A Non-Parametric Structural Hybrid Modeling Approach for Electricity Prices

Somayeh Moazeni[†*], Michael Coulon[‡], I. Arciniegas Rueda[§], B. Song[†] and Warren B. Powell[†]

[†]Department of Operations Research and Financial Engineering, Princeton University, Sherrerd Hall, Charlton Street, Princeton, New Jersey, USA 08544
[‡]Department of Business and Management, University of Sussex, Brighton BN1 9SL, UK
[§]PSEG Energy Resources and Trade, Public Service Enterprise Group (PSEG), 80 Park Plaza, Newark, New Jersey, USA 07101

July 13, 2015

Abstract

We develop a stochastic model of zonal/regional electricity prices, designed to reflect information in fuel forward curves and aggregated capacity and load as well as zonal or regional price spreads. We use a nonparametric model of the supply stack that captures heat rates and fuel prices for all generators in the market operator territory, combined with an adjustment term to approximate congestion and other zone-specific behavior. The approach requires minimal calibration effort, is readily adaptable to changing market conditions and regulations, and retains sufficient tractability for the purpose of forward price calibration. The model is illustrated for the spot and forward electricity prices of the PS zone in the PJM market, and the set of time-dependent risk premiums are inferred and analyzed.

Keywords: Electricity market; Electricity price modeling; Energy trading; Supply stack

1 Introduction

In today's dynamic deregulated electricity markets, energy firms need reliable and tractable models for electricity price behavior in order to address a variety of valuation, trading and risk management challenges. Such models can help to determine appropriate strategies for investing in new supply, accurately assessing existing risk exposures and effectively hedging these risks via derivative products such as forwards and options. Furthermore, accurate price forecasting also facilitates decisions by market operators regarding system planning and management.

In a deregulated electricity market, the market operator determines electricity prices by means of an optimization scheme involving the unit commitment and economic dispatch problems at different time scales,

*corresponding author

see, e.g., (Schweppe et al., 1988; Conejo et al., 2010) or specifically for the PJM market, (Ott, 2003). This procedure involves many input parameters provided by the market participants, and several types of constraints, making the required computation rather expensive. As a result, full-blown fundamental price models that accommodate all such details soon become inadequate for the needs of fast-paced trading and hedging. Various alternative approaches for modeling electricity prices have been proposed in the literature and are often categorized in different ways. These range from 'reduced-form' models focused purely on specifying stochastic price processes to a variety of 'hybrid' or 'structural' models, which attempt to remain closer to the full price formation mechanism but retain greater tractability through various approximation techniques.

Pure reduced-form price models are often adapted from traditional financial markets, and by definition fail to take into account the impact of any fundamental power price drivers such as fuel prices, capacity or load. Examples of such models include geometric mean-reverting processes with jumps (Johnson and Barz, 1999), stochastic volatility jump-diffusion models (Deng, 2000), regime-switching jump diffusion models (Thompson et al., 2004), or mean-reverting jump-diffusion models with seasonality (Cartea and Figueroa, 2005). For an overview of reduced-form electricity price models, see, e.g., (Barz and Johnson, 1999), Chapter 4 of (Eydeland and Wolyniec, 2002), or (Swindle, 2014). Given the clear links between electricity prices and observable supply and demand factors, many authors have suggested adapting reduced-form price processes to incorporate these influential exogenous factors. For example, the volatility or the jump probability of the electricity price can depend on temperature or on the reserve margin, see, e.g., (Skantze and Ilic, 2001; Davison et al., 2002), or Chapter 7 of (Eydeland and Wolyniec, 2002). These models acknowledge the impact of driving factors on the electricity price distribution, but they greatly simplify the relationship by neglecting the information contained in the known price formation mechanism and supply-demand equilibrium frameworks.

Moving one step further from reduced-form stochastic processes, so-called structural models represent electricity prices as explicit functions of a subset of the observable input variables for a given market, such as fuel prices or load. Examples of structural models can be found in (Barlow, 2002; Skantze et al., 2004; Kanamura and Ōhashi, 2007; Carmona and Coulon, 2013). These models differ in the transformation function specified or the number of input factors considered, but share an aim of approximating the important features of the price formation mechanism, typically while retaining a fair degree of tractability for derivative pricing purposes. Since establishing an explicit transformation function to exactly relate solutions of the economic dispatch optimization scheme to its input parameters is not trivial, the parametric transformation functions in these models can lead to an inaccurate representation of the sensitivity of electricity prices to input factors, as we shall discuss in greater detail in Section 3.

Both reduced-form and structural models can face significant challenges when adapting to new electricity markets or to different subregions, and often considerable modifications to the choices of stochastic processes or parametric functional approximations may be required. Furthermore, as they rely heavily on the use of historical data for model calibration, technological developments or changes in market regulations, generation mix, or participants' behavior can all reduce the reliability of the estimated parameter values. In addition, when these models are adopted for modeling regional or zonal electricity prices, only the (local) supply and demand variables associated with that specific region are incorporated and the impact of interconnected capacity and load from outside the studied region is typically ignored.

In light of the above observations, the present paper aims to enhance the structural approach by instead establishing a detailed generation supply curve directly from generator cost data, avoiding parametric transformation functions and overreliance on history, while still retaining sufficient simplicity for derivatives pricing. Our proposed zonal spot price model consists of an energy price component and an adjustment term. The energy price (fundamental component) is computed from an approximate economic dispatch

optimization problem for the entire market operator territory, which captures the full generation stack (a non-parametric transformation) and the forecasted demand for the entire market operator region. This stack structure allows the modeling framework to capture the correlation between electricity prices and total capacity and demand, as well as fuel prices. As the second step of the methodology, an adjustment term is computed to map the energy price to the zonal electricity price. This component reflects the effects of inter-zonal congestion[1] and marginal losses related to the target zone or the neighboring zones, as well as the approximation error in the derivation of the first term. This component is determined through calibration to historical real-time electricity prices for the region of interest and thus takes into account the information content in the historical data.

We investigate the performance of the proposed model using historical price data for the PS zone in the PJM electricity market. Our analysis indicates that the model captures important properties of electricity prices such as fat tail behavior; the simulated hourly probability distributions match the historical probability distributions quite well, and the moments are in agreement. Although the framework involves an embedded optimization, we demonstrate that forward electricity prices can be derived by means of a computationally tractable method. This procedure allows us to match the model generated forwards to the market data and infer the set of time-dependent risk premiums. We analyze the pattern and stability of the inferred risk premiums using the available PS zone power forward prices. In particular, we observe that for a given pricing date, the risk premiums for different maturities are far from being a constant, as assumed in (de Maere d'Aertrycke and Smeers, 2010), and maintains some seasonality pattern. For all pricing dates we considered, we observe that the risk premiums are often higher during the summer month maturities than winter month maturities. In addition, risk premiums for on-peak hour deliveries are significantly higher than risk premiums for off-peak hours maturities. For a given pricing date, the inferred time-dependent risk premium curve can be extrapolated which then allows us, along with available long natural gas and coal forward prices, to compute power forward prices for those long maturities. The pattern of the risk premium curve changes gradually over different pricing dates, i.e., the risk premium curve of the pricing date August 1, 2013 is different from that of August 1, 2014. However, the difference is insignificant for consecutive pricing dates, indicating that applying the risk premium curve from yesterday can lead to a relatively good approximation of today's forward prices.

The proposed modeling framework has the advantage of being computationally tractable and requiring minimal estimation effort, as historical prices are only used to calibrate the adjustment term via a non-parametric probability distribution. The model mimics some of the computational benefits of the alternative parametric approaches, such as the important ability to rapidly compute expected spot prices and hence calibrate the model to entire fuel and electricity forward curves observed in the market. The developed bottom-up spot price modeling approach can easily be implemented for the entire market operator territory, for a hub, or a specific zone, by incorporating both aggregated and local capacity and load variables. In addition, since the derivation of the energy price component relies on an approximate price formation optimization problem and takes into account the key datasets from the entire market operator territory, the modeling approach can promptly (dynamically) adapt to changes in the price formation mechanism, regulations, or the levels of total supply or demand in the market operator region, by simply modifying the embedded optimization and its constraints or inputs. For instance, when several electricity markets are integrated (coupled), e.g., Midwest Independent Transmission System Operator (MISO) joins the PJM market, then the model can be readily adapted to predict impact of MISO capacity or demand on PJM zonal electricity prices. Similar accommodations can be made when a current power plant is shut down, a new generation unit starts working, or market regulations change. Each of these events may abruptly move

[1]The congestion effect is decreasing over years at least in the PJM market, see (Monitoring Analytics LLC Report, 2013).

the electricity prices in manner which traditional reduced-form models or regression-based models cannot capture due to an over-reliance on lengthy and local historical datasets. The introduced model in this paper combines advantages of both full fundamental and structural or hybrid models. Remaining close to the true price formation mechanism and exploiting generator-specific data, we provide greater intuition from an engineering perspective and capture more precisely the sensitivity of electricity prices to driving factors such as fuel prices.

Several authors have suggested related approaches that advocate cost-based approximations of the generation stack as starting points for spot price models. In Chapter 7 of (Eydeland and Wolyniec, 2002), a 'fundamental hybrid model' is suggested, which consists of a stack function constructed from all units' costs and models fuel prices, emissions costs, outage rates and load as underlying stochastic processes. The authors then introduce three tuning parameters that scale various components of the model in order to calibrate to observed spot or forward prices. However, the model is discussed only for the entire PJM market, and a full implementation, forward price calibration and empirical analysis is not included. Another related approach to ours is found in (de Maere d'Aertrycke and Smeers, 2010), where a simplified optimal dispatch model is embedded into a PDE based approach to solving for forward prices, building on the approach of (Pirrong and Jermakyan, 2008). Costs of individual generators are again included, but with the assumption of only fuel and emissions costs determining bids (ie, no other operational costs) in order to reduce dimensionality. A constant market price of risk is assumed due to the computation time required to solve the PDE, and this is estimated from data via an optimization routine. Instead, along the lines of some structural models (Coulon et al., 2013), we propose here a time dependent risk premium which can be solved for computationally efficiently and therefore used to exactly match the entire forward curve observed in the market. This is a crucial feature when aiming to then use a model for other purposes such as option pricing or plant valuation, or indeed for constructing forward curves beyond the range of liquid maturities in the market.

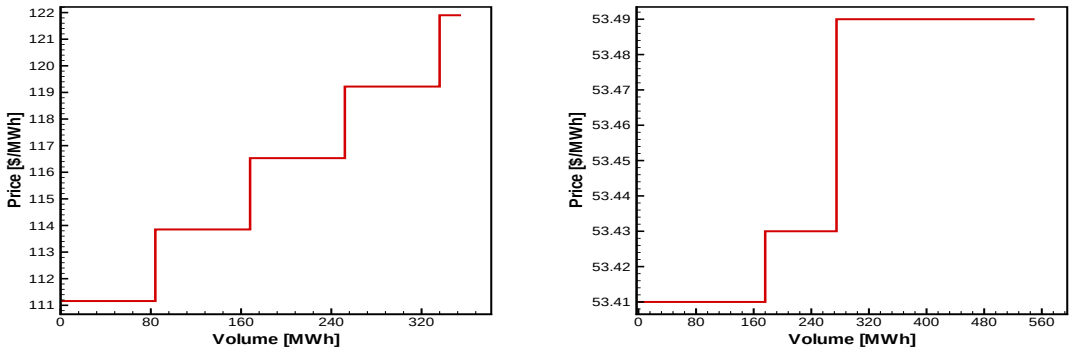
The rest of the paper is organized as follows. Section 2 briefly explains the price formation mechanism in deregulated electricity markets and investigate sensitivity of electricity prices to fuel prices in parametric structural modeling approaches. Section 3 introduces the proposed non-parametric structural spot price modeling framework. The parameter estimation and model performance are discussed in Section 4. Section 5 presents a computational method to derive the corresponding forward prices and demonstrates the inferred time-dependent risk premiums from the market data. Finally, we conclude in Section 6, where some directions for future work are also addressed.

