	70.509	7.355	114.646	0.610
$Load^{avg}$	2722	0.000	2985.228	−9498.074	12371.730	0.608	4.236	0.625
$Load^{max}$	2722	0.000	4039.404	−12508.650	18528.080	0.619	4.764	0.595
CDD	2722	0.000	31.394	−102.100	121.781	0.465	3.901	0.541
HDD	2722	0.000	26.039	−75.025	147.475	0.898	7.825	0.566
T^{max}	2722	0.000	51.739	−156.830	171.750	0.134	3.152	0.524
T^{min}	2722	0.000	44.329	−168.583	172.547	−0.006	3.270	0.533
T^{avg}	2722	0.000	44.605	−152.101	169.149	0.131	3.396	0.568

Table 3
Projecting Demand and Supply Variables on Electricity Futures PCs

We report the results from projecting the demand and supply variables M_t on the first five principal components (PC_t^{1-5}) of electricity futures, i.e.

$$M_t = \alpha + \gamma_{pc} PC_t^{1-5} + UM_t$$

Columns (1) to (3) report the regression results for the natural gas price (PX), the maximum load ($Load^{max}$), and cooling degree days (CDD) respectively. For each regression, the adjusted R^2 and the number of observations are reported. Standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

	(1)	(2)	(3)
	PX_t	$Load_t^{max}$	CDD_t
PC_t^1	0.658***	0.163***	0.004
	(0.014)	(0.010)	(0.009)
PC_t^2	1.314***	1.561***	0.345***
	(0.049)	(0.034)	(0.031)
PC_t^3	1.088***	−1.257***	−0.448***
	(0.100)	(0.069)	(0.064)
PC_t^4	−1.671***	1.454***	0.502***
	(0.202)	(0.139)	(0.129)
PC_t^5	0.535**	0.860***	0.218
	(0.205)	(0.139)	(0.131)
Constant	0.000	0.000	0.000
	(0.012)	(0.008)	(0.011)
Adj. R^2	0.530	0.061	0.279
N. Obs.	2722	2722	2722

Table 4
Forecasting Changes in Principal Components

We forecast changes in the first two principal components (ΔPC^{1-2}) using the unspanned component of demand and supply variables (UMs), i.e.

$$\Delta PC^{1-2}_{t\to t+1} = \alpha + \beta_{pc} PC^{1-5}_t + \beta_{um} UM_t + \epsilon_t$$

Lags of the first five PCs, i.e. PC^{1-5}_t, are used as control variables. Panel A reports results for the first PC. Panel B reports results for the second PC. The first column reports results using the PCs only. The second column adds the unspanned natural gas price (unPX). The third column adds the unspanned natural gas price (unPX) and the unspanned maximum load ($unLoad^{max}$) as regressors. The fourth column adds the unspanned natural gas price and the unspanned CDD ($unCDD$). Standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

Panel A: Forecasting Changes in the First Principal Component

	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$
$unPX_t$		0.026***	0.027***	0.026***
		(0.002)	(0.002)	(0.002)
$unLoad^{max}_t$			0.010***	
			(0.003)	
$unCDD_t$				0.017***
				(0.003)
PC^1 to PC^5	Yes	Yes	Yes	Yes
Adj. R^2	0.044	0.126	0.131	0.138
N. Obs.	2117	2117	2117	2117

Panel B: Forecasting Changes in the Second Principal Component

	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$
$unPX_t$		0.076***	0.080***	0.080***
		(0.004)	(0.004)	(0.004)
$unLoad^{max}_t$			0.033***	
			(0.006)	
$unCDD_t$				0.042***
				(0.007)
PC^1 to PC^5	Yes	Yes	Yes	Yes
Adj. R^2	0.123	0.251	0.261	0.364
N. Obs.	2117	2117	2117	2117

Table 5
Risk-Neutral Dynamics

We report maximum likelihood estimates of the unspanned model. s_t denotes the electricity spot price. PC^2 denotes the second principal component extracted from the electricity futures. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

	K_0^Q	K_1^Q	
		s_t	PC_t^2
s_{t+1}	-1.773×10^{-5}	1.000***	−0.047***
	(1.928×10^{-4})	(1.995×10^{-4})	(0.001)
PC_{t+1}^2	-1.635×10^{-5}	2.039×10^{-4}	0.957***
	(1.045×10^{-4})	(7.898×10^{-4})	(0.001)

Table 6
Model Fit

We report the root mean squared error (RMSE) and relative root mean squared error (RRMSE) for the unspanned model, the spanned model with demand and supply variables, and the Lucia and Schwartz (2002) model. RMSE and RRMSE are defined as

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{F}_i - F_i)^2} \qquad RRMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\frac{(\hat{F}_i - F_i)^2}{F_i^2}}$$

where $\hat{F}_i$ is the fitted futures price and F_i is the realized deseasonalized futures price. $F_{1/30}$ denotes the futures maturing in 1 day. F_1 to F_{12} denote the futures maturing in 1 month and in 12 months respectively. Overall RMSE and RRMSE are calculated as simple averages using all futures contracts. The sample period is from May 1, 2003 to May 30, 2014.

	Unspanned Model with Economic Variables		Spanned Model with Economic Variables		Lucia-Schwartz Model	
	RMSE	RRMSE	RMSE	RRMSE	RMSE	RRMSE
$F_{1/30}$	0.049	0.039	0.233	0.177	0.198	0.152
F_1	0.080	0.074	0.074	0.074	0.069	0.062
F_2	0.073	0.063	0.074	0.067	0.058	0.051
F_3	0.068	0.056	0.078	0.070	0.062	0.052
F_4	0.061	0.046	0.080	0.074	0.054	0.046
F_5	0.049	0.038	0.090	0.081	0.057	0.046
F_6	0.041	0.034	0.092	0.081	0.046	0.039
F_7	0.039	0.034	0.100	0.088	0.044	0.038
F_8	0.042	0.038	0.105	0.096	0.036	0.034
F_9	0.049	0.044	0.113	0.105	0.038	0.035
F_{10}	0.054	0.047	0.114	0.105	0.036	0.032
F_{11}	0.058	0.050	0.112	0.105	0.047	0.041
F_{12}	0.058	0.051	0.114	0.106	0.051	0.046
Overall	0.055	0.047	0.106	0.095	0.061	0.052

Table 7
Physical Dynamics

We report estimates of the physical dynamics for different demand and supply variables. Panel A reports results for the natural gas price and the maximum load. Panel B reports results for the natural gas price and CDD. Bootstrapped standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

Panel A: Physical Dynamics using the Natural Gas Price and the Maximum Load

	K_0^P	K_1^P			
		s_t	PC_t^2	PX_t	$Load_t^{max}$
s_{t+1}	−0.014***	0.775***	−0.102***	0.089***	0.011**
	(0.003)	(0.016)	(0.018)	(0.004)	(0.006)
PC_{t+1}^2	−0.014***	−0.227***	0.899***	0.083***	0.019**
	(0.003)	(0.016)	(0.017)	(0.004)	(0.006)
PX_{t+1}	−0.021*	1.047***	−0.919***	0.559***	−0.046*
	(0.009)	(0.049)	(0.053)	(0.013)	(0.019)
$Load_{t+1}^{max}$	−0.028	0.025	0.131*	0.041**	0.718***
	(0.009)	(0.048)	(0.052)	(0.012)	(0.018)

Panel B: Physical Dynamics using the Natural Gas Price and CDD

	K_0^P	K_1^P			
		s_t	PC_t^2	PX_t	CDD_t
s_{t+1}	−0.014***	0.766***	−0.095***	0.089***	0.040***
	(0.003)	(0.015)	(0.016)	(0.004)	(0.007)
PC_{t+1}^2	−0.013***	−0.216***	0.905***	0.083***	0.037***
	(0.002)	(0.014)	(0.016)	(0.003)	(0.006)
PX_{t+1}	−0.020*	1.013***	−0.947***	0.563***	−0.020
	(0.008)	(0.045)	(0.050)	(0.012)	(0.021)
CDD_{t+1}	0.006	0.040	−0.022	−0.009	0.688***
	(0.006)	(0.034)	(0.037)	(0.009)	(0.015)

Table 8
Spot Premiums

We report the sample mean and the standard deviation of daily spot premiums for different models. The second column reports the spot premium for the unspanned model using the natural gas price and the CDD. The third column reports the spot premium of the spanned model. The fourth column reports the spot premium of the spanned model which uses the natural gas price and the CDD. The last column reports the spot premium of the Lucia and Schwartz model. The spot premium is the difference between the expected spot return and the one-period basis. See the appendix for more details. The sample is from May 1, 2003 to May 30, 2014.

	Unspanned Model with Economic Variables	Unspanned Model with Latent Factors	Spanned Model with Economic Variables	Lucia-Schwartz Model
2003	0.022	−0.012	0.036	0.007
2004	0.005	−0.016	−0.011	0.007
2005	−0.012	−0.029	−0.113	0.007
2006	0.013	0.013	−0.077	0.007
2007	−0.023	−0.012	−0.093	0.007
2008	−0.032	−0.023	−0.134	0.007
2009	0.018	0.011	0.026	0.007
2010	−0.041	−0.027	0.018	0.007
2011	−0.037	−0.017	0.024	0.007
2012	−0.014	−0.004	0.079	0.007
2013	−0.030	−0.017	0.060	0.007
2014	−0.052	−0.065	−0.041	0.007
Mean	−0.015	−0.015	−0.018	0.007
Standard Deviation	0.070	0.036	0.093	0.000

Table 9
Decomposing the Spot Premium. Unspanned Model with Natural Gas Price and Maximum Load

We report the decomposition of the average daily spot premium using the unspanned model with the natural gas price and the maximum load. The decomposition is given by

$$\pi_{s,t} = Constant + \pi_{Electricity,t} + \pi_{PX,t} + \pi_{Load_{max},t}$$

where $\pi_{Electricity,t}$ is the spot premium associated with the electricity, $\pi_{PX,t}$ is the spot premium associated with the natural gas price and $\pi_{Load_{max},t}$ is the spot premium associated with the maximum load. See the appendix for details. The sample period is from May 1, 2003 to May 30, 2014.