2 Price Formation Mechanism and Sensitivity Estimation

In this section, we briefly explain how electricity prices are formed in deregulated competitive markets and address some disadvantages of using overly simplified structural models which then have motivated us to develop our non-parametric structural hybrid modeling framework.

The electricity prices, at which all cash-energy transactions clear, are established in the pool market. Examples of the pool market include New England Power Pool (NEPOOL), New York Intrastate Access Settlement Pool (NYPOOL), and the California Independent System Operator (CAISO). In a typical pool market, the price is determined through an auction mechanism, in which the market operator (MO) is often the auctioneer. In this process, the generators and power marketers submit electricity supply bids to the MO. A bid is a set of pairs $(price,\ volume)$, from which a *bid curve* for a particular power supplier is constructed. The bid curve determines at which price a generator is willing to supply a given volume of electricity. Two examples of day-ahead bid curves from $1-\text{Jan}-2010$ for a natural gas-fired unit and a coal-based unit in the PJM market are illustrated in Figure 1. The PJM daily energy market bid data are publicly available

at the PJM website[2]. Note that in PJM generators are not obliged to bid step function supply curves, but may for example connect pairs of bids with diagonal lines. Furthermore, they may submit additional unit-specific costs and constraints such as start-up costs, no load costs, and maximum or minimum run times. Simultaneously, the load serving entities submit their demand bids to the market operator. The MO collects



(a) Natural Gas-based Unit

(b) Coal-based Unit

Figure 1: Bid curves for a gas-based and a coal-based generation units in the PJM market.

the supply and demand bids from all the generation or load serving units and determines an optimal output of each generator by means a market clearing process.

Nowadays sophisticated unit commitment and economic dispatch optimization problems are employed in the market clearing process in order to handle a wide variety of unit-specific and system-wide constraints (e.g., transmission constraints, min-start constraints, spinning reserve constraints, minimum up/down times, hydro constraints, or fuel constraints), see, e.g., Chapter 5 of (Wood and Wollenberg, 1996). However, such complex optimization problems can lead to high computation times and are thus not tractable for our purposes of simulating spot prices and finding forward curves over long time horizons.

In contrast, one of the simplest market clearing procedures is the *priority-list method*, see, e.g., Section 5.2.1 of (Wood and Wollenberg, 1996). In this method, the MO sorts the supply bids by price to obtain the *supply stack curve*, also called the *system bid stack*. The market clearing price is then defined as the highest price on the system bid stack, at which the total generation matches the total electricity demand. This is the price paid by all buyers (demands) to all suppliers, and thus it is identical for all market agents in the entire MO region. The market clearing price is sometimes referred to as the *pool price* or *system marginal price*, as it is the bid price of the marginal generator. Note that after the market clearing price is set, additional markets with similar procedures will open to handle any system constraints.

In this paper, we approximate the price formation mechanism by constructing the full PJM supply stack, as we shall discuss further in Section 3. To motivate this choice, we first compare briefly with the typical structural modeling approach of simplifying the stack's shape via parametric functions. A broad range of parametric structural models has been introduced in the literature, see, e.g., (Carmona and Coulon, 2013; Weron, 2014), which differ in the number of fundamental relationships they choose to capture or the

[2] `www.pjm.com/markets-and-operations/energy/real-time/historical-bid-data/`

adopted transformation function to capture them. The main component of a structural model is the explicit parametric function to relate the electricity price with the electricity demand, capacity, marginal fuel, or multiple fuels. Exponential functions have been a persistent part of the parametric transformation in many structural models.

In the very beginning proposal for this class of models, Barlow (2002) suggests

$$P_t = \begin{cases} f_\alpha(L_t) & 1+\alpha L_t > \varepsilon_0 \\ \varepsilon_0^{\frac{1}{\alpha}} & 1+\alpha L_t \leq \varepsilon_0 \end{cases}, \text{ where } f_\alpha(L_t) = \begin{cases} (1+\alpha L_t)^{1/\alpha} & \alpha \neq 0 \\ \exp(L_t) & \alpha = 0 \end{cases},$$

where the (zonal) electricity demand L_t is assumed to follow an Ornstein-Uhlenbeck stochastic process. Later on, other authors have proposed writing prices directly as a function of both the electricity demand L_t and total market capacity U_t using an exponential form such as:

$$P_t = g_t\left(L_t, U_t, P_t^{\mathbf{fuels}}\right) \exp\left(f_t(L_t, U_t, P_t^{\mathbf{fuels}})\right) \qquad . \tag{1}$$

Skantze et al. (2000); Cartea and Villaplana (2008); Lyle and Elliott (2009) consider $g_t(L_t, U_t, P_t^{\mathbf{fuels}}) = 1$ and $f_t(L_t, U_t, P_t^{\mathbf{fuels}}) = \alpha L_t + \beta U_t$ with $\alpha > 0$ and $\beta < 0$. Skantze et al. (2004) choose $f_t(L_t, U_t, P_t^{\mathbf{fuels}}) = \alpha + \beta U_t$ with the same g_t as before. Burger et al. (2004) propose $g_t(U_t, L_t, P_t^{\mathbf{fuels}}) = 1$ and $f_t(L_t, U_t, P_t^{\mathbf{fuels}}) = k_t(L_t/U_t) + X_t$, in which k_t is a non-parametric function and X_t is a portfolio of noise terms. Villaplana (2005) estimates $g_t(L_t, U_t, P_t^{\mathbf{fuels}}) = \gamma_1 U_t^{\gamma_2}$ and $f_t(L_t, U_t, P_t^{\mathbf{fuels}}) = \gamma_t L_t$. Pirrong (2012) and Pirrong and Jermakyan (2008) consider $g_t(L_t, U_t, P_t^{\mathbf{fuels}}) = \left(P_t^{\mathbf{fuels}}\right)^\gamma$ and $f_t(L_t, U_t, P_t^{\mathbf{fuels}}) = \alpha L_t^2 + S(t)$, where $S(t)$ is a deterministic seasonal function.

Coulon et al. (2013) make use of the exponential form in (1) with a second exponential for a spike regime, whose probability is linear in the quantile of demand. More precisely, the spot electricity price P_t,

$$P_t = P_t^{\mathbf{G}} \exp\left(\alpha_m + \beta_m L_t + \gamma_m X_t\right), \tag{2}$$

where $m = 1$ with probability $1 - p_s \Phi\left(\frac{\bar{L}_t - \mu_s}{\sigma_s}\right)$ and $m = 2$ with probability $p_s \Phi\left(\frac{\bar{L}_t - \mu_s}{\sigma_s}\right)$. Here, $P_t^{\mathbf{G}}$ denotes the natural gas price at time t, $\Phi(\cdot)$ is the standard Gaussian cumulative distribution function (cdf), μ_s, σ_s, and p_s are positive constants.

The electricity demand L_t includes two components,

$$L_t = S_t^{\mathbf{L}} + \bar{L}_t. \tag{3}$$

The seasonal component $S_t^{\mathbf{L}}$ is estimated using hourly data,

$$S_t^{\mathbf{L}} \overset{\text{def}}{=} a_{1,h_t} + a_{2,h_t}\cos(2\pi t + a_{3,h_t}) + a_{4,h_t}\cos(4\pi t + a_{5,h_t}) + a_{6,h_t} t + a_{7,h_t} 1_{\mathbf{we}}, \tag{4}$$

where, h_t is the hour corresponding to time t.

The de-seasonalized electricity demand process $\bar{L}_t$ follows a zero-mean Ornstein-Uhlenbeck stochastic process with $\kappa_{\mathbf{L}} > 0$:

$$d\bar{L}_t = -\kappa_{\mathbf{L}} \bar{L}_t dt + \eta_{\mathbf{L}} dW_t^{\mathbf{L}}. \tag{5}$$

In equation (2), the additional factor X_t proxies for the effect of capacity outages and grid congestion, and is given by the summation of a seasonal component and a zero-mean Ornstein-Uhlenbeck stochastic process which is correlated to the the Wiener process $W_t^{\mathbf{L}}$ with a constant correlation parameter.

The use of a pre-specified parametric transformation function to relate the electricity prices to the underlying factors such as fuel prices make most structural models tractable, simple, and attractive for many applications. However, it is extremely challenging to accurately represent the solution of the large-scale market clearing optimization procedure simply as an explicit function of a few input parameters. Therefore, a parametric structural price model, which correctly approximates price levels or distributions, may fail to accurately estimate partial derivative of electricity prices to driving factors like fuel prices. In the following, we further investigate such sensitivities.

In a multi-fuel electricity market like PJM, the shape of the supply curve varies with the level of fuel prices, including the natural gas or coal prices. For example, if the price of natural gas is very low or very high, a small change in the gas price is less likely to impact the merit order of fuels, or the marginal generator in the stack. Therefore, in this case, the electricity price will remain unchanged. In contrast, when the fuel price perturbation impacts the merit order, particularly in those circumstances that this change impacts the type of the marginal generator, the change in the natural gas price can significantly impact the electricity price. Thus an electricity price model should suggest the sensitivity of the electricity price to the natural gas price, $\frac{\Delta P_t}{\Delta P_t^{\mathbf{G}}}(\cdot)$, to depend on the level of gas price and to be a non-constant function of the fuel price. For a detailed discussion, the reader is referred to Chapter 6 in (Swindle, 2014).

Consider the structural model (2) with $p_s = 0$. Given a fixed L_t and X_t, independent of the natural gas price, the electricity price model (2) yields

$$\frac{\Delta P_t}{\Delta P_t^{\mathbf{G}}} = \exp\left(\alpha + \beta L_t + \gamma X_t\right). \tag{6}$$

Equation (6) suggests that the sensitivity of the electricity price to the natural gas price is constant with respect to the natural gas price level. In addition, this sensitivity is an exponential function of the electricity load L_t, only in the region under study and not the aggregated load. We investigate this issue when the electricity price model (2) is applied for the entire market operator territory.