	$\pi_{s,t}^{PX,Load_{max}}$	$\pi_{Electricity,t}^{PX,Load_{max}}$	$\pi_{PX,t}^{PX,Load_{max}}$	$\pi_{Load_{max},t}^{PX,Load_{max}}$
2003	0.022	0.051	−0.009	−0.006
2004	0.005	0.011	0.009	0.000
2005	−0.012	−0.079	0.078	0.004
2006	0.013	0.006	0.022	0.000
2007	−0.023	−0.047	0.034	0.006
2008	−0.032	−0.097	0.080	0.001
2009	0.018	0.073	−0.038	−0.003
2010	−0.041	−0.001	−0.028	0.003
2011	−0.037	0.017	−0.039	0.000
2012	−0.014	0.082	−0.077	−0.004
2013	−0.030	0.043	−0.054	−0.003
2014	−0.052	−0.095	0.057	0.002
2003 - 2014	−0.015	0.000	0.000	0.000

Table 10
Decomposing the Spot Premium. Unspanned Model with Natural Gas Price and CDD

We report the decomposition of the average daily spot premium using the unspanned model with the natural gas price and CDD. The decomposition is given by

$$\pi_{s,t} = Constant + \pi_{Electricity,t} + \pi_{PX,t} + \pi_{CDD,t}$$

where $\pi_{Electricity,t}$ is the spot premium associated with the electricity, $\pi_{PX,t}$ is the spot premium associated with the natural gas price and $\pi_{CDD,t}$ is the spot premium associated with the CDD. See the appendix for details. The sample period is from May 1, 2003 to May 30, 2014.

	$\pi^{PX,CDD}_{s,t}$	$\pi^{PX,CDD}_{Electricity,t}$	$\pi^{PX,CDD}_{PX,t}$	$\pi^{PX,CDD}_{CDD,t}$
2003	0.020	0.049	−0.009	−0.005
2004	0.004	0.010	0.009	−0.001
2005	−0.011	−0.070	0.078	0.000
2006	0.012	−0.002	0.022	−0.001
2007	−0.022	−0.046	0.034	0.004
2008	−0.031	−0.093	0.080	−0.003
2009	0.014	0.069	−0.038	−0.002
2010	−0.040	0.000	−0.028	0.002
2011	−0.035	0.017	−0.039	0.002
2012	−0.011	0.078	−0.077	0.003
2013	−0.028	0.041	−0.055	0.000
2014	−0.050	−0.089	0.057	−0.004
2003 - 2014	−0.015	0.000	0.000	0.000

Table 11
Average Spot Premiums: By Seasons

This table reports the daily average of spot premiums of different models by season. The spot premium is defined as follows.

$$\text{Spot Premium}_t = E_t^P[\text{Spot Price}_{t+1}] - E_t^Q[\text{Spot Price}_{t+1}]$$

Winter is defined as December, January, and February. Spring is defined as March, April, and May. Summer is defined as June, July, and August. Fall is defined as September, October, and November. Unspanned model denotes the model with both latent and macro variables Latent model is the model with only latent variables. Macro model is the one with only macro variables. Lucia-Schwartz denotes the Lucia and Schwartz (2002) model. The macro variables here are natural gas prices and CDDs. The reported numbers are daily returns in raw numbers. The sample period is from May 1, 2003 to May 30, 2014.

	Winter				
	N. Obs	Mean	Std. Dev.	Minimum	Maximum
Unspanned Model	656	-0.0148	0.1049	-0.5065	1.1416
Latent Model	656	-0.0145	0.0478	-0.3668	0.0992
Macro Model	656	-0.0183	0.1093	-1.0160	0.1520
Lucia-Schwartz	656	0.0074	0.0000	0.0074	0.0074
	Spring				
	N. Obs	Mean	Std. Dev.	Minimum	Maximum
Unspanned Model	715	-0.0148	0.0502	-0.4487	0.1292
Latent Model	715	-0.0145	0.0278	-0.2536	0.0514
Macro Model	715	-0.0182	0.0818	-0.2639	0.1618
Lucia-Schwartz	715	0.0074	0.0000	0.0074	0.0074
	Summer				
	N. Obs	Mean	Std. Dev.	Minimum	Maximum
Unspanned Model	687	-0.0148	0.0666	-0.3576	0.1314
Latent Model	687	-0.0145	0.0389	-0.1861	0.0705
Macro Model	687	-0.0183	0.0861	-0.2967	0.1592
Lucia-Schwartz	687	0.0074	0.0000	0.0074	0.0074
	Fall				
	N. Obs	Mean	Std. Dev.	Minimum	Maximum
Unspanned Model	664	-0.0148	0.0442	-0.2409	0.1438
Latent Model	664	-0.0145	0.0267	-0.1406	0.0751
Macro Model	664	-0.0183	0.0939	-0.3401	0.1362
Lucia-Schwartz	664	0.0074	0.0000	0.0074	0.0074

Table 12
Relationship between Spot Premiums and Variance and Skewness of Day-Ahead Prices

This table reports the estimated coefficients of spot premiums on variance and skewness of day-ahead prices. The regression specification is given as follows.

$$\text{Spot Premium}_t = \alpha + \beta_{\text{Variance}} \times \text{Variance}_t + \beta_{\text{Skewness}} \times \text{Skewness}_t + \epsilon_t$$

where Spot Premium_t is the average daily spot premium of unspanned model in month t, Variance_t is the variance of daily natural gas prices in month t, and Skewness_t is the skewness of daily natural gas prices in month t. I use them to approximate the variance and skewness of daily prices. The sample period is from May, 2003 to May, 2014. *,**,*** denotes the significance level at 10%, 5%, and 1% respectively.

	β_{Variance}	β_{Skewness}	α	N.Obs	Adjusted R^2
$SP_{\text{PX, CDD}}$	0.0025***	-0.0078***	-0.0123***	133	0.096
	(0.0008)	(0.0029)	(0.0030)		
$SP_{\text{PX, HDD}}$	0.0027***	-0.0076***	-0.0130***	133	0.113
	(0.0007)	(0.0028)	(0.0029)		
$SP_{\text{PX, Max Load}}$	0.0024***	-0.0081***	-0.0123***	133	0.091
	(0.0008)	(0.0029)	(0.0031)		
$SP_{\text{PX, Avg Load}}$	0.0026***	-0.0078***	-0.0125***	133	0.107
	(0.0007)	(0.0028)	(0.0029)		
$SP_{\text{PX, Max T}}$	0.0026***	-0.0074***	-0.0131***	133	0.105
	(0.0007)	(0.0028)	(0.0029)		
$SP_{\text{PX, Min T}}$	0.0025***	-0.0079***	-0.0126***	133	0.106
	(0.0007)	(0.0028)	(0.0029)		
$SP_{\text{PX, Avg T}}$	0.0026***	-0.0076***	-0.0129***	133	0.105
	(0.0007)	(0.0028)	(0.0029)		

Table 13
Model Estimates from 2003 to 2013

This table reports the estimated P and Q dynamics of a subsample from 2003 to 2013. Bootstrapped standard errors are reported in parentheses. The sample period is from May, 2003 to December, 2013.

Panel A: P Dynamics

	K_0^P	K_1^P			
		s_t	PC_t^2	PX_t	T_t
s_{t+1}	−0.146	0.752	−0.040	0.097	0.038
	(0.003)	(0.022)	(0.021)	(0.008)	(0.007)
PC_{t+1}^2	−0.014	−0.231	0.959	0.091	0.034
	(0.003)	(0.020)	(0.020)	(0.007)	(0.007)
PX_{t+1}	−0.012	0.590	−0.562	0.740	−0.009
	(0.006)	(0.045)	(0.044)	(0.016)	(0.014)
CDD_{t+1}	0.007	0.092	−0.047	−0.032	0.691
	(0.007)	(0.051)	(0.049)	(0.018)	(0.016)

Panel B: Q Dynamics

	K_0^Q	K_1^Q	
		s_t	PC_t^2
s_{t+1}	-4.590×10^{-5}	1.000	-0.047
	(8.259×10^{-5})	(3.454×10^{-4})	(0.002)
PC_{t+1}^2	-4.592×10^{-5}	2.680×10^{-4}	0.957
	(7.842×10^{-5})	(3.493×10^{-4})	(0.002)

Futures Prices

We plot the time series of the futures prices. The top two panels plot the price of the one-month (F1) contract and the deseasonalized F1 price. The bottom two panels plot the price of the twelve-month (F12) contract and the deseasonalized F12 price. The deseasonalized price is the residual from regressing the raw series on monthly dummies. The sample period is from May 1, 2003 to May 30, 2014.

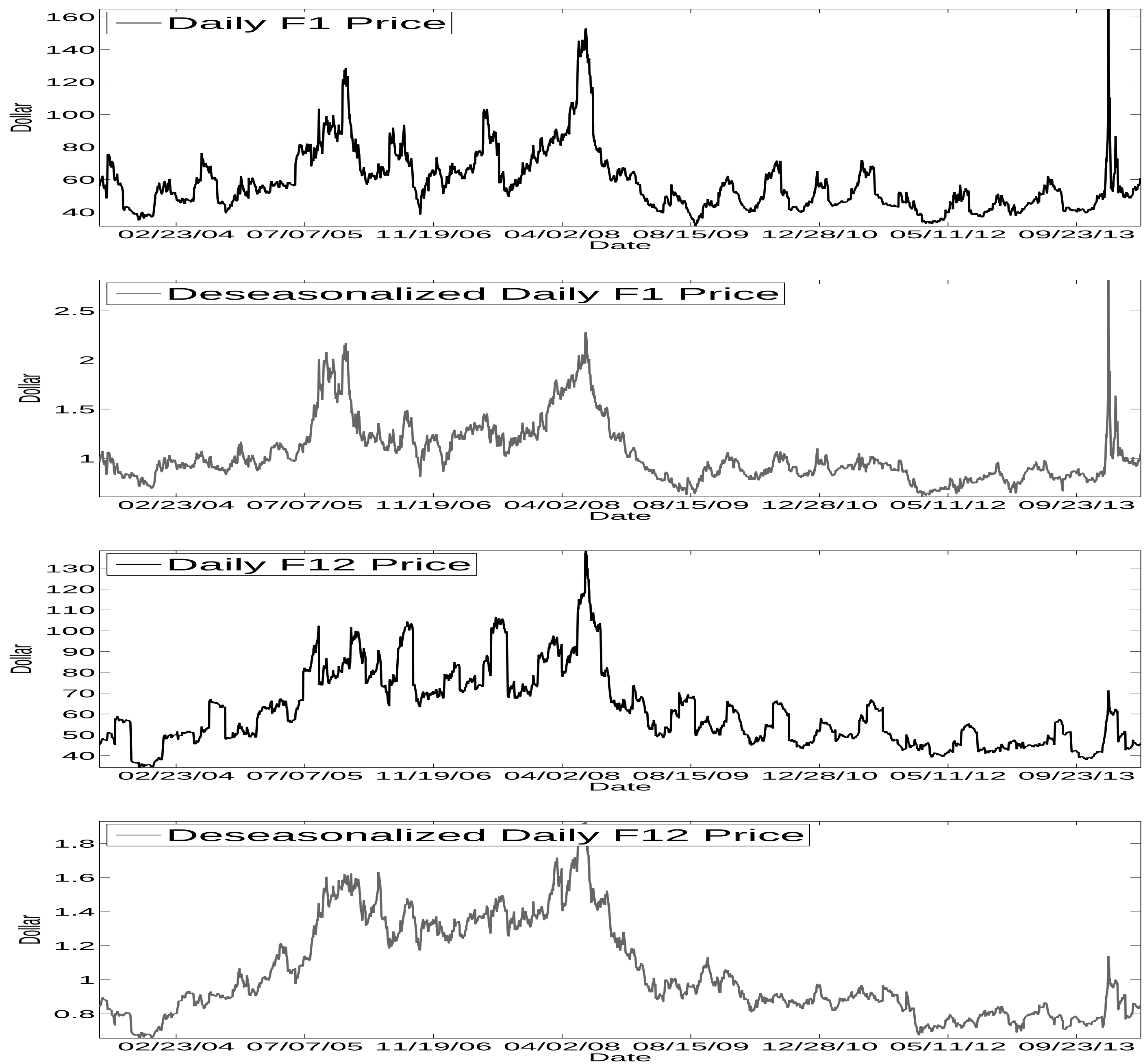


Figure 2

The Natural Gas Price

We plot the time series of the supply variable, the natural gas price. The top panel plots the raw price series and the bottom panel plots the corresponding deseasonalized part. The deseasonalized part is the residual from regressing the raw series on monthly dummies. The sample period is from May 1, 2003 to May 30, 2014.

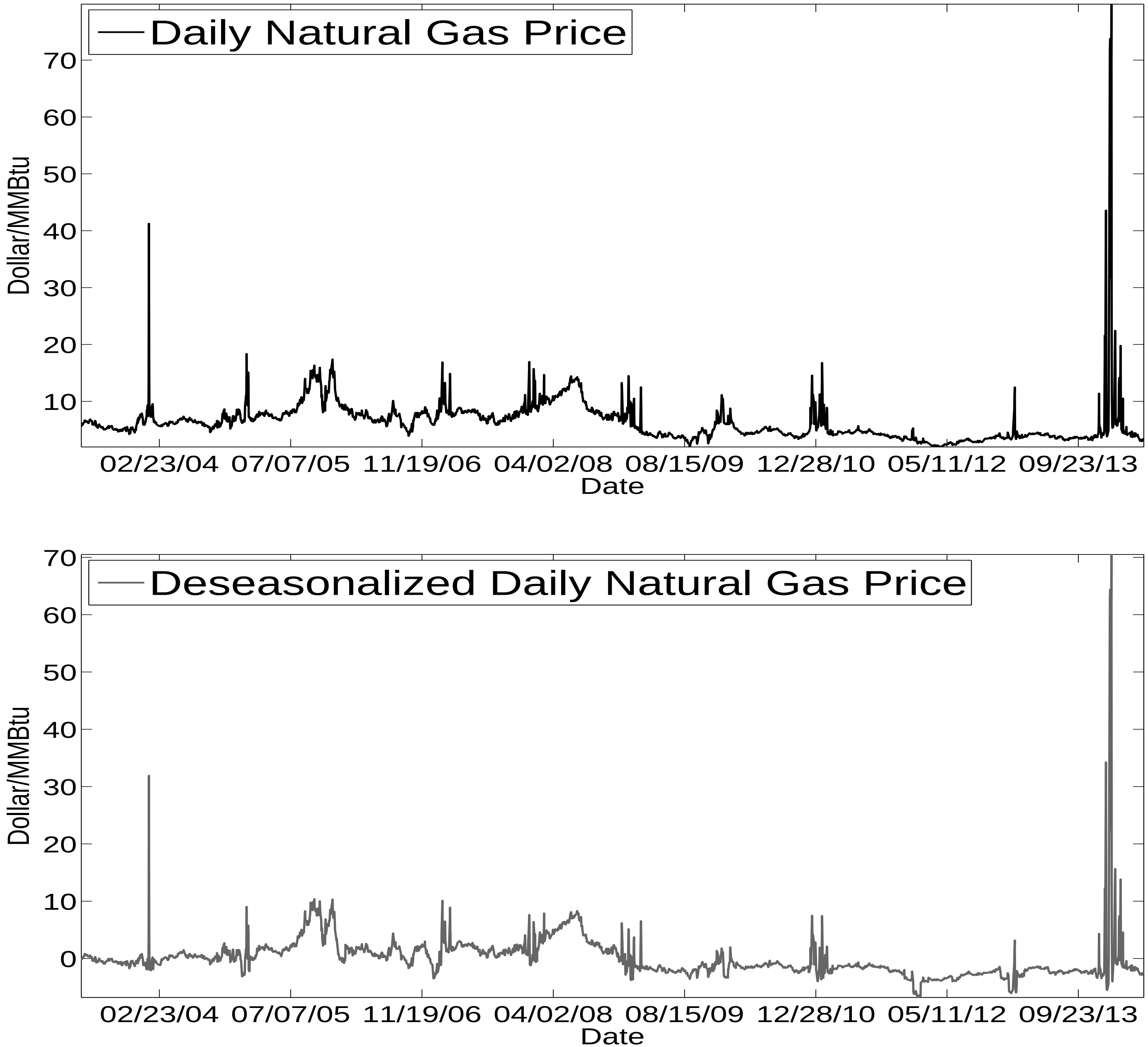


Figure 3

The Demand Variables

We plot the time series of the demand variables. The top two panels plot the daily maximum load and the deseasonalized maximum load. The bottom two panels plot the CDD and the deseasonalized CDD. The deseasonalized part is the residual from regressing the raw series on monthly dummies. The sample period is from May 1, 2003 to May 30, 2014.

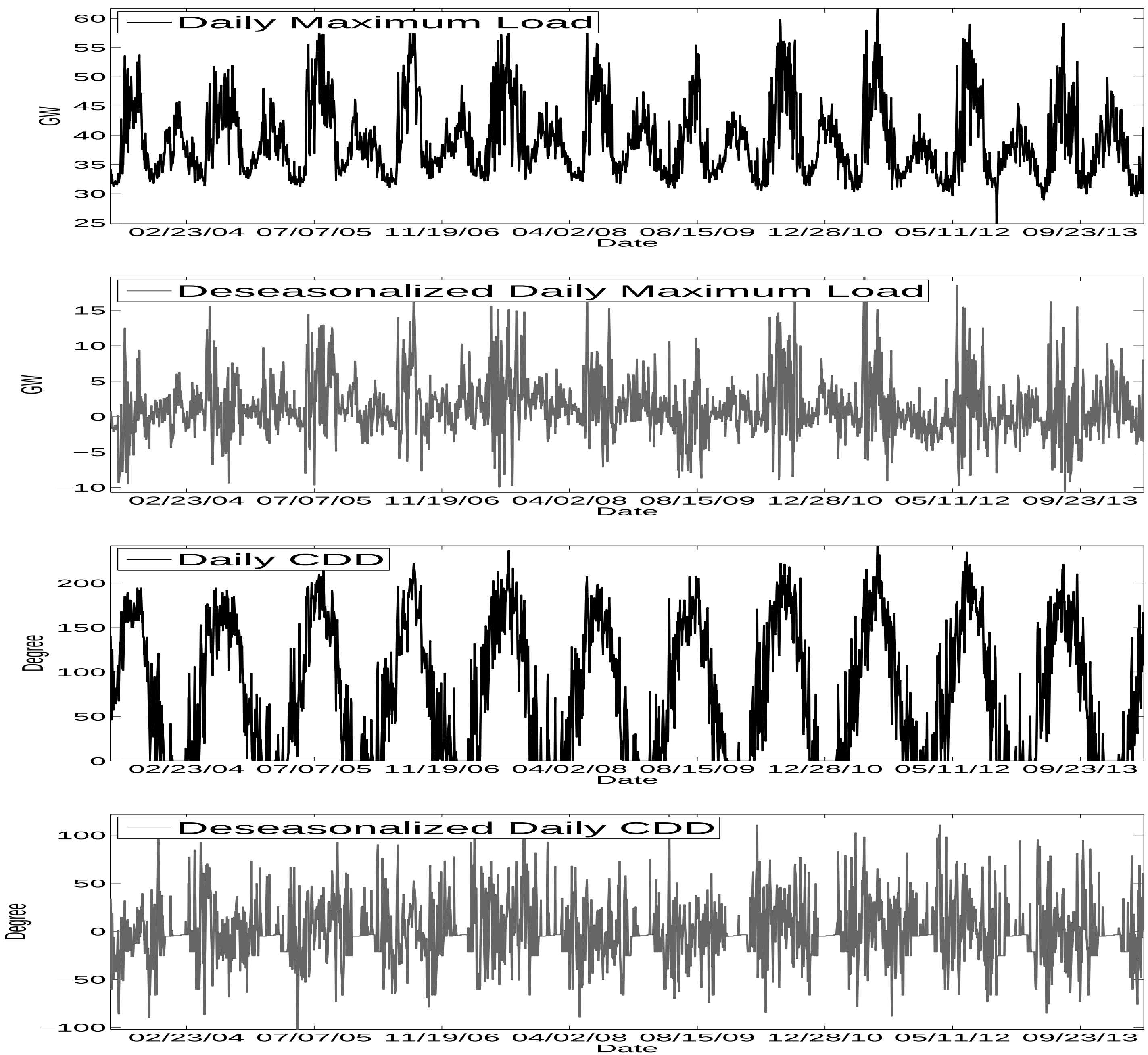


Figure 4
Principal Components of Electricity Futures

We plot the loadings of the first three principal components (PCs) of log futures prices. The first PC is plotted in blue, the second PC is plotted in green, and the third PC is plotted in red. The legend displays the fraction of the total variance explained by each of the principal components. The sample period is from May 1, 2003 to May 30, 2014.