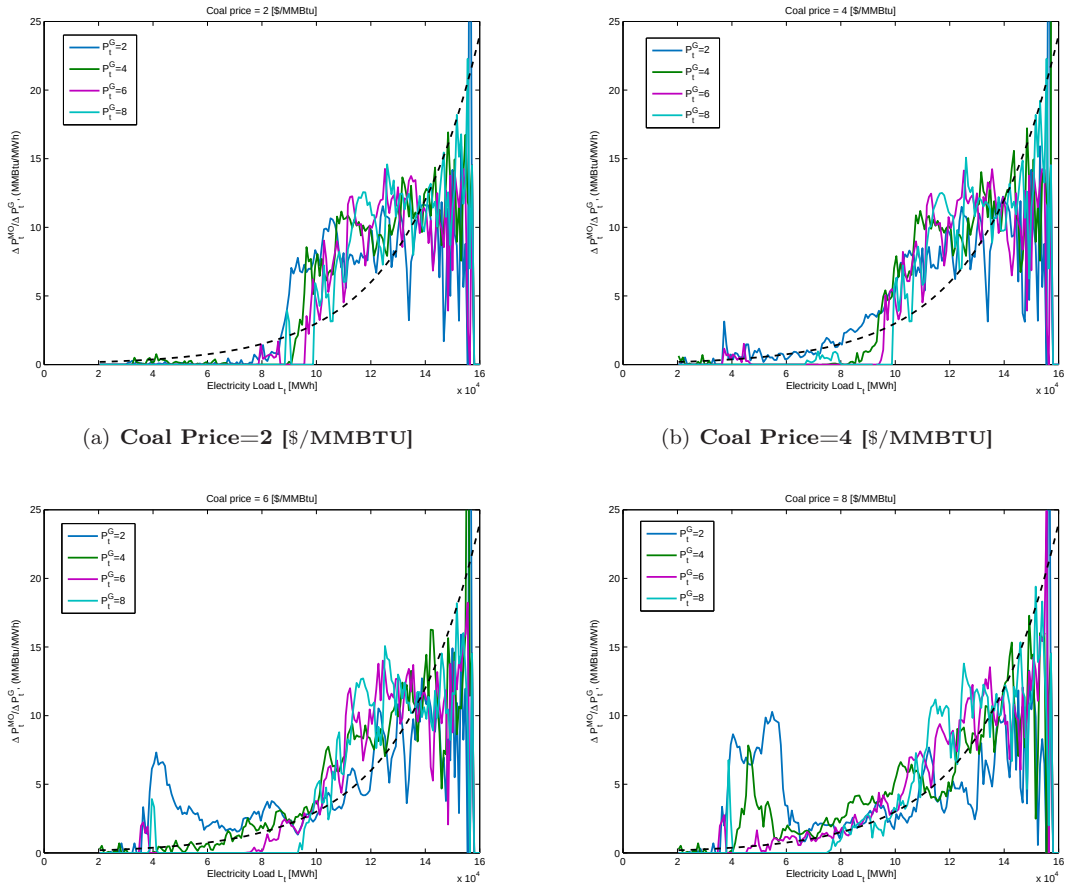
The plots in Figure 2 illustrate the sensitivity of the observed supply curve in PJM to changes in the natural gas price as a function of the aggregated electricity load for $P_t^{\mathbf{G}} = 2, 4, 6, 8$ [\$/MMBTU] when the coal price level is fixed for each plot[3]. We set $\Delta P_t^{\mathbf{G}} = 1$ throughout. As the plots in Figure 2 clearly indicate the sensitivity of the market clearing price to changes in the gas price can significantly deviate from an exponential curve.

At low levels of the electricity load, the marginal generator is most likely a coal-based one. Thus, a change in the gas price is likely to have no impact (see the left end of the curves) on the supply curve and consequently the system marginal price. When the load level is very high, the marginal generator may be a less efficient gas-based unit or quite often an oil-based or diesel-based generator. Therefore, the sensitivity of the energy price to the gas price becomes unstable towards the top of the stack and reduces for high demand (see the right end of the curves). The sensitivity of the price is however consistently large for the medium to high levels of the electricity load at which most possibly a gas-based generator will be the marginal one. This overall pattern is observed for various choices of the coal price (compare different plots in Figure 2). Clearly, such a pattern for the sensitivity of electricity prices to natural gas prices is very hard to predict and cannot be explained by an exponential function of the load as in (6) and the dashed curve in Figure 2.

Furthermore, the pattern of this sensitivity depends on the coal price level. By comparing Figure 2(a) with Figure 2(d), we see that as the coal price increases, the change in the pool price relative to the change

[3]Henry hub natural gas spot price in 2013 (1-Jan-2013 till 31-Dec-2013) had an average of 3.72 [\$/MMBTU] with the minimum 3.08 [\$/MMBTU] and maximum 4.52 [\$/MMBTU]. TETCO-M3 natural gas spot price in 2013 had an average of 3.97 [\$/MMBTU], with the minimum 3.11 [\$/MMBTU] and maximum 11.59 [\$/MMBTU]. Henry hub natural gas spot price from 1-Jan-2008 till 31-Dec-2012 varied between 1.83 [\$/MMBTU] and 13.32 [\$/MMBTU] with an average of 4.77 [\$/MMBTU].

in the gas price increases at the lower levels of the electricity demand; the sensitivity at the medium or high levels of the electricity demand further decreases. This is expected since as the coal price increases, some bids from coal generators move above those from gas generators. As the merit order changes, the probability of each fuel being marginal changes at each load level. Therefore, the sensitivity of the electricity price to the gas price should depend not only on the gas price level but also on the other fuels' prices.



(a) **Coal Price=2 [$/MMBTU]** (b) **Coal Price=4 [$/MMBTU]**

(c) **Coal Price=6 [$/MMBTU]** (d) **Coal Price=8 [$/MMBTU]**

Figure 2: Sensitivity of electricity price P_t to gas price as a function of the electricity load L_t.

A number of multi-fuel structural models have recently been proposed (e.g., Coulon and Howison (2009); Aid et al. (2009, 2013); Carmona et al. (2013)) to better capture complex correlation structures like those discussed above. However, this comes at the expense of either more challenging parameter estimation procedures, or more limited closed-form solutions for forwards. Furthermore, these models require a simplification of the detailed generation cost structure in the market, for example via smooth-shaped clusters of bids for each fuel type in Coulon and Howison (2009), or via an assumption of equal heat rates within fuel types

and a fixed ordering of fuels in Aid et al. (2013). While the more sophisticated structural models cause estimation challenges and require significant modifications for different markets, the simpler structural models (like the one in equation (2)) may not be transferrable from one electricity market to another, without substantial effort and care taken to adapt the model to the key differences and identify dominant risk factors, as discussed in the survey papers (Carmona and Coulon, 2013; Weron, 2014).

On the other hand, as long as sufficient market data is available on generating units, our non-parametric structural hybrid modeling approach based on the generation stack construction can be built for each market without making any compromises on the shape of the stack transformation. This model is described in the subsequent section.

3 The Non-Parametric Structural Hybrid Modeling Approach

We propose to decompose the regional electricity prices into two components and adopt different methods to model each piece:

$$P_t^{\mathbf{zone}} = P_t^{\mathbf{MO}} + \bar{P}_t^{\mathbf{zone}}. \tag{7}$$

The energy price component $P_t^{\mathbf{MO}}$ is computed through a computationally tractable approximation of the economic dispatch optimization scheme, that can be as simple as a priority-list approach. By including an approximation of the market clearing price $P_t^{\mathbf{MO}}$, the fuel effects as well as the total demand and capacity effects on the regional electricity price have been explicitly disentangled. The derivation of this element of the zonal electricity price, as described in the next subsection, does not involve a parameter fitting procedure, avoiding an overreliance on historical price data which may no longer be representative of the market. As the electricity price formation mechanism adopted by the market operator changes, for example due to a change in the regulations or the geography of the region, etc, the optimization scheme to compute $P_t^{\mathbf{MO}}$ can promptly be modified accordingly. The information content in the historical data is not however fully discarded and is taken into account in the estimation of the adjustment component $\bar{P}_t^{\mathbf{zone}}$. This second term captures the effect of congestion, marginal losses, and other sources of noise including any approximation error in the derivation of the first term [4].

3.1 Energy Price

We set up a simple supply-demand model as an approximate economic dispatch optimization problem, to obtain an estimation for the market clearing price. We refer to the resulting price as the energy price. In the subsequent discussion, we assume that all generation units are single-fuel, i.e., no fuel switching is possible, and all of the units are available throughout the year, i.e., no planned or forced unit outages.

Let $\mathcal{I}$ be the set of all power plants in the MO territory. Denote the bid price for power generation unit i at time t by $P_{i,t}^{\mathbf{bid}}$. The domain of this function is $[0,\ V_i^{\max}]$, where $V_i^{\max}$ is the maximum capacity of the

[4] As Monitoring Analytics LLC Report (2005) reports, in calendar year 2004 several geographic areas in the PJM Mid-Atlantic region and the Western region experienced frequent congestion and showed high local market concentration. Particularly, the PJM-PS zone experienced $1,784$ congestion-event hours, the most of any control zone. However, congestion has decreased over years. Day-ahead congestion costs decreased by 43.3% and balancing congestion cost decreased by 5.8% (Monitoring Analytics LLC Report, 2013). It has decreased by 51.4% from the first nine months of 2011 to the first nine months of 2012. In fact, PJM has incentive to minimize the cost for everybody, so the decrease in congestion is going to be observed as well over the coming years. Therefore, we expect that congestion will not be the dominant element of $\bar{P}_t^{\mathbf{zone}}$ in the coming years. This also confirms the importance of focusing on an accurate estimation of $P_t^{\mathbf{MO}}$.

i^{th} power plant. We assume that each bid curve is continuous piecewise constant with at most K pieces, the i^{th} power plant's day-ahead bid curve as a function of the electricity volume V can be represented as:

$$P_{i,t}^{\mathbf{bid}}(V) = \begin{cases} P_{i,t,1}^{\mathbf{bid}} & 0 \leq V \leq V_{i,1} \\ \vdots & \\ P_{i,t,K}^{\mathbf{bid}} & V_{i,K-1} < V \leq V_{i,K} \end{cases} \tag{8}$$

Here, $\{V_{i,1}, \cdots, V_{i,K}\}$ denote the different volume levels (in [MWh]) per day. These volume levels can vary from power plant to power plant, or from day to day for a single power plant. However, in this paper, we let the set of volume levels remain unchanged for all days during the analysis and $K = 10$, for every power plant $i \in \mathcal{I}$.

The balancing act between supply and demand in the price formation leads to mean reversion of prices towards production costs (Carmona and Coulon, 2013). Figures 2 and 8a in (Carmona and Coulon, 2013) indicate that the long term levels of prices tend to match closely with costs of production, as does the long term behavior of the bids (Coulon and Howison, 2009). Therefore, we let for model tractability the bidding decisions of each generator be determined by the cost of power generation for the unit, i.e.,

$$P_{i,t}^{\mathbf{bid}} \approx P_{i,t}^{\mathbf{gen}},$$

where $P_{i,t}^{\mathbf{gen}}$ is the cost of generating 1 [MWh] electrical energy. This assumption excludes the importance of start-up and no-load costs in the bid structure, such that generator costs do not depend on the unit's previous states of operation. The generation cost depends on the efficiency of the power plant and the fuel price, see, e.g., Chapter 7 of (Eydeland and Wolyniec, 2002).

Efficiency of the generation unit is measured by its ability to convert fuel energy content, expressed in [BTU], into electrical energy, measured in [MWh]. This measure is referred to as the *heat rate*. Following the equation (7.6) in (Eydeland and Wolyniec, 2002), for every power plant $i \in \mathcal{I}$ and for every time t, the production cost $P_{i,t}^{\mathbf{gen}}$ can be written as

$$P_{i,t}^{\mathbf{gen}} = H_i P_{i,t}^{\mathbf{fuel}} + C_i^{\mathbf{gen}}, \tag{9}$$

where H_i is the heat rate, $P_{i,t}^{\mathbf{fuel}}$ is the spot fuel price in [\$/MMBTU], and $C_i^{\mathbf{gen}}$, in [\$/MWh], proxies other variable costs of generation such as emission costs or other operational costs. The heat rate H_i and the cost $C_i^{\mathbf{gen}}$ are usually nonconstant and vary with the generation level or other factors such as ambient temperature.

For the generation level V, the heat rate can be expressed by

$$H_i \stackrel{\text{def}}{=} \frac{a_i}{V} + b_i + d_i V, \tag{10}$$

where the coefficients a_i, b_i, d_i are heat rate parameters associated with the i^{th} power plant. Each power plant has its own particular heat rate coefficients, monitored and calculated by *Continuous Emission Monitoring System* (CEMS); the heat rate parameters associated with each power plant can be obtained from the Ventyx database (Ventyx Inc., 2012).

Denote the heat rate and cost at the power level $V_{i,k}$ by $H_{i,k}$ and $C_{i,k}^{\mathbf{gen}}$, respectively. Thus, using equation (9), we have

$$P_{i,t,k}^{\mathbf{gen}} = H_{i,k} P_{i,t}^{\mathbf{fuel}} + C_{i,k}^{\mathbf{gen}}. \tag{11}$$

The marginal increment of production cost due to the change in the fuel price can be approximated by:

$$\Delta P^{\mathbf{gen}}_{i,t,k} \approx H_{i,k} \Delta P^{\mathbf{fuel}}_{i,t}. \tag{12}$$

Therefore, given a bid curve $P^{\mathbf{bid}}_{i,0}$ for the i^{th} generation unit when the fuel price equals $P^{\mathbf{fuel}}_{i,0}$, the new generation curve $P^{\mathbf{gen}}_{i,t}$ for this power plant due to the new fuel prices $P^{\mathbf{fuel}}_{i,t}$ can be computed by:

$$P^{\mathbf{gen}}_{i,t,k} \approx P^{\mathbf{bid}}_{i,0,k} + H_{i,k}\left(P^{\mathbf{fuel}}_{i,t} - P^{\mathbf{fuel}}_{i,0}\right) = H_{i,k} P^{\mathbf{fuel}}_{i,t} + \left(P^{\mathbf{bid}}_{i,0,k} - H_{i,k} P^{\mathbf{fuel}}_{i,0}\right). \tag{13}$$

Hence, $(P^{\mathbf{bid}}_{i,0,k} - H_{i,k} P^{\mathbf{fuel}}_{i,0})$ can be an approximation for $C^{\mathbf{gen}}_{i,k}$.

At every time t, once the production costs $P^{\mathbf{gen}}_{i,t}$ and consequently approximations for the generation units' bid curves are determined, we can construct the generation stack function by sorting the units according to their costs. Therefore, the equilibrium price, denoted by $P^{\mathbf{MO}}_t$, corresponding to the total electricity demand $\tilde{L}^{\mathbf{MO}}_t$ equals

$$P^{\mathbf{MO}}_t \stackrel{\text{def}}{=} \max_{i,k \quad \text{s.t.} \quad x^*_{i,k}>0} \tilde{P}^{\mathbf{gen}}_{i,t,k}. \tag{14}$$

where $\{x^*_{i,k}\}_{i\in I, k=1,\cdots,K}$ is a solution of the following optimization problem:

$$\begin{aligned} \min_{x_{i,k}} \quad & \sum_{i\in\mathcal{I}} \sum_{k=1}^{K} \tilde{P}^{\mathbf{gen}}_{i,t,k}\, x_{i,k} \\ \text{subject to} \quad & \sum_{i\in\mathcal{I}} \sum_{k=1}^{K} x_{i,k} \geq \tilde{L}^{\mathbf{MO}}_t, \\ & 0 \leq x_{i,k} \leq V_{i,k}, \quad \text{for all } k = 1,\cdots,K, \text{ and } i \in \mathcal{I}. \end{aligned} \tag{15}$$

The price $P^{\mathbf{MO}}_t$ is the optimal dual variable of the constraint $\sum_{i\in\mathcal{I}} \sum_{k=1}^{K} x_{i,k} \geq \tilde{L}^{\mathbf{MO}}_t$. Note that the dual problem is as below:

$$\begin{aligned} \max_{P,\ \lambda_{i,k}} \quad & \tilde{L}^{\mathbf{MO}}_t P - \sum_{i\in\mathcal{I}} \sum_{k=1}^{K} V_{i,k} \lambda_{i,k}, \\ \text{subject to} \quad & P - \lambda_{i,k} \leq \tilde{P}^{\mathbf{gen}}_{i,k,t}, \quad i \in \mathcal{I},\ k = 1,\cdots,K \\ & P \geq 0, \quad \lambda_{i,k} \geq 0, \quad i \in \mathcal{I},\ k = 1,\cdots,K. \end{aligned} \tag{16}$$

The optimization problem (16) can get more sophisticated and closer to the economic dispatch optimization scheme by imposing other constraints. For example, box constraints such as $P^{\mathbf{MO}}_t \leq P^{\max}$ for a given constant $P^{\max}$ can be included. Notice that, when each LMP at each bus is restricted to be less than $P^{\max}$ by the market operator, the weighted average of the LMPs and consequently the system price should also satisfy this bound constraint. In our simulation in Section 5, we let $P^{\max} = 1000$ [\$/MWh], which is equal to the current PJM's price cap, although it may lift[5].

In addition to addressing changes in the maximum price cap, since the derivation of the price $P^{\mathbf{MO}}_t$ relies on the total electricity demand $\tilde{L}^{\mathbf{MO}}_t$ and the capacity of all generation units in the MO region, $P^{\mathbf{MO}}_t$ is

[5]see `http://www.rtoinsider.com/pjm-lifts-1k-cap-0114/`

able to adapt to various changes in the entire MO territory. Examples of these events include addition or reduction in electricity generating sources or demands in other zones, updating the generation retirement decisions or maintenance schedules, modifying the objective function for instance to an unequally-weighted social welfare function, or revising the bidding policy particularly for storage devices.

In this paper, we use a subset of the PJM generation units in our analysis [6]. We consider $1,014$ power plants in PJM, 304 of them are gas-based (natural gas, landfill gas, other gas, and Kerosene), and 303 of them are running on coal (bituminous coal, subbituminous coal, waste coal), and the rest of them (407 units) contain oil-based (distillate fuel oil or residual fuel oil) units or those generators for which we do not have precise information.

Figure 3 depicts the daily average real time electricity prices for the PJM-PS zone versus the daily Henry hub natural gas and NYMEX QX (QXc1) coal prices, multiplied by a factor of 10 for comparison purposes. The historical electricity price path in Figure 3 shows high correlations of the PJM-PS electricity prices with both the natural gas and coal prices. Therefore, in our analyses, we only adjust the generation costs of the gas-fired and coal-fired generators with the changes in the natural gas and coal prices. For those generation units which are not running on natural gas or coal, we assume that their production costs and thus their bids remain unchanged over time. The approach, however, can easily accommodate sensitivity to other fuels' price changes.

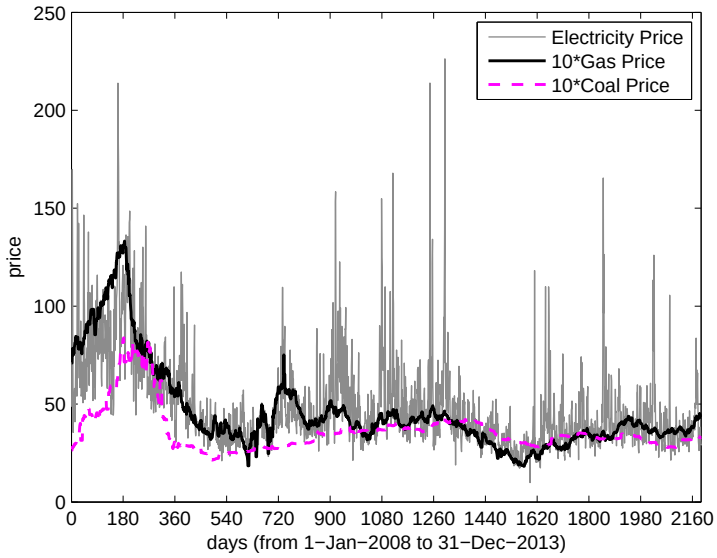
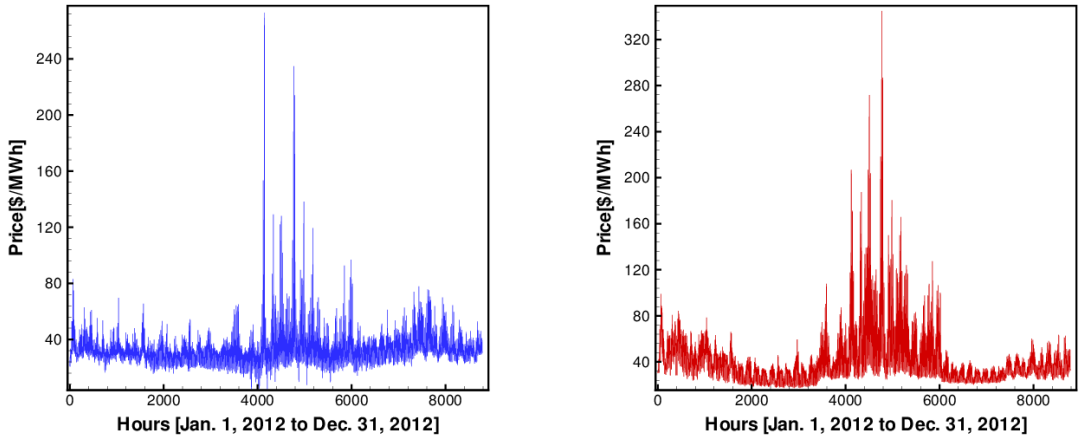


Figure 3: Daily average real time electricity prices in [\$/MWh] for PJM-PS zone as well as daily gas and coal prices in [\$/MMBTU] from $1-\text{Jan}-2008$ to $31-\text{Dec}-2013$.

We use the day-ahead bid curves from $1-\text{Jan}-2010$ as $\{P_{i,0}^{\mathbf{bid}}\}_{i\in\mathcal{I}}$ in equation (13) to approximate the costs $C_{i,k}^{\text{gen}}$ and consequently to construct the (day-ahead) generation curves from $1-\text{Jan}-2012$ to

[6] According to the report dated 3-Mar-2014, PJM has $1,375$ generating sources, see `http://www.pjm.com/~/media/about-pjm/newsroom/fact-sheets/pjm-statistics.ashx`

31 – Dec – 2012. For $P_{i,t}^{\mathbf{gas}}$ and $P_{i,t}^{\mathbf{coal}}$, we respectively apply 2012 daily historical data from Henry hub natural gas prices and the continuous front month for NYMEX QX (`QXc1`) coal prices. For $\tilde{L}_t^{\mathbf{MO}}$, we use hourly historical demands for the PJM market in 2012 to compute $P_t^{\mathbf{MO}}$ for every hour. The right plot in Figure 4 illustrates the computed price $P_t^{\mathbf{MO}}$. The left plot in Figure 4 depicts the historical system marginal prices for the same time window obtained from the PJM website [7]. A comparison between the two graphs clearly shows that the computed prices $P_t^{\mathbf{MO}}$ can capture the dominant behavior in the historical system marginal prices.



(a) Historical PJM System Price

(b) Computed $P_t^{\mathbf{MO}}$

Figure 4: Historical PJM system prices and computed energy prices $P_t^{\mathbf{MO}}$ for 2012.

The derivation of the pool price may be further improved by taking into account the variation in the generation costs of all (including non-coal/gas based) power plants, considering start up costs and maintenance schedules, and updating the total generating resources in PJM over time, as new generation units are installed and some units are deactivated. In addition, the natural gas or coal prices may vary given different points in the PJM area, for example due to the congestion in the natural gas pipeline network. Nevertheless, here, we assumed that the natural gas price is identical for all generating units across the PJM territory and equals the Henry Hub gas price. Our modeling approach, however, is capable to employ distinct regional natural gas prices $P_{i,t}^{\mathbf{fuel}}$ for different units i.

The price $P_t^{\mathbf{MO}}$ constitutes the fundamental component in our electricity price modeling approach. However, this term by itself cannot fully reproduce the observed stochastic behavior of regional electricity prices.