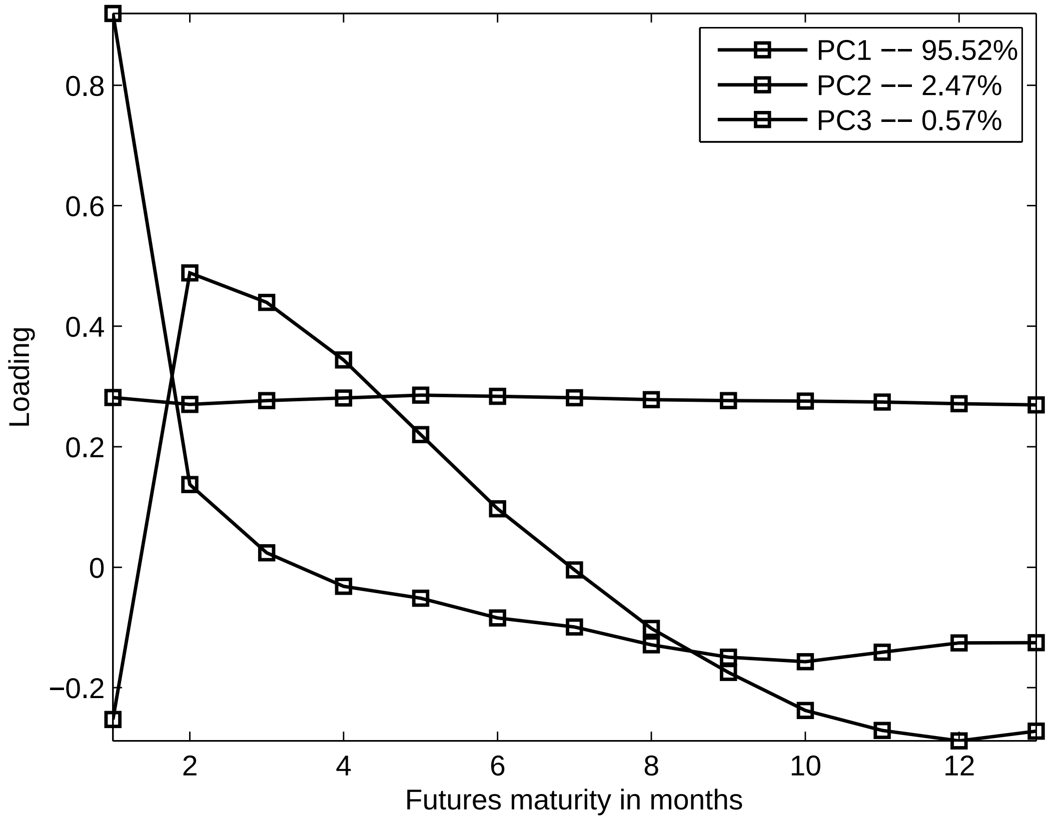


Figure 5

The Unspanned Model. In-Sample Fit

We plot model prices and actual prices for two futures contracts. The top panel reports on the 1-month futures contract F1. The bottom panel reports on the 12-month futures contract F12. The black solid line refers to the realized futures price and the grey dotted line refers to the futures price estimated by the unspanned model. The sample period is from May 1, 2003 to May 30, 2014.

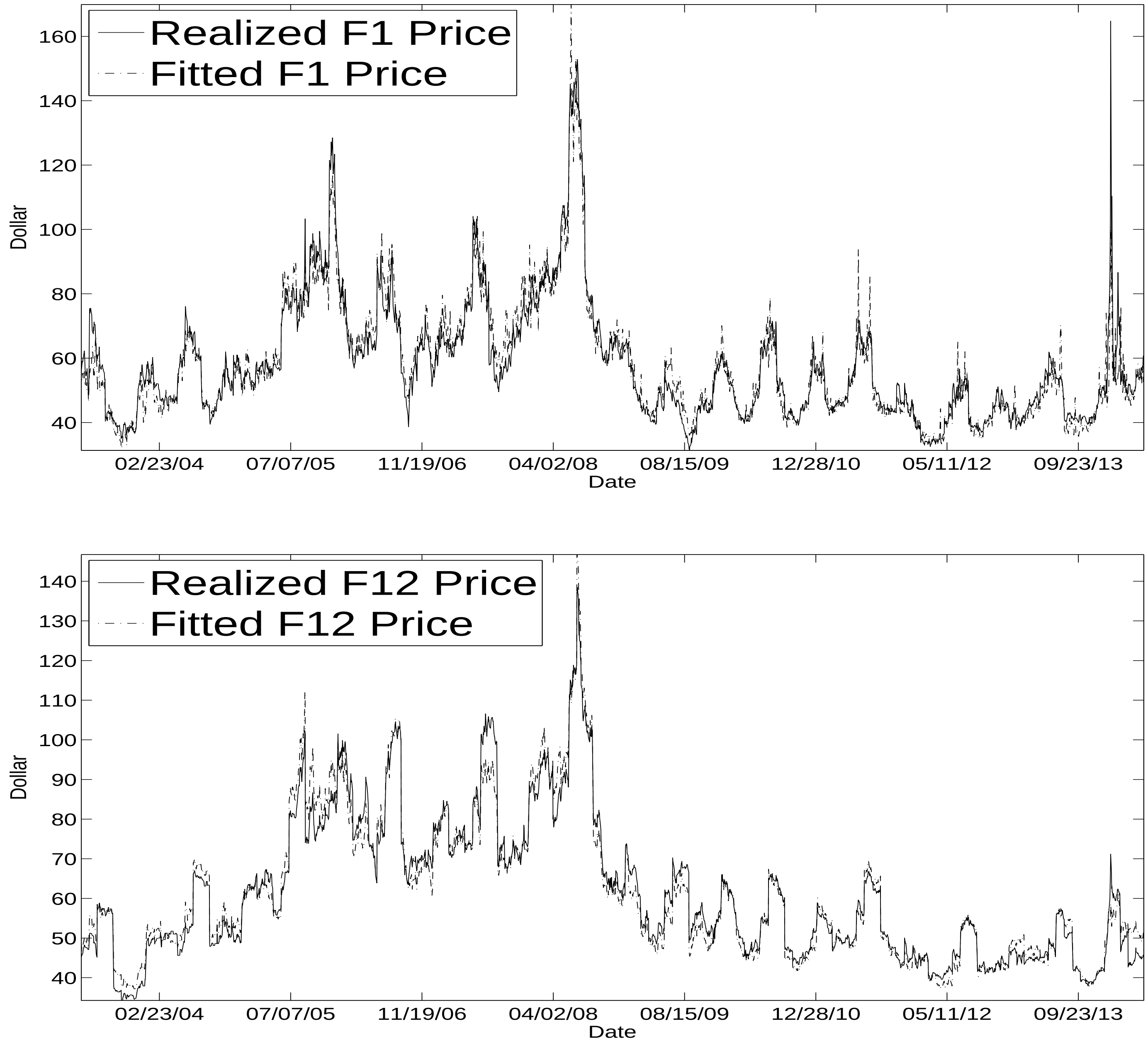


Figure 6

Spot Premiums. Various Models

We plot the spot premium for different models. The top panel plots the spot premium for the unspanned model with the natural gas price and the CDD. The second panel plots the spot premium for the spanned model. The third panel plots the spot premium for the spanned model with the natural gas price and the CDD. The bottom panel plots the spot premium of the Lucia-Schwartz model. See the appendix for details about the spot premium. The sample period is from May 1, 2003 to May 30, 2014.

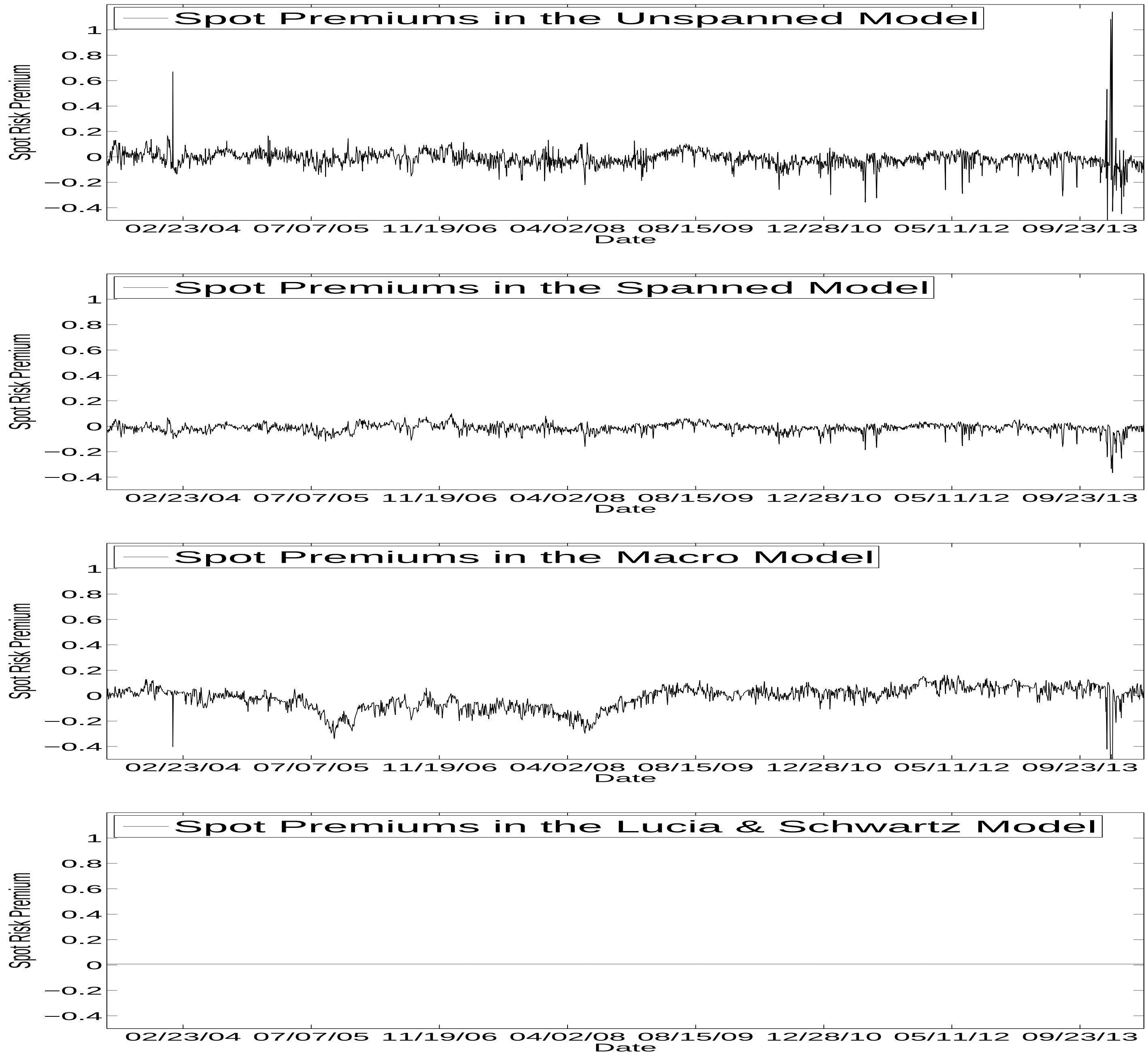


Figure 7

Decomposing the Spot Premium. Unspanned Model with Natural Gas Price and Maximum Load

We plot the decomposition of the spot premium for the unspanned model with the natural gas price and the maximum load. We plot the model-implied spot premium, the contribution from electricity itself , the contribution from the natural gas price, and the contribution from the maximum load. See the appendix for details about the decomposition. The sample period is from May 1, 2003 to May 30, 2014.

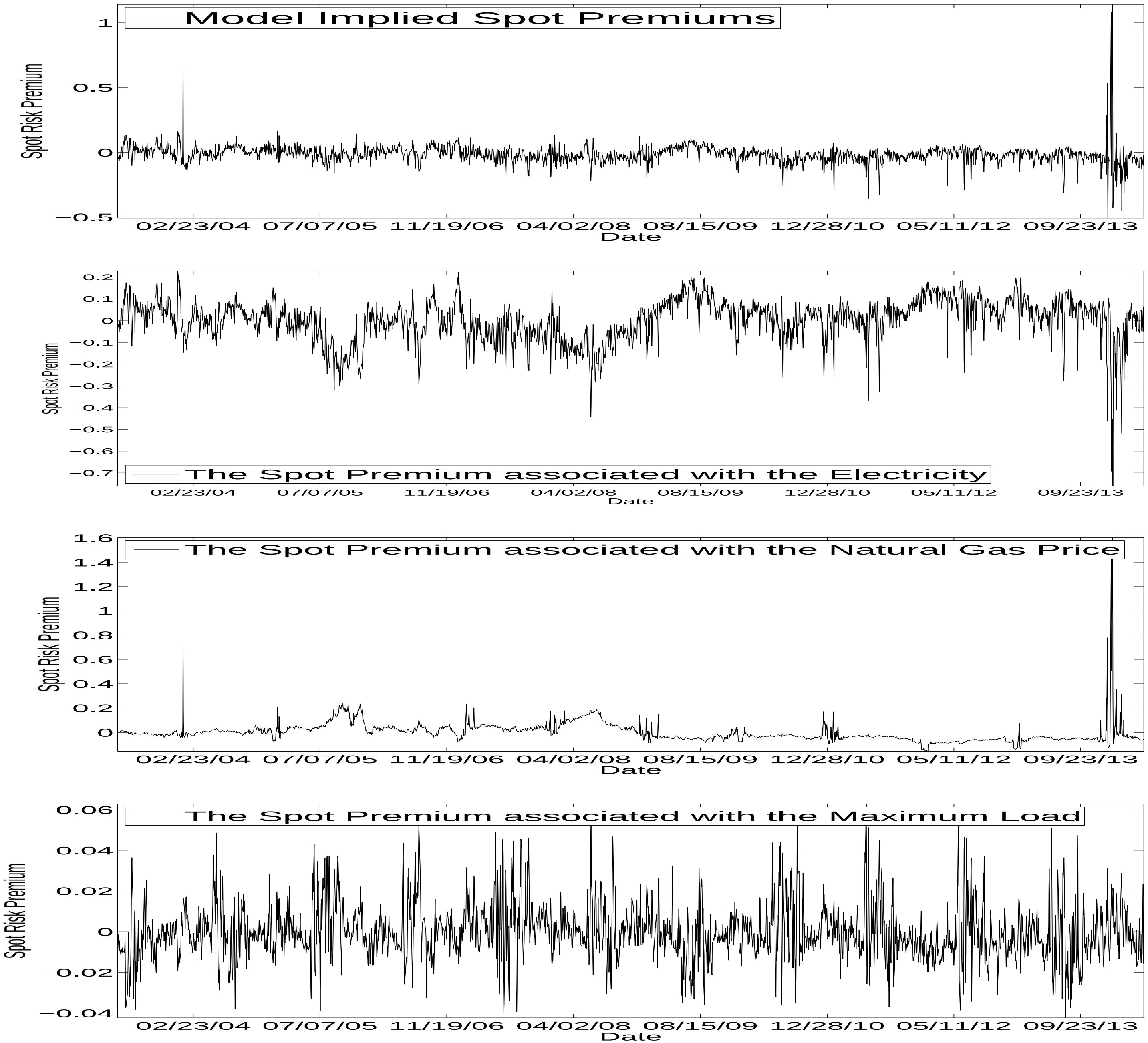


Figure 8

Decomposing the Spot Premium. Unspanned Model with Natural Gas Price and CDD

We plot the decomposition of the spot premium for the unspanned model with the natural gas price and the CDD. We plot the model-implied spot premium, the contribution from electricity itself, the contribution from the natural gas price, and the contribution from the CDD. See the appendix for details about the decomposition. The sample period is from May 1, 2003 to May 30, 2014.

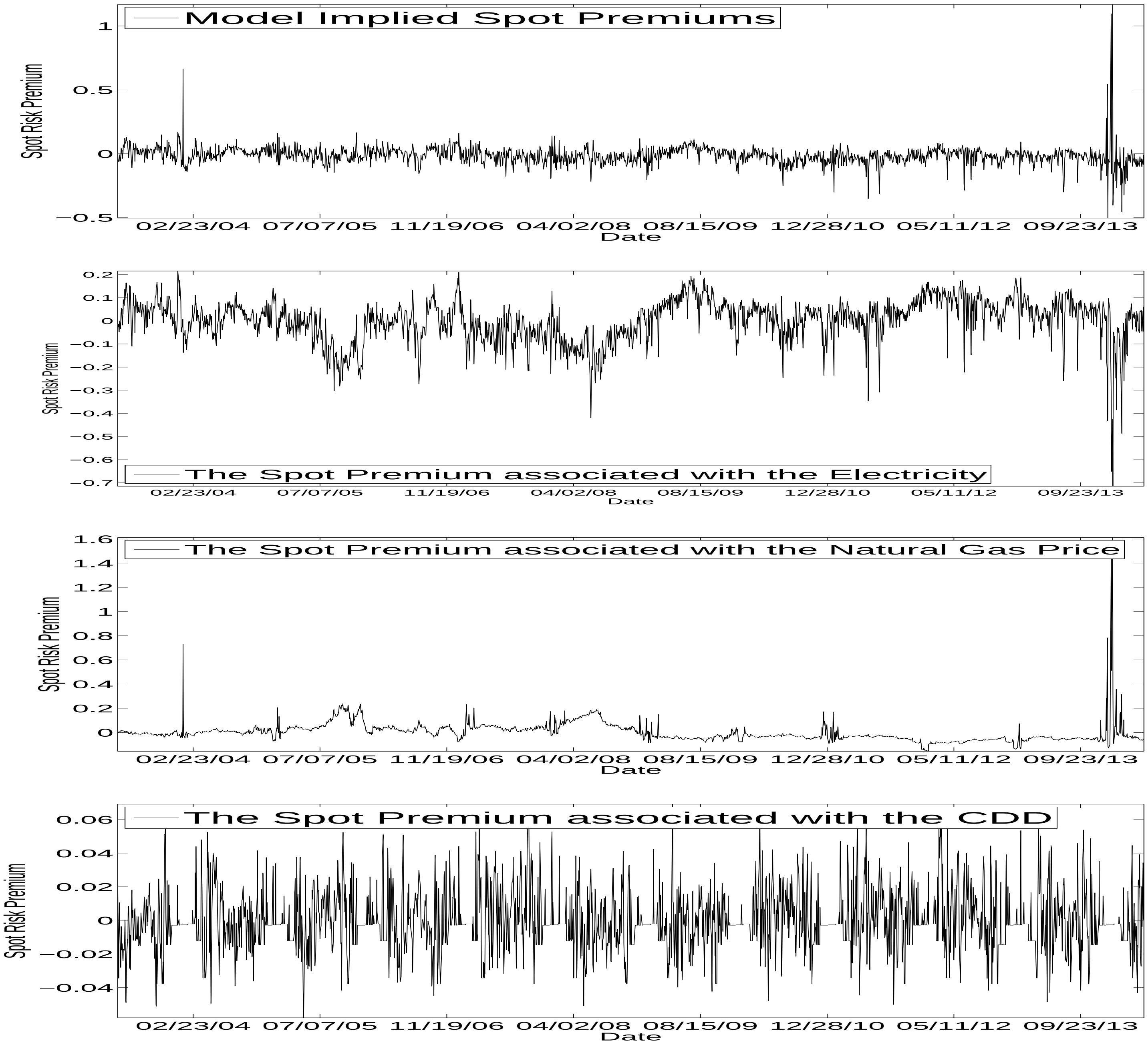


Figure 9
Forward Premiums of Unspanned Model

This figure plots the model implied forward premiums of different horizons. The forward premium is defined as the difference between the expected forward rate in P and the expected forward rate in Q. The definition is as follows.