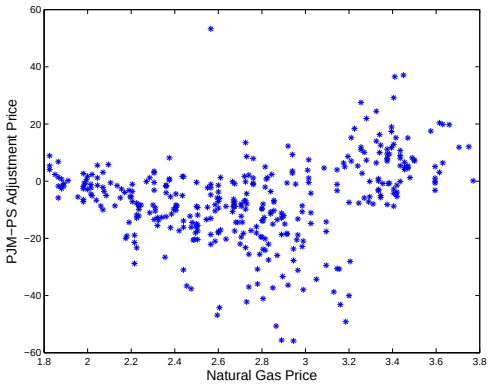
3.2 Adjustment Term

The adjustment term can be modeled either by a parametric statistical model, fitted to the historical residual prices, or through a nonparametric model. In this work, we choose to follow a nonparametric approach for

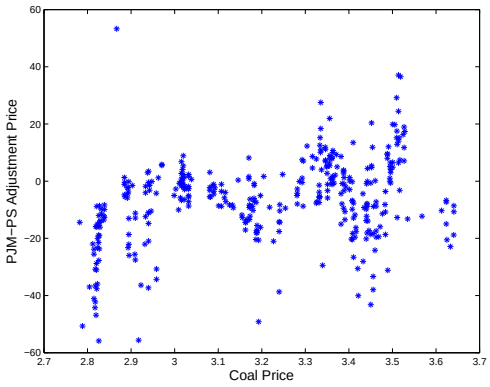
[7] `http://www.pjm.com/markets-and-operations/energy/day-ahead/lmpda.aspx`

addressing $\bar{P}_t^{\mathbf{zone}}$.

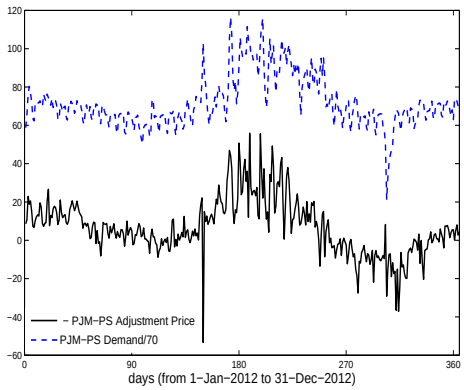
To gain some insight into the main drivers of the adjustment term, we first examine the difference between the historical PJM-PS zonal electricity prices and computed $P_t^{\mathbf{MO}}$ for the PJM market, as described in the previous subsection. The plots in Figure 5(a) and Figure 5(b) illustrate the daily average of the historical $\bar{P}_t^{\mathbf{zone}}$ from $1-\text{Jan}-2012$ to $31-\text{Dec}-2012$ in terms of the daily gas price and coal price in 2012. These plots do not suggest a strong correlation between $\bar{P}_t^{\mathbf{zone}}$ and the gas price or coal price. Therefore, it seems that the dominant part of the impact of the fuels prices on the zonal electricity prices has already been captured by the energy price component $P_t^{\mathbf{MO}}$. The plots in Figure 5(c), in contrast, suggest a significant correlation between $\bar{P}_t^{\mathbf{zone}}$ and the PJM-PS zonal electricity demand over the calendar year 2012.

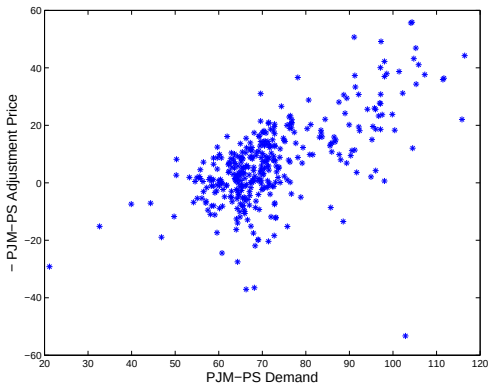


(a) $\bar{P}_t^{\mathbf{zone}}$ **and Natural Gas Price**



(b) $\bar{P}_t^{\mathbf{zone}}$ **and Coal Price**





(c) $\bar{P}_t^{\mathbf{zone}}$ **and PJM-PS Electricity Demands**

Figure 5: Historical adjustment prices $\bar{P}_t^{\mathbf{zone}}$ for PJM-PS zone versus the fuel prices and zonal electricity demands in 2012.

To further analyze the impact of PJM-PS zonal electricity demand on $\bar{P}_t^{\mathbf{zone}}$, we analyze the deviation

of $\bar{P}_t^{\mathbf{zone}}$ from its expected value as a function of $L_t^{\mathbf{zone}} - \mathbb{E}[L_t^{\mathbf{zone}}]$ for the PJM-PS zone. We partition the data points for $L_t^{\mathbf{zone}} - \mathbb{E}[L_t^{\mathbf{zone}}]$ in 2012 into five and ten buckets, and compute the standard deviation of the corresponding $\bar{P}_t^{\mathbf{zone}}$ for each bucket. For five buckets, the break points are chosen to be $L^{(1)} = -2000$, $L^{(2)} = 0$, $L^{(3)} = 2000$, and $L^{(4)} = 4000$ [MWh], and the resulting buckets contain 86, 4887, 3152, 594, 65 data points, respectively. For ten buckets, the break points include $L^{(1)} = -3000$, $L^{(2)} = -2000$, $L^{(3)} = -1000$, $L^{(4)} = 0$, $L^{(5)} = 1000$, $L^{(6)} = 2000$, $L^{(7)} = 3000$, $L^{(8)} = 4000$, and $L^{(9)} = 5000$ [MWh], which results in 32, 54, 1556, 3331, 2524, 628, 393, 201, 61, 4 data points. Figure 6 depicts the levels of standard deviation for the $\bar{P}_t^{\mathbf{zone}}$ data points in each bucket in each case. The plots in Figure 6 indicate that the volatility in $\bar{P}_t^{\mathbf{zone}}$ increases, as $L_t^{\mathbf{zone}}$ farther deviates from its expected value, i.e., $|L_t^{\mathbf{zone}} - \mathbb{E}[L_t^{\mathbf{zone}}]|$ deviates from zero. As the plots illustrate $\bar{P}_t^{\mathbf{zone}}$ has almost the same standard deviation in the two intervals $-2000 \leq L_t^{\mathbf{zone}} \leq -1000$ and $-1000 \leq L_t^{\mathbf{zone}} \leq 0$. Such an analysis can assist the model user to determine a right number of buckets for modeling $\bar{P}_t^{\mathbf{zone}}$.

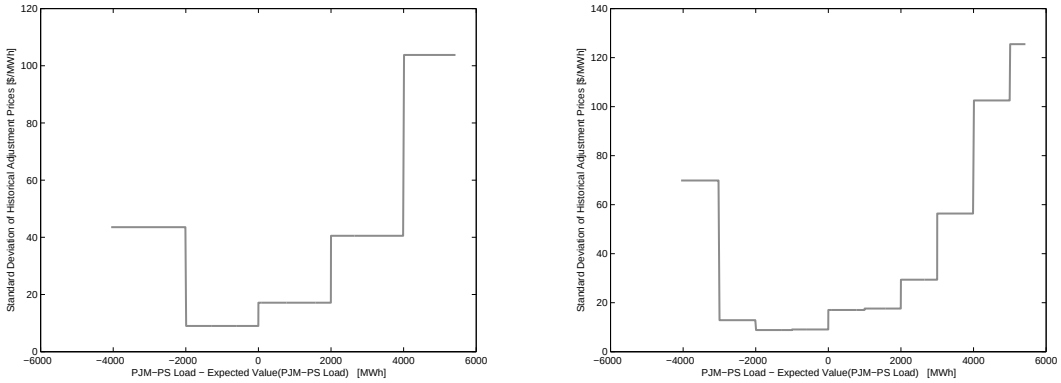


Figure 6: Standard deviation of partitioned 2012 historical $\bar{P}_t^{\mathbf{zone}}$ in terms of $L_t^{\mathbf{zone}} - \mathbb{E}[L_t{}^{\mathbf{zone}}]$.

The above observation motivates us to consider a piecewise probability distribution function for modeling $\bar{P}_t^{\mathbf{zone}}$. We partition the regional electricity demand into M buckets and let the cumulative distribution function for $\bar{P}_t^{\mathbf{zone}}$ be,

$$F_{\bar{P}_t^{\mathbf{zone}}}(x) \stackrel{\text{def}}{=} \begin{cases} F^{(1)}_{\bar{P}_t^{\mathbf{zone}}}(x) & \text{if } L_t^{\mathbf{zone}} < L^{(1)} \\ F^{(2)}_{\bar{P}_t^{\mathbf{zone}}}(x) & \text{if } L^{(1)} \leq L_t^{\mathbf{zone}} < L^{(2)} \\ \vdots & \\ F^{(M-1)}_{\bar{P}_t^{\mathbf{zone}}}(x) & \text{if } L^{(M-2)} \leq L_t^{\mathbf{zone}} < L^{(M-1)} \\ F^{(M)}_{\bar{P}_t^{\mathbf{zone}}}(x) & \text{if } L^{(M-1)} \leq L_t^{\mathbf{zone}} \end{cases} \tag{17}$$

In this work, we let $F^{(i)}_{\bar{P}_t^{\mathbf{zone}}}$ be the empirical distribution function associated with the data points in the i^{th} bucket $\mathcal{L}_i \stackrel{\text{def}}{=} [L^{(i-1)},\ L^{(i)}]$. For notational convenience, we set $L_0 = -\infty$ and $L_M = +\infty$. To generate a sample $\bar{P}_t^{\mathbf{zone}}$, we use the inversion method, see, e.g., Theorem 2.8 in (Seydel, 2006). In this method, we first forecast $L_t^{\mathbf{zone}}$ and generate a uniform random variable $\tilde{u} \sim \mathcal{U}[0,1]$. Then $\bar{P}_t^{\mathbf{zone}} = \hat{F}_i^{-1}(\tilde{u})$, where i is

the load index where $L^{(i-1)} \leq L_t^{\mathbf{zone}} < L^{(i)}$. Since the mean of the empirical distribution equals the sample mean, we have,

$$\mathbb{E}\left[\bar{P}_t^{\mathbf{zone}}\right] = \sum_{i=1}^{M} \left(\frac{1}{|\mathcal{L}_i|} \sum_{\ell \in \mathcal{L}_i} \bar{P}_{t,\ell}^{\mathbf{zone}} \right) \Pr\left(L^{(i-1)} \leq L_t^{\mathbf{zone}} < L^{(i)} \right), \tag{18}$$

where $|\mathcal{L}_i|$ indicates the number of points in our historical data set $\mathcal{L}_i$, and $\{\bar{P}_{t,\ell}^{\mathbf{zone}}\}_{\ell \in \mathcal{L}_i}$ are the corresponding adjustment prices in this bucket.

4 Model Implementation and Performance

We implement our structural hybrid model for the PJM-PS zone, when the energy prices are computed as in subsection 3.1 and a piecewise empirical distribution as in (17) is applied for $\bar{P}_t^{\mathbf{zone}}$. We consider 7 buckets in our analysis, $L^{(1)} = -3000$, $L^{(2)} = -2000$, $L^{(3)} = 0$, $L^{(4)} = 2000$, $L^{(5)} = 3000$, $L^{(6)} = 4000$, to avoid dealing with too small number of data points in a bucket.

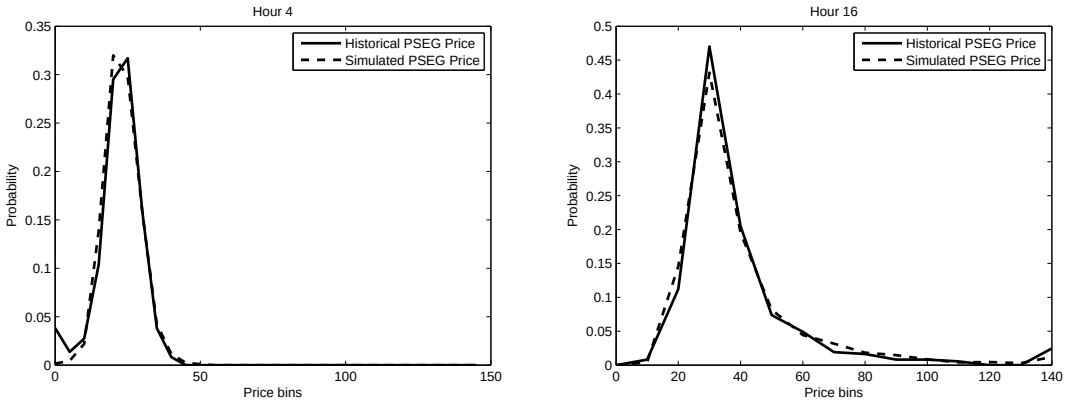
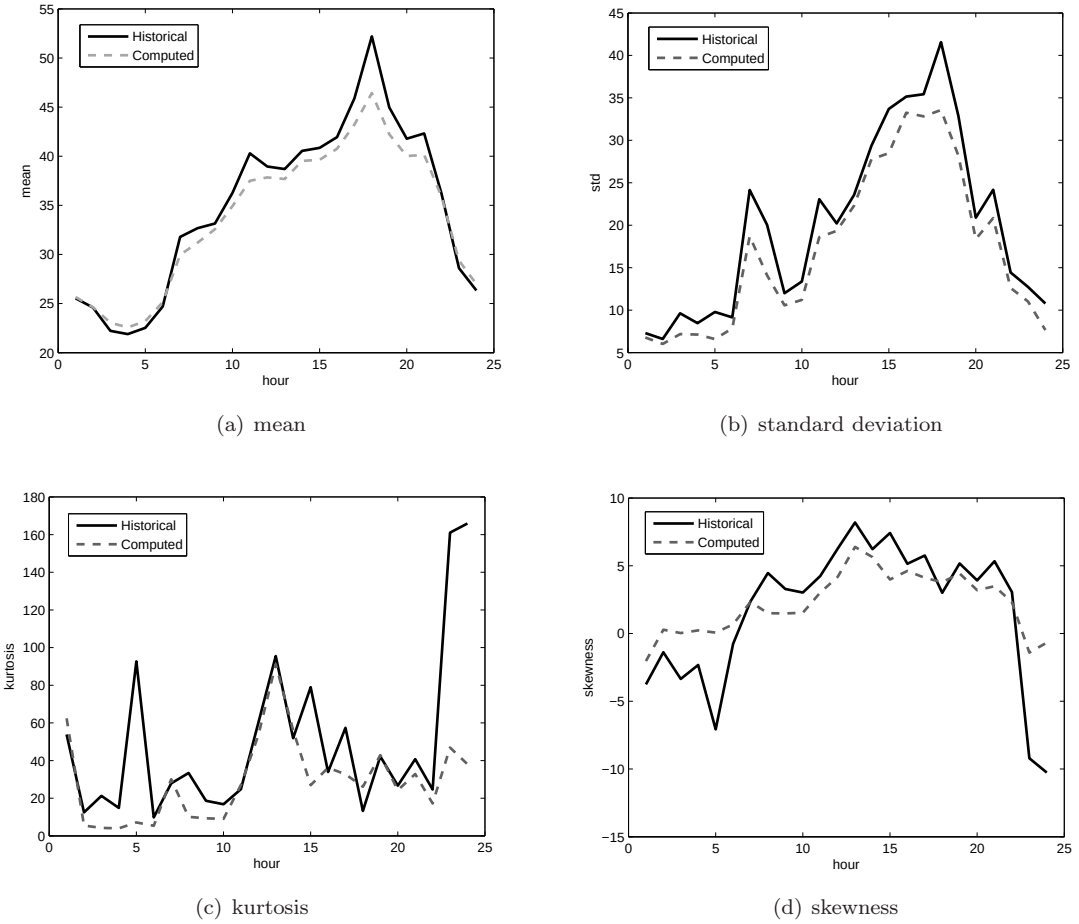


Figure 7: Histograms for P_t: model simulation vs. historical data for PJM-PS Zone in 2012 (left: Hour 4 AM; right: Hour 4 PM).

Figure 7 illustrates the histograms of the historical zonal electricity prices in 2012 and the simulated prices from the proposed structural hybrid model. For illustration purposes, we plot one representative peak hour (4pm) and one representative off-peak hour (4am), which can be observed to differ significantly. These graphs provide some evidence of the ability of the model to capture hourly price distributions accurately. These plots were generated by simulating 1000 paths of one year of electricity price and load dynamics, while always using the historical natural gas and coal prices for the calendar year 2012. The graphs in Figure 7 indicate that the proposed structural hybrid model is successful at explaining the price distributions. We capture well many of the important characteristics of electricity prices, including most notably the moments of prices, as presented in Figure 8.

To assess the performance of out-of-sample forecasts, using the 2012 empirical distribution for $\bar{P}_t^{\mathbf{MO}}$, hourly PJM-PS electricity demands in 2013, and the daily natural gas and coal prices in 2013, we estimate

the PJM-PS zonal electricity prices in 2013. The histograms for model simulated P_t and historical PJM-PS electricity prices for 2013 are illustrated in Figure 9. These plots show that for some hours the proposed model is not able to capture the hourly price distributions accurately. It might be due to the fact that the adjustment term $\bar{P}_t^{\mathbf{MO}}$ has a non-stationary distribution or due to changes in the generation stack over time.



(a) mean

(b) standard deviation

(c) kurtosis

(d) skewness

Figure 8: Descriptive statistics of historical PJM-PS electricity spot prices and 100 simulated prices for 2012.

5 Electricity Forward Contract Prices

To manage the risk associated with the inherent volatility of the spot market, market participants frequently enter into forward or futures contracts, agreeing at time t to buy or sell an asset at a fixed time $T > t$ and a fixed forward price $F(t, T)$. Forward prices for financial assets can be directly linked to today's spot prices

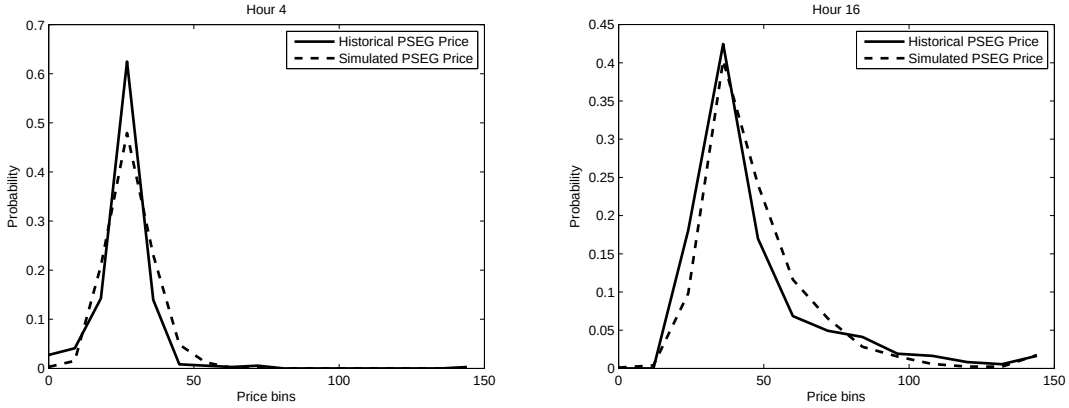


Figure 9: Histograms for P_t: model simulation vs. historical data for PJM-PS Zone in 2013.

by the following simple no arbitrage relationship, see, e.g., Chapter 16 of (McDonald, 2013)

$$F(t,T) = P_t \exp\left(-(\delta - r)(T-t)\right),$$

where P_t is the asset spot price, r is the constant continuously compounded interest rate, and δ is the constant continuous dividend yield on the asset.

However, commodities differ from financial assets in many important aspects, as a result of considerations that do not arise with financial assets such as storage and delivery costs and the so-called 'convenience yield' benefit of holding physical commodities. Therefore, in contrast to the forward price for a stock which is largely redundant, commodity forward prices provide price discovery, revealing otherwise unobtainable information about the (expected) future price of the commodity and its unique economic characteristics or underlying driving factors.The forward price $F(t,T)$ for a commodity is the price determined today (at time t) to receive or deliver one unit of the commodity on the future date T. Simple no arbitrage relationships can no longer be used to directly link $F(t,T)$ to P_t in a model-free manner, but a fair forward price can be expressed as a risk-neutral expectation of the future spot price P_T. In addition, we may define a 'risk premium' $m(t,T)$ in order to relate the time-T commodity forward price with the true expected spot price $\mathbb{E}_t(P_T)$ (conditional on time t information), e.g.,

$$F(t,T) = \mathbb{E}_t(P_T) \exp\left(-m(t,T)\right), \tag{19}$$

Note that variations of this formula are equally acceptable, for example with the risk premium defined additively instead of multiplicatively, or with the exponential function omitted. However, the above definition allows for a natural analogy with the equity markets where $m(t,T) = (\mu - r)(T-t)$, with μ representing the constant drift or expected rate of return of the stock. For stocks and other financial assets, m_T is typically assumed positive due to risk aversion of investors, hence the name 'risk premium'. On the other hand, in commodity markets, m_T may be positive of negative, is linked to hedging pressures from producers, consumers and other market participants, and has been shown in various studies to potentially change sign either through time or across maturities, see, e.g., Longstaff and Wang (2004); Pirrong and Jermakyan (2008); Veraart and Veraart (2013).

Given the observed historical forward price $F(t,T)$ for the maturity T, risk premiums in our proposed model can therefore easily be computed in terms of the forecasted spot price $\mathbb{E}_t(P_T)$:

$$m(t,T) = -\log\left(\frac{F(t,T)}{\mathbb{E}_t(P_T)}\right) = -\log\left(\frac{F(t,T)}{\mathbb{E}_t\left(P_T^{\mathbf{MO}}\right) + \mathbb{E}_t\left(\bar{P}_T^{\mathbf{zone}}\right)}\right). \tag{20}$$

In this equation, the term $\mathbb{E}_t\left(\bar{P}_T^{\mathbf{zone}}\right)$ can be computed from the equation (18). The expected value of $P_T^{\mathbf{MO}}$ is, by definition

$$\mathbb{E}_t(P_T^{\mathbf{MO}}) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} P_T^{\mathbf{MO}}\left(P_T^{\mathbf{G}}, P_T^{\mathbf{C}}, L_T^{\mathbf{MO}}\right)\; f^{\mathbf{GC}}\left(P_T^{\mathbf{G}}, P_T^{\mathbf{C}}\right)\; f^{\mathbf{L}}\left(L_T^{\mathbf{MO}}\right)\; dP_T^{\mathbf{G}}\; dP_T^{\mathbf{C}}\; dL_T^{\mathbf{MO}}, \tag{21}$$

where $f^{\mathbf{GC}}$ is the joint probability density function of the fuel prices at time T, $(P_T^{\mathbf{G}}, P_T^{\mathbf{C}})$, and $f^{\mathbf{L}}$ is the probability density function of the total electricity load at time T, $L_T^{\mathbf{MO}}$. Here, we have assumed that the electricity demand distribution is independent of the joint distribution of the natural gas and coal prices. In the rest of this section, we assume that $P_T^{\mathbf{G}} \in [a^{\mathbf{G}}, b^{\mathbf{G}}]$, $P_T^{\mathbf{C}} \in [a^{\mathbf{C}}, b^{\mathbf{C}}]$, and $L_T^{\mathbf{MO}} \in [a^{\mathbf{L}}, b^{\mathbf{L}}]$. These bounds can be estimated by the maximum and minimum prices and loads in the historical data set considered.