$$\text{Foward Premium}_t^{i\to i+j} = \text{Implied Foward Rate in P}_t^{i\to i+j} - \text{Implied Foward Rate in Q}_t^{i\to i+j}$$

$$\text{Implied Foward Rate in P}_t^{i\to i+j} = (E^P[f_t^{i+j}]/E^P[f_t^{i+j}])^{\frac{1}{j}} - 1$$

$$\text{Implied Foward Rate in Q}_t^{i\to i+j} = (E^Q[f_t^{i+j}]/E^Q[f_t^{i+j}])^{\frac{1}{j}} - 1$$

where f_t^i is the expected price of futures maturing in j at time t. The blue line is the 3 month to 1 month forward premium. The red line is the 6 month to 1 month. The yellow line is the 9 month to 1 month. The sample period is from November. 2013 to April, 2014.

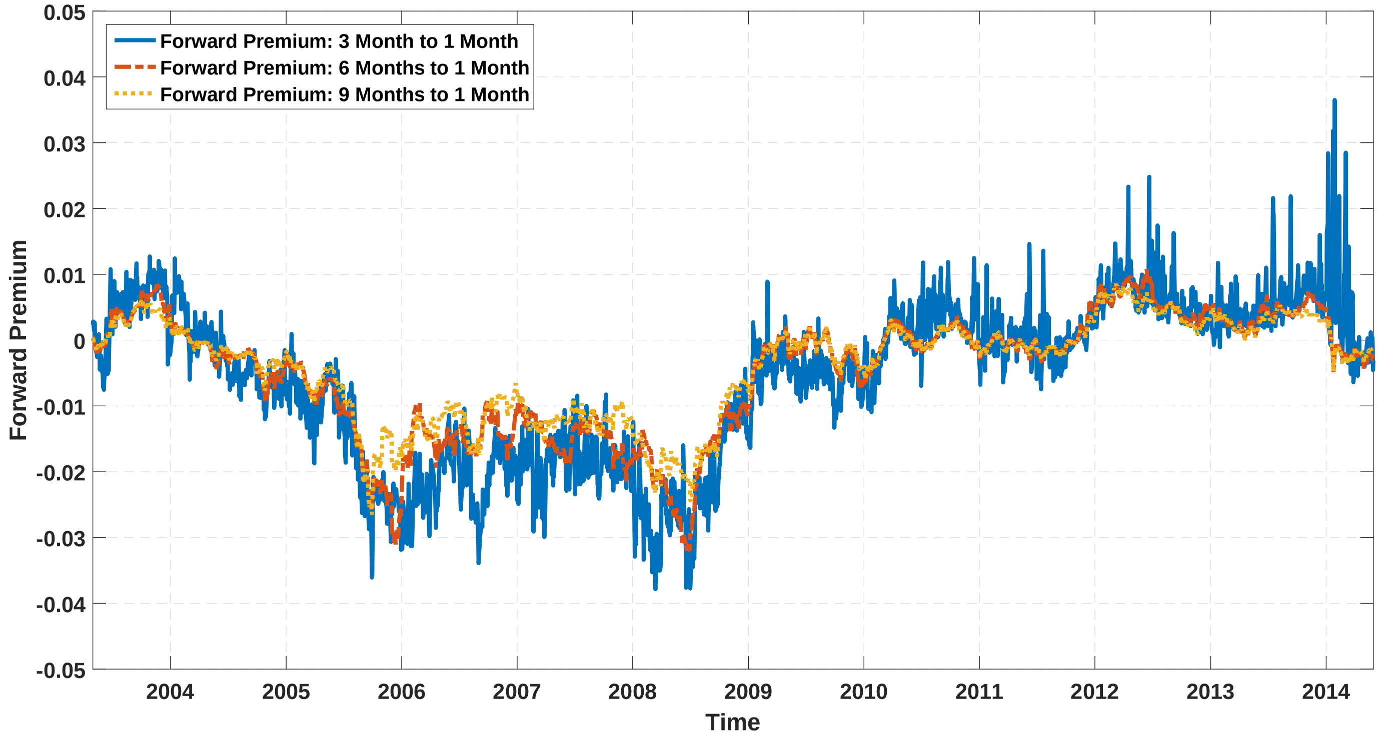


Figure 10
Comparison between Forward Premiums of Unspanned Model with Forward Premiums of Spanned Model

This figure plots the model implied 3 to 1 forward premiums of unspanned model (top panel), spanned model (middle) panel, and their differences (bottom panel). The sample period is from November. 2013 to April, 2014.

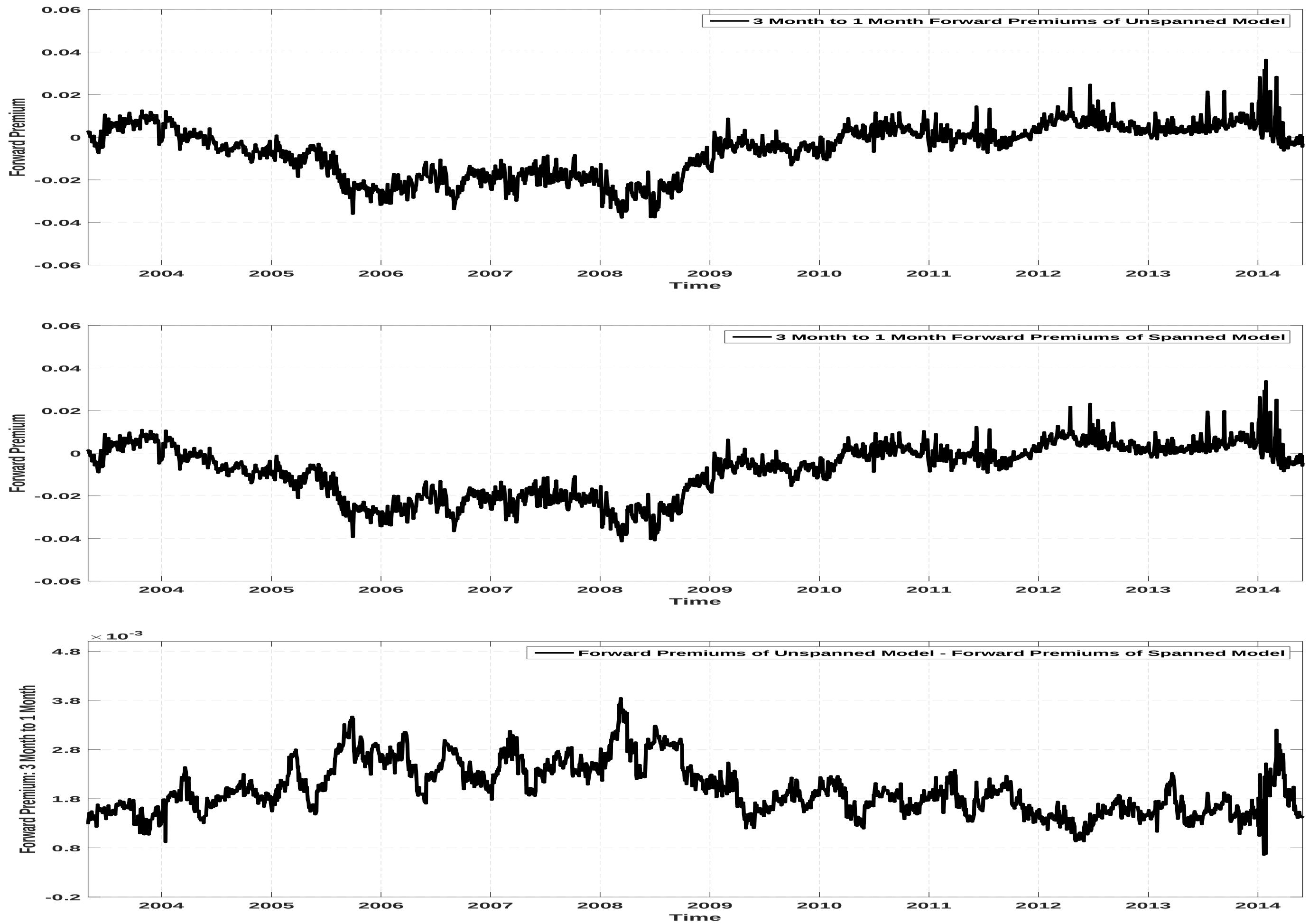
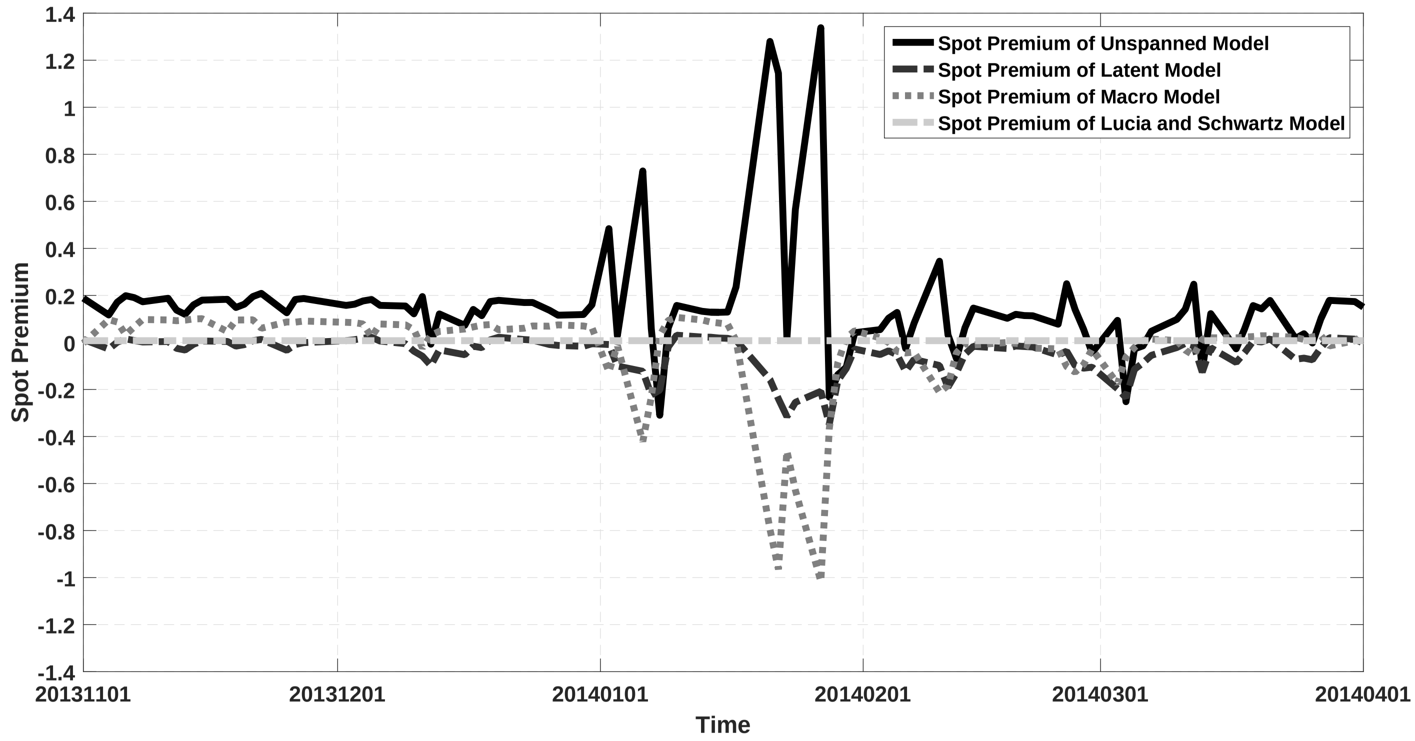