The integrals in (21) can be approximated using the quadrature rule, see, e.g., (Golub and Meurant, 2010; Brass and Petras, 2011):

$$\mathbb{E}_t(P_T^{\mathbf{MO}}) \approx \sum_{i=1}^{n^{\mathbf{G}}}\sum_{j=1}^{n^{\mathbf{C}}}\sum_{k=1}^{n^{\mathbf{L}}} P_{T,i,j,k}^{\mathbf{MO}}\; f^{\mathbf{GC}}\left(P_{T,i}^{\mathbf{G}}, P_{T,j}^{\mathbf{C}}\right)\; f^{\mathbf{L}}\left(L_{T,k}^{\mathbf{MO}}\right)\; h^{\mathbf{G}}h^{\mathbf{C}}h^{\mathbf{L}}, \tag{22}$$

where $P_{T,i,j,k}^{\mathbf{MO}}$ refers to the energy price computed using $P_{T,i}^{\mathbf{G}}$ for the natural gas price, $P_{T,j}^{\mathbf{C}}$ for the coal price, and $L_{T,k}^{\mathbf{MO}}$ for the total electricity demand. Here, given the number of discretization points $n^{\mathbf{G}}$, $n^{\mathbf{C}}$, $n^{\mathbf{L}}$ for the natural gas, coal, and load, respectively, we have,

$$h^{\mathbf{G}} \stackrel{\text{def}}{=} \frac{b^{\mathbf{G}} - a^{\mathbf{G}}}{n^{\mathbf{G}}}, \qquad h^{\mathbf{C}} \stackrel{\text{def}}{=} \frac{b^{\mathbf{C}} - a^{\mathbf{C}}}{n^{\mathbf{C}}}, \qquad h^{\mathbf{L}} \stackrel{\text{def}}{=} \frac{b^{\mathbf{L}} - a^{\mathbf{L}}}{n^{\mathbf{L}}},$$

and for $i = 1, \cdots, n^{\mathbf{G}}$, $j = 1, \cdots, n^{\mathbf{C}}$, and $k = 1, \cdots, n^{\mathbf{L}}$,

$$P_{T,i}^{\mathbf{G}} \stackrel{\text{def}}{=} a^{\mathbf{G}} + \left(i - \frac{1}{2}\right)h^{\mathbf{G}}, \qquad P_{T,j}^{\mathbf{C}} \stackrel{\text{def}}{=} a^{\mathbf{C}} + \left(j - \frac{1}{2}\right)h^{\mathbf{C}}, \qquad L_{T,k}^{\mathbf{MO}} \stackrel{\text{def}}{=} a^{\mathbf{L}} + \left(k - \frac{1}{2}\right)h^{\mathbf{L}}.$$

Let the electricity demand $L_t^{\mathbf{MO}}$ be as in equation (3), in which $\bar{L}_t$ follows the Ornstein-Uhlenbeck stochastic process (5) and the seasonality component $S^{\mathbf{L}}(t)$ is as in (4).

$$f^{\mathbf{L}}\left(L_{T,k}^{\mathbf{MO}}\right) = \frac{1}{\sigma_{\mathbf{L}}(t,T)\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{L_{T,k}^{\mathbf{MO}} - S^{\mathbf{L}}(T) - \mu_{\mathbf{L}}(t,T)}{\sigma_{\mathbf{L}}(t,T)}\right)^2\right),$$

where

$$\mu_{\mathbf{L}}(t,T) = \bar{L}_t \exp\left(-\kappa_{\mathbf{L}}\left(T-t\right)\right), \qquad \sigma_{\mathbf{L}}^2(t,T) = \frac{\eta_{\mathbf{L}}^2}{2\kappa_{\mathbf{L}}}\left(1 - \exp\left(-2\kappa_{\mathbf{L}}(T-t)\right)\right).$$

The risk premium $m(t,T)$ in equation (20) consists of an aggregation of risk premia for each of the underlying stochastic factors in the model, namely load, natural gas and coal prices. However, as forward

contracts are traded in the coal and gas markets, the fuel risk premiums can be inferred from their observed fuel forward prices, providing valuable information when constructing power forward curves. Therefore, by identifying the parameters of the risk-neutral distribution $\hat{f}^{\mathbf{GC}}$ from the market, we can disentangle the multiple risk premia and instead effectively allow $m(t,T)$ to capture only the remaining load-related risk premium. This approach is similar to that of Aid et al. (2013), who describing using a 'local risk minimizing (LRM) strategy' to choose an equivalent martingale measure under which model-implied forward prices are found using risk-neutral probabilities for fuel prices but physical probabilities for load.

Indeed, several authors advocate the use of observed fuel forward prices as inputs when calculating electricity forward prices in structural models, see, e.g., Coulon et al. (2013); Carmona et al. (2013). As power forwards are effectively derivatives on fuel prices, such a procedure is comparable to fitting the yield curve before pricing options in interest rate markets or matching the implied volatility of vanilla equity options before pricing more exotic equity derivatives. Analogously to those cases, we require a flexible enough model for fuels in order to exactly reproduce fuel forward prices, which can be achieved by letting their long term mean levels be time-dependent (piecewise constant[8]) under $\mathbb{Q}$. The following procedure formalizes this procedure, explaining precisely how the joint probability density function $\hat{f}^{\mathbf{GC}}$ under Q is related to the forward prices for the natural gas and coal, determined by the forward market. In the subsequent discussion, we refer to the forward prices with the maturity T at time t for natural gas and coal by $F^{\mathbf{G}}(t,T)$ and $F^{\mathbf{C}}(t,T)$, respectively. Thus, $F^{\mathbf{G}}(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[P_t^{\mathbf{G}}\right]$ and $F^{\mathbf{C}}(t,T) = \mathbb{E}_t^{\mathbb{Q}}\left[P_t^{\mathbf{C}}\right]$, where $\mathbb{E}_t^{\mathbb{Q}}$ denotes time t conditional expectation under the risk-neutral pricing measure $\mathbb{Q}$. Assume that at pricing time t, N data points on the historical natural gas and coal forward curves are given at maturities $T_1, \cdots, T_N$. Note that N is often greater than the number of data points on the historical electricity forward curve.

Proposition 1. *Assume that the gas price $P_t^{\boldsymbol{G}}$ and the coal price $P_t^{\boldsymbol{C}}$ under $\mathbb{Q}$ are the exponentials of Ornstein-Uhlenbeck processes, i.e.,*

$$d\log P_t^{\boldsymbol{G}} = \kappa_{\boldsymbol{G}}\left(\hat{m}^{\boldsymbol{G}}(t) - \log P_t^{\boldsymbol{G}}\right)dt + \eta_{\boldsymbol{G}}\, d\hat{W}_t^{\boldsymbol{G}}, \tag{23}$$

$$d\log P_t^{\boldsymbol{C}} = \kappa_{\boldsymbol{C}}\left(\hat{m}^{\boldsymbol{C}}(t) - \log P_t^{\boldsymbol{C}}\right)dt + \eta_{\boldsymbol{C}}\, d\hat{W}_t^{\boldsymbol{C}}, \tag{24}$$

where the $\mathbb{Q}$-Brownian motions $\hat{W}_t^{\boldsymbol{G}}$ and $\hat{W}_t^{\boldsymbol{C}}$ are correlated with a parameter $\rho_{\boldsymbol{GC}}$, and $\hat{m}^{\boldsymbol{G}}(u)$ and $\hat{m}^{\boldsymbol{C}}(u)$ are piecewise constant functions with one jump per maturity time, i.e.,

$$\hat{m}^{\boldsymbol{G}}(u) \overset{\text{def}}{=} \begin{cases} \hat{m}_1^{\boldsymbol{G}} & t \le u < T_1 \\ \hat{m}_2^{\boldsymbol{G}} & T_1 \le u < T_2 \\ \vdots & \\ \hat{m}_{T_N}^{\boldsymbol{G}} & T_{N-1} \le u \le T_N \end{cases}, \qquad \hat{m}^{\boldsymbol{C}}(u) \overset{\text{def}}{=} \begin{cases} \hat{m}_1^{\boldsymbol{C}} & t \le u < T_1 \\ \hat{m}_2^{\boldsymbol{C}} & T_1 \le u < T_2 \\ \vdots & \\ \hat{m}_{T_N}^{\boldsymbol{C}} & T_{N-1} \le u \le T_N \end{cases} \tag{25}$$

Then, $(P_t^{\boldsymbol{G}}, P_t^{\boldsymbol{C}})$ has a multivariate log-normal distribution (see, e.g., (Kleiber and Kotz, 2003)), i.e.,

$$\hat{f}^{\boldsymbol{GC}}\left(P_{T,i}^{\boldsymbol{G}}, P_{T,j}^{\boldsymbol{C}}\right) = \frac{1}{2\pi P_{T,i}^{\boldsymbol{G}} P_{T,j}^{\boldsymbol{C}}\sqrt{|\hat{\Sigma}(t,T)|}} \exp\left(-\frac{1}{2}\begin{pmatrix} \log P_{T,i}^{\boldsymbol{G}} - \hat{\mu}_{\boldsymbol{G}}(t,T) \\ \log P_{T,j}^{\boldsymbol{C}} - \hat{\mu}_{\boldsymbol{C}}(t,T) \end{pmatrix}^{\top} \hat{\Sigma}^{-1}(t,T) \begin{pmatrix} \log P_{T,i}^{\boldsymbol{G}} - \hat{\mu}_{\boldsymbol{G}}(t,T) \\ \log P_{T,j}^{\boldsymbol{C}} - \hat{\mu}_{\boldsymbol{C}}(t,T) \end{pmatrix}\right),$$

[8] If starting from the model in (23) and (24) under the physical measure, this corresponds to a time-dependent and piecewise constant market price risk for both gas and coal.

where $\hat{\mu}_G(t,T)$, $\hat{\mu}_C(t,T)$, and $\hat{\Sigma}(t,T)$ are as follows,

$$\hat{\mu}_G(t,T) \stackrel{\text{def}}{=} e^{-\kappa_{\boldsymbol{G}}(T-t)} \log P_t^{\boldsymbol{G}} + \sum_{i=1}^{\ell_T-1} \hat{m}_i^{\boldsymbol{G}} \left(e^{\kappa_G(T_i-T)} - e^{\kappa_G(T_{i-1}-T)} \right) + \hat{m}_{\ell_T}^{\boldsymbol{G}} \left(1 - e^{\kappa_G(T_{\ell_T-1}-T)} \right), \tag{26}$$

$$\hat{\mu}_C(t,T) \stackrel{\text{def}}{=} e^{-\kappa_{\boldsymbol{C}}(T-t)} \log P_t^{\boldsymbol{C}} + \sum_{i=1}^{\ell_T-1} \hat{m}_i^{\boldsymbol{C}} \left(e^{\kappa_C(T_i-T)} - e^{\kappa_C(T_{i-1}-T)} \right) + \hat{m}_{\ell_T}^{\boldsymbol{C}} \left(1 - e^{\kappa_{\boldsymbol{C}}(T_{\ell_T-1}-T)} \right), \tag{27}$$

$$\hat{\Sigma}(t,T) \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\eta_G^2}{2\kappa_G} \left(1 - e^{-2\kappa_{\boldsymbol{G}}(T-t)} \right) & \frac{\rho_{\boldsymbol{GC}}\ \eta_{\boldsymbol{G}}\ \eta_{\boldsymbol{C}}}{\kappa_{\boldsymbol{G}}+\kappa_{\boldsymbol{C}}} \left(1 - e^{(-(\kappa_{\boldsymbol{G}}+\kappa_{\boldsymbol{C}})(T-t))} \right) \\ \frac{\rho_{\boldsymbol{GC}}\ \eta_{\boldsymbol{G}}\ \eta_{\boldsymbol{C}}}{\kappa_{\boldsymbol{G}}+\kappa_{\boldsymbol{C}}} \left(1 - e^{(-(\kappa_{\boldsymbol{G}}+\kappa_{\boldsymbol{C}})(T-t))} \right) & \frac{\eta_C^2}{2\kappa_C} \left(1 - e^{-2\kappa_C(T-t)} \right) \end{pmatrix}. \tag{28}$$

with the $\hat{m}^{\boldsymbol{G}}(u)$ and $\hat{m}^{\boldsymbol{C}}(u)$ defined as below (for $s \geq 1$):

$$\hat{m}_s^{\boldsymbol{G}} = \frac{\log F^{\boldsymbol{G}}(t,T_s) - e^{-\kappa_{\boldsymbol{G}}(T_s-t)} \log P_t^{\boldsymbol{G}} - \frac{\eta_G^2}{4\kappa_{\boldsymbol{G}}} \left(1 - e^{-2\kappa_{\boldsymbol{G}}(T_s-t))} \right) - \sum_{i=1}^{s-1} \hat{m}_i^{\boldsymbol{G}} \left(e^{-\kappa_{\boldsymbol{G}}(T_s-T_i)} - e^{-\kappa_{\boldsymbol{G}}(T_s-T_{i-1})} \right)}{\left(1 - e^{-\kappa_{\boldsymbol{G}}(T_s-T_{s-1})} \right)},$$