Figure 11
Spot Premiums During Polar Vortex

This figure plots the model implied spot premiums of different models during the 2014 Polar Vortex period. The sample period is from November. 2013 to April, 2014.



Appendix

Definition of the Spot Premium

Szymanowska et al. (2014) define the spot risk premium $\pi_{s,t}$ as the expected spot return in excess of the one-period basis,

$$\pi_{s,t} = E_t^P[r_{s,t+1}] - y_t^{(1)} = E_t^P[ln(S_{t+1}) - ln(S_t)] - y_t^{(1)} = E_t^P[s_{t+1} - s_t] - y_t^{(1)} \tag{50}$$

The one-period basis $y_t^{(1)}$ is defined as

$$y_t^{(1)} = log(F_t^{(1)}) - log(S_t) = f_t^{(1)} - s_t \tag{51}$$

where $F_t^{(1)}$ is the price of the futures contract maturing in 1 period. $y_t^{(1)}$ measures the one period cost of carry, analogous to a bond's one-period interest rate.

The spot premium, $\pi_{s,t}$ is defined as the expected return in excess of the one-period basis, in the manner of stock returns in excess of the short-term interest rate (and adjusted for the dividend yield).

Heath (2016) shows the spot premium can be expressed as follows in the unspanned model.

$$\begin{aligned}
\pi_{s,t} &= E_t^P[s_{t+1} - s_t] - y_t^{(1)} \\
&= E_t^P[s_{t+1}] - E_t^P[s_t] - y_t^{(1)} \\
&= E_t^P[s_{t+1}] - s_t - y_t^{(1)} \\
&= E_t^P[s_{t+1}] - (s_t + y_t^{(1)}) \\
&= E_t^P[s_{t+1}] - f_t^{(1)} \\
&= E_t^P[s_{t+1}] - (E_t^Q[s_{t+1}] + \frac{1}{2}\sigma_s^2) \\
&= E_t^P[s_{t+1}] - E_t^Q[s_{t+1}] - \frac{1}{2}\sigma_s^2
\end{aligned}$$

Decomposition of the Spot Premium

The unspanned model implies that

$$E_t^P[s_{t+1}] = K_0^P(1) + K_{1,s}^P s_t + K_{1,PC^2}^P PC_t^2 + K_{1,PX}^P PX_t + K_{1,Load/T}^P Load/T_t \tag{52}$$

$$E_t^Q[s_{t+1}] = K_0^Q(1) + K_{1,s}^Q s_t + K_{1,PC^2}^Q PC_t^2 + K_{1,PX}^Q PX_t + K_{1,Load/T}^Q Load/T_t \tag{53}$$

where $K_0^P(1)$ is the first element of the column vector K_0^P, $K_0^Q(1)$ is the first element of the column vector K_0^Q, $K_{1,s}^P$ is the coefficient of the matrix K_1^P for s and so on.

Then the spot premium can be written as

$$
\begin{aligned}
\pi_{s,t} &= E_t^P[s_{t+1}] - E_t^Q[s_{t+1}] - \frac{1}{2}\sigma_s^2 \\
&= (K_0^P(1) - K_0^Q(1) - \frac{1}{2}\sigma_s^2) + \underbrace{(K_{1,s}^P s_t - K_{1,s}^Q s_t + K_{1,PC^2}^P PC_t^2 - K_{1,PC^2}^Q PC_t^2)}_{\text{Spot premium associated with electricity}} \\
&+ \underbrace{(K_{1,PC^2}^P - K_{1,PX}^Q)PX_t}_{\text{Spot premium associated with PX}} + \underbrace{(K_{1,Load/T}^P - K_{1,Load/T}^Q)Load/T_t}_{\text{Spot premium associated with Load/Temperature}} \\
&= Constant + \pi_{s,t}^{electricity} + \pi_{s,t}^{PX} + \pi_{s,t}^{Load/T}
\end{aligned}
$$

The Spot Premium in the Lucia-Schwartz Model

Lucia and Schwartz model uses two factors to model the log price, i.e. $ln(S_t) = X_t + \epsilon_t$. From the equation (1) and (2) in SS model, the dynamics of two factors are given by

$$dX_t = -\kappa X_t dt + \sigma_X dz_X \tag{54}$$

$$d\epsilon_t = \mu_\epsilon dt + \sigma_\epsilon dz_\epsilon \tag{55}$$

and

$$dz_X dz_\epsilon = \rho_{x\epsilon} dt \tag{56}$$

Thus we can have

$$
\begin{aligned}
lnE[(S_{t+1})] &= E[ln(S_t)] + \frac{1}{2}Var[ln(S_t)] \\
&= e^{-\kappa}X_0 + \epsilon_0 + \mu_\epsilon \\
&+ \frac{1}{2}((1 - e^{-2\kappa})\frac{\sigma_X^2}{2\kappa} + \sigma_\epsilon^2 + 2(1 - e^{-\kappa})\frac{\rho_{x\epsilon}\sigma_X\sigma_\epsilon}{\kappa})
\end{aligned} \tag{57}
$$

The Q dynamic is given in the equation (7a) and (7b), which are

$$dX_t = (-\kappa X_t - \lambda_X)dt + \sigma_X dz_X^* \tag{58}$$

$$d\epsilon_t = (\mu_\epsilon - \lambda_\epsilon)dt + \sigma_\epsilon dz_\epsilon^* \tag{59}$$

and

$$dz_X^* dz_\epsilon^* = \rho_{x\epsilon} dt \tag{60}$$

Thus the expected futures price at time t+1 is

$$\begin{aligned} ln(F_{t+1}) = ln(E^Q[S_{t+1}]) &= E^Q[ln(S_t)] + \frac{1}{2}Var^Q[ln(S_t)] \\ &= e^{-\kappa}X_0 + \epsilon_0 + \mu_\epsilon^* - (1-e^\kappa)\frac{\lambda_X}{\kappa} \\ &+ \frac{1}{2}((1-e^{-2\kappa})\frac{\sigma_X^2}{2\kappa} + \sigma_\epsilon^2 + 2(1-e^{-\kappa})\frac{\rho_{x\epsilon}\sigma_X\sigma_\epsilon}{\kappa}) \end{aligned} \tag{61}$$

where $\mu_\epsilon^* = \mu_\epsilon - \lambda_\epsilon$.

The spot premium is defined as

$$lnE_t[(S_{t+1})] - ln(F_{t+1}) = \lambda_\epsilon + (1-e^\kappa)\frac{\lambda_X}{\kappa} \tag{62}$$

Online Appendix

Table A.1
Projecting Demand and Supply Variables on Principal Components

We report results from the projection of the demand and supply variables M_t on the first five principal components (PC_t^{1-5}) of electricity futures, i.e.

$$M_t = \alpha + \gamma_{pc} PC_t^{1-5} + UM_t$$

Columns (1) to (5) are the regression results for the average load ($Load^{avg}$), heating degree days (HDD), the maximum temperature (T_t^{max}), the minimum temperature (T_t^{min}) and the average temperature (T_t^{avg}) respectively. For each regression, the adjusted R^2 and the number of observations are reported. Standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

	(1)	(2)	(3)	(4)	(5)
	$Load_t^{avg}$	HDD_t	T_t^{max}	T_t^{min}	T_t^{avg}
PC_t^1	0.187***	0.052***	−0.014	−0.019	−0.017
	(0.010)	(0.012)	(0.010)	(0.010)	(0.010)
PC_t^2	1.676***	1.403***	−0.244***	−0.290***	−0.028***
	(0.034)	(0.045)	(0.037)	(0.035)	(0.034)
PC_t^3	−1.298***	−0.660***	−0.103	−0.048	−0.078
	(0.068)	(0.090)	(0.075)	(0.071)	(0.069)
PC_t^4	1.391***	0.217	0.137	0.443**	0.287*
	(0.137)	(0.181)	(0.150)	(0.142)	(0.138)
PC_t^5	0.860***	0.259	0.013	0.112	0.061
	(0.139)	(0.184)	(0.152)	(0.145)	(0.140)
Constant	0.000	0.000	0.000	0.000	0.000
	(0.008)	(0.008)	(0.009)	(0.009)	(0.009)
Adj. R^2	0.551	0.279	0.018	0.028	0.026
N. Obs.	2722	2722	2722	2722	2722

Table A.2
Forecasting PC Changes with Unspanned Demand and Supply Variables

We report results from forecasting PC changes using the unspanned components of five different demand and supply variables. For all pairs, the supply variable is the unspanned natural gas price (unPX). The demand variables are the unspanned average load ($unLoad^{avg}$), heating degree days ($unHDD$), the maximum temperature (unT_t^{max}), the minimum temperature (unT_t^{min}) and the average temperature (unT_t^{avg}) respectively. The regression is specified as follows.

$$\Delta PC_{t\to t+1}^{1-2} = \alpha + \beta_{pc} PC_t^{1-5} + \beta_{um} UM_t + \epsilon_t$$

The first five PCs, i.e. PC_t^{1-5}, are included as control variables. The top panel reports results for the first PC. The bottom panel reports results for the second PC. For each regression, the adjusted R^2 and the number of observations are reported. Standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

	$\Delta PC_{t\to t+1}^1$	$\Delta PC_{t\to t+1}^1$	$\Delta PC_{t\to t+1}^1$	$\Delta PC_{t\to t+1}^1$	$\Delta PC_{t\to t+1}^1$	$\Delta PC_{t\to t+1}^1$
$unPX_t$		0.027***	0.026***	0.026***	0.026***	0.026***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$unLoad_t^{avg}$		0.004				
		(0.003)				
$unHDD_t$			−0.004*			
			(0.002)			
unT_t^{max}				0.012***		
				(0.003)		
unT_t^{max}					0.013***	
					(0.003)	
unT_t^{avg}						0.0134***
						(0.003)
PC^1 to PC^5	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.044	0.138	0.127	0.133	0.135	0.135
N. Obs.	2117	2117	2117	2117	2117	2117
	$\Delta PC_{t\to t+1}^2$	$\Delta PC_{t\to t+1}^2$	$\Delta PC_{t\to t+1}^2$	$\Delta PC_{t\to t+1}^2$	$\Delta PC_{t\to t+1}^2$	$\Delta PC_{t\to t+1}^2$
$unPX_t$		0.027***	0.077***	0.079***	0.079***	0.079***
		(0.002)	(0.004)	(0.004)	(0.004)	(0.004)
$unLoad_t^{avg}$		0.011				
		(0.006)				
$unHDD_t$			−0.010*			
			(0.004)			
unT_t^{max}				0.033***		
				(0.006)		
unT_t^{min}					0.026***	
					(0.006)	
unT_t^{avg}						0.034***
						(0.006)
PC^1 to PC^5	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.123	0.364	0.257	0.262	0.257	0.261
N. Obs.	2117	2117	2117	2117	2117	2117

Table A.3
Forecasting PC Changes with Unspanned Demand and Supply Variables and Lagged PCs

We report results from forecasting changes in the first two principal components (ΔPC^{1-2}) with the first five principal components (PC_t^{1-5}), the unspanned components of the demand and supply variables (UMs), and the lagged change of the PCs, i.e.

$$\Delta PC_{t\to t+1}^{1-2} = \alpha + \beta_{pc} PC_t^{1-5} + \beta_{um} UM_t + \beta_{L.pc} \Delta PC_{t-1\to t}^{1-2} + \epsilon_t$$

We include PC_t^{1-5} as control variables. For each regression, the adjusted R^2 and the number of observations are reported. Standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$	$\Delta PC^1_{t\to t+1}$
$\Delta PC^1_{t-1\to t}$	0.096***	0.091***	0.088***	0.084***	0.087***	0.086***	0.086***	0.085***
	0.019	0.018	0.018	0.018	0.018	0.018	0.018	0.018
$unPX_t$		0.026***	0.027***	0.027***	0.026***	0.027***	0.027***	0.027***
		(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
$unLoad_t^{avg}$		0.006*						
		(0.003)						
$unLoad_t^{max}$			0.010***					
			(0.003)					
$unCDD_t$				0.016***				
				(0.003)				
$unHDD_t$					−0.005*			
					(0.002)			
unT_t^{max}						0.011***		
						(0.003)		
unT_t^{min}							0.013***	
							(0.003)	
unT_t^{avg}								0.0134***
								(0.003)
PC^1 to PC^5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.057	0.137	0.141	0.147	0.137	0.143	0.144	0.145
N. Obs.	2116	2116	2116	2116	2116	2116	2116	2116
	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$	$\Delta PC^2_{t\to t+1}$
$\Delta PC^2_{t-1\to t}$	0.149***	0.138***	0.133***	0.125***	0.126***	0.124***	0.126***	0.124***
	0.019	0.018	0.017	0.017	0.018	0.017	0.018	0.018
$unPX_t$		0.076***	0.078***	0.078***	0.075***	0.077***	0.077***	0.078***
		(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
$unLoad_t^{avg}$		0.021***						
		(0.006)						
$unLoad_t^{max}$			0.036***					
			(0.006)					
$unCDD_t$				0.041***				
				(0.007)				
$unHDD_t$					−0.009			
					(0.005)			
unT_t^{max}						0.031***		
						(0.006)		
unT_t^{min}							0.025***	
							(0.006)	
unT_t^{avg}								0.033***
								(0.006)
PC^1 to PC^5	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Adj. R^2	0.148	0.273	0.280	0.282	0.270	0.279	0.274	0.278
N. Obs.	2116	2116	2116	2116	2116	2116	2116	2116

Table A.4
Estimating the Physical Dynamics. Various Demand and Supply Variables

We report estimates of the P dynamics using various demand and supply pairs. For all pairs, the supply variable is the natural gas price (PX). The demand variables are the average load ($Load^{avg}$), the heating degree days (HDD), the maximum temperature (T_t^{max}), the minimum temperature (T_t^{min}) and the average temperature (T_t^{avg}) respectively. Standard errors are reported in parentheses. *, **, and *** denote the significance at the 5%, 1%, and 0.1% respectively. The sample period is from May 1, 2003 to May 30, 2014.

	K_0^P	K_1^P			
		s_t	PC_t^2	PX_t	$Load_t^{avg}$
s_{t+1}	−0.014***	0.775***	−0.085***	0.086***	0.001
	(0.003)	(0.017)	(0.018)	(0.004)	(0.006)
PC_{t+1}^2	−0.013***	−0.208***	0.915***	0.080***	−0.001
	(0.003)	(0.016)	(0.017)	(0.004)	(0.006)
PX_{t+1}	−0.021*	1.055***	−0.913***	0.559***	−0.046*
	(0.009)	(0.049)	(0.053)	(0.013)	(0.019)
$Load_{t+1}^{avg}$	−0.008	0.041	0.261***	0.053***	0.706***
	(0.008)	(0.046)	(0.049)	(0.011)	(0.018)
		s_t	PC_t^2	PX_t	HDD_t
s_{t+1}	−0.014***	0.775***	−0.065***	0.087***	−0.014**
	(0.003)	(0.015)	(0.018)	(0.004)	(0.005)
PC_{t+1}^2	−0.013***	−0.208***	0.932***	0.081***	−0.013***
	(0.003)	(0.014)	(0.017)	(0.004)	(0.004)
PX_{t+1}	−0.020*	1.009***	−0.919***	0.567***	−0.025*
	(0.009)	(0.045)	(0.054)	(0.012)	(0.015)
HDD_{t+1}	0.002	−0.428***	0.402***	0.201***	0.640***
	(0.009)	(0.050)	(0.059)	(0.013)	(0.016)

Table A.4
Estimating the Physical Dynamics. Various Demand and Supply Variables (Continued)

	K_0^P	K_1^P			
		s_t	PC_t^2	PX_t	T_t^{max}
s_{t+1}	−0.014***	0.770***	−0.074***	0.088***	0.035***
	(0.003)	(0.015)	(0.017)	(0.004)	(0.006)
PC_{t+1}^2	−0.013***	−0.213***	0.925***	0.083***	0.033***
	(0.002)	(0.015)	(0.016)	(0.003)	(0.006)
PX_{t+1}	−0.020*	1.011***	−0.955***	0.565***	−0.009
	(0.008)	(0.045)	(0.050)	(0.012)	(0.018)
T_{t+1}^{max}	0.000	0.201***	−0.159***	−0.090***	0.647***
	(0.007)	(0.039)	(0.044)	(0.010)	(0.016)
		s_t	PC_t^2	PX_t	T_t^{min}
s_{t+1}	−0.014***	0.770***	−0.076***	0.088***	0.029***
	(0.003)	(0.015)	(0.017)	(0.004)	(0.006)
PC_{t+1}^2	−0.014***	−0.213***	0.923***	0.083***	0.026***
	(0.003)	(0.015)	(0.016)	(0.003)	(0.006)
PX_{t+1}	−0.021*	1.001***	−0.947***	0.567***	0.0174
	(0.009)	(0.045)	(0.050)	(0.012)	(0.019)
T_{t+1}^{min}	0.008	0.192***	−0.157***	−0.081***	0.642***
	(0.007)	(0.037)	(0.041)	(0.010)	(0.016)
		s_t	PC_t^2	PX_t	T_t^{avg}
s_{t+1}	−0.015***	0.770***	−0.073***	0.089***	0.037***
	(0.003)	(0.015)	(0.017)	(0.004)	(0.003)
PC_{t+1}^2	−0.014***	−0.214***	0.926***	0.083***	0.034***
	(0.003)	(0.015)	(0.016)	(0.004)	(0.006)
PX_{t+1}	−0.021*	1.001***	−0.951***	0.566***	0.003
	(0.009)	(0.045)	(0.050)	(0.012)	(0.019)
T_{t+1}^{avg}	0.003	0.193***	−0.144***	−0.083***	0.694***
	(0.007)	(0.034)	(0.038)	(0.009)	(0.015)