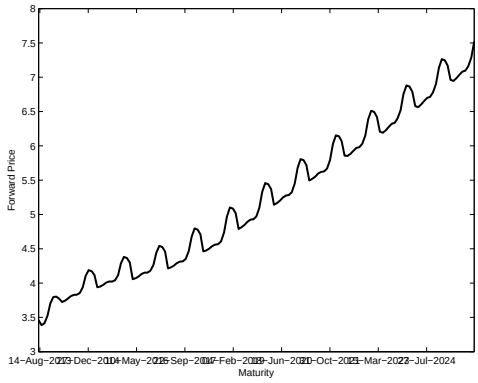
$$\hat{m}_s^{\boldsymbol{C}} = \frac{\log F^{\boldsymbol{C}}(t,T_s) - e^{-\kappa_{\boldsymbol{C}}(T_s-t)} \log P_t^{\boldsymbol{C}} - \frac{\eta_C^2}{4\kappa_{\boldsymbol{C}}} \left(1 - e^{-2\kappa_{\boldsymbol{C}}(T_s-t))} \right) - \sum_{i=1}^{s-1} \hat{m}_i^{\boldsymbol{C}} \left(e^{-\kappa_{\boldsymbol{C}}(T_s-T_i)} - e^{-\kappa_{\boldsymbol{C}}(T_s-T_{i-1})} \right)}{\left(1 - e^{-\kappa_{\boldsymbol{C}}(T_s-T_{s-1})} \right)}.$$

A proof for Proposition 1 is provided in Appendix A.

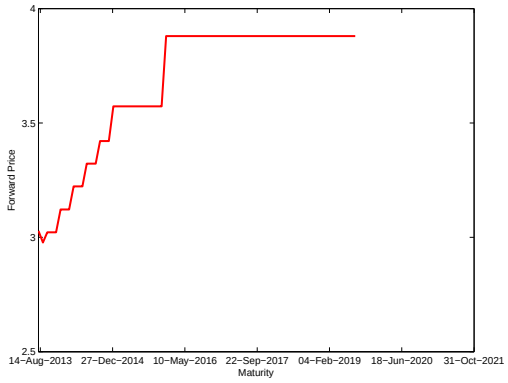
Figures 10(a) and 10(b) illustrate the natural gas forward curve for pricing date 1-August-2013 with maturities ranging from 1-September-2013 to 1-December-2025, and the coal forward curve for pricing date 1-August-2013 with maturities of 1-September-2013 to 1-July-2019. The risk adjusted drifts $\{\hat{m}_T^{\mathbf{G}}\}_{T=1}^{72}$ and $\{\hat{m}_T^{\mathbf{C}}\}_{T=1}^{72}$ found via Proposition 1 are illustrated in Figures 10(c) and 10(d). Here, the derivation use $\kappa_{\mathbf{G}} = 1.2034$, $\kappa_{\mathbf{C}} = 1.0608$, $\eta_{\mathbf{G}} = 0.5964$, $\eta_{\mathbf{C}} = 0.3711$, and $\rho_{\mathbf{GC}} = 0.5983$ estimated by calibrating the exponential Ornstein-Uhlenbeck processes from daily historical data of natural gas prices and coal prices for 2-January-2008 to 31-May-2013.

For every given pricing date, using Proposition 1, we can then use the inferred risk adjusted drifts to compute the joint $\mathbb{Q}$-probability density function $f^{\mathbf{GC}}(P_{T,i}^{\mathbf{G}}, P_{T,j}^{\mathbf{C}})$ and consequently to derive electricity price expectations $\mathbb{E}_t(P_T^{\mathbf{MO}})$ as in equation (22). In this computation, we set $a^{\mathbf{G}} = a^{\mathbf{C}} = 1$ [\$/MMBTU], $b^{\mathbf{G}} = b^{\mathbf{C}} = 10$ [\$/MMBTU], and $n^{\mathbf{G}} = n^{\mathbf{C}} = 30$. We also use $b^{\mathbf{L}} = 180,000$ [MWh], $a^{\mathbf{L}} = 20,000$ [MWh], and $n^{\mathbf{L}} = 100$. To compute $f^{\mathbf{L}}\left(L_{T,k}^{\mathbf{MO}}\right)$, we set the parameter values $\kappa_{\mathbf{L}} = 376.1865$ and $\eta_{\mathbf{L}} = 328,611.0959$, estimated using hourly historical PJM electricity demand for 2-January-2008 to 31-May-2013. The estimated load seasonality parameters in $S_t^{\mathbf{L}}$ are presented in Appendix B. We then calculate $\mathbb{E}[\bar{P}_t^{\mathbf{zone}}]$ via (18) to obtain zonal price expectations, before finally obtaining the risk premiums $m(t,T)$. Figure 11 shows the computed risk premium curves $\{m(t,T)\}_T$ for both on-peak and off-peak power forward prices (from PJM market, PSEG component) corresponding to pricing dates t =1-August-2013 and t =1-August-2014. For on-peak forwards, for simplicity we assume that the delivery period is from 7 am to 11 pm of the first day of the maturity month; hence the averaged expectation $\mathbb{E}_t(P_T^{\mathbf{MO}})$ is used in equation (20) to infer risk premiums for on-peak forwards. Similarly, for off-peak forwards, we assume that the delivery period is from 12 am to 6 am of the first day of the maturity month and the corresponding averages are used.

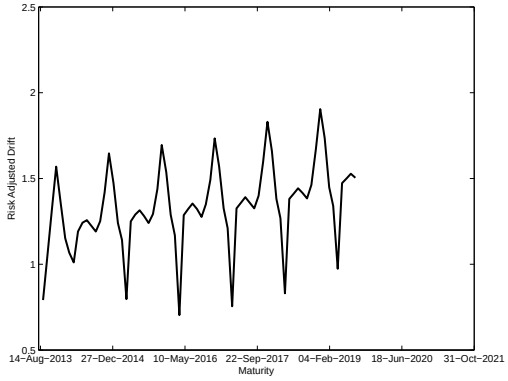
The plots in Figure 11 indicate for both pricing dates, the risk premiums for on-peak hours are generally higher than the risk premiums for off-peak hours. In addition, the risk premiums vary with maturity and



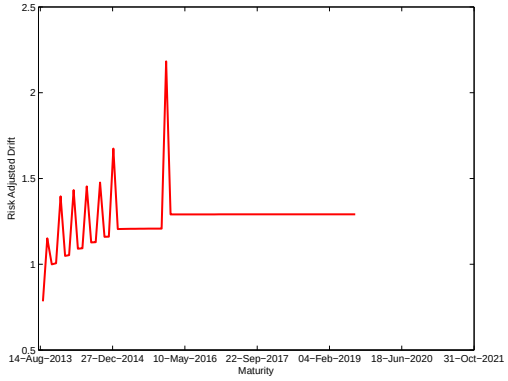
(a) Natural Gas Forward Curve



(b) Coal Forward Curve



(c) Natural Gas Risk Adjusted Drifts

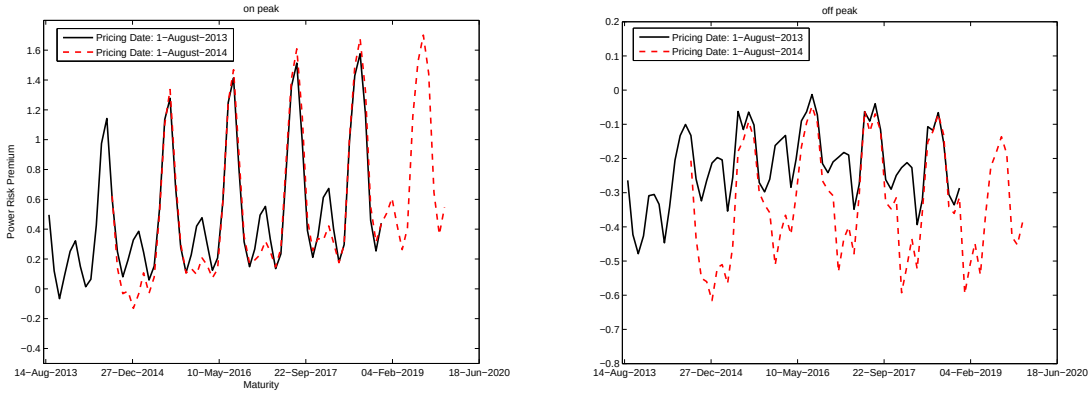


(d) Coal Risk Adjusted Drifts

Figure 10: Natural gas forward prices (from NYMEX market, NG Component) and coal forward prices (from NYMEX market, QX Component) for pricing date 1-August-2013 and the corresponding inferred risk adjusted drifts.

maintain some seasonality pattern. The change in the magnitude of the risk premium by maturity is more significant for on-peak hours. A comparison between the two plots in Figure 11 shows that the pattern varies somewhat between the two chosen pricing dates, but is remarkably stable for summer months. In contrast, the difference between the inferred risk premiums for winter months using the available forward prices on 1-August-2013 and 1-August-2014 is quite prominent, particularly for off-peak contracts. However, this change in the risk premiums is rather gradual from day to day.

For a pricing time t and given an electricity forward curve for time t with N_e maturities and fuel forward curves for pricing time t with $N >> N_e$ data points, the risk premium rates $m_1, \cdots, m_{N_e}$ are first obtained from equation (20). Then to project the electricity forward curve for maturities beyond what is currently



(a) On Peak

(b) Off Peak

Figure 11: Inferred on-peak risk premiums and off-peak risk premiums for electricity forwards for two pricing dates.

available in the market and to leverage long fuel forward curves, a common practitioners' approach is to extrapolate the computed market risk premiums $m_{T_1}, \cdots, m_{T_{N_e}}$, see, e.g., Benth et al. (2008). Being able to discover the pattern of risk premia for different maturities allows us to better estimate the forward curve beyond liquidly tradable maturities, a procedure which can be very valuable for accurately valuing or hedging very long maturity contracts or assets such as physical power plants. Furthermore, this process does not require any optimization routine to infer the set of risk premiums that ensures that the model computed forward prices match the market data.

6 Conclusions

This paper develops a structural hybrid spot electricity price model constituting an energy price and an adjustment term. The first term is derived from an approximate economic dispatch problem, e.g., simple supply-demand model, which takes into account local fuel prices, total capacity, and total demand for the entire market operator territory. The adjustment term maps the energy price to zonal electricity prices through calibrating a piecewise probability distribution and captures the information content in the historical data. The proposed model seems to capture the critical features observed in the spot electricity market and can be calibrated to perfectly match the forward market data.

Relying on a computationally tractable approximate economic dispatch optimization problem and the full generation stack instead of an explicit parametric transformation function enables the model to promptly adapt to changes in the total demand or capacity levels in the market operator region, or some changes in the imposed regulation or the price formation mechanism. This feature makes the modeling framework robust and well suited for today's dynamic electricity market. In addition, since in contrast to most structural models the model does not assume an explicit parametric supply stack function, the energy price component of the model is capable of explaining better the sensitivity of the electricity prices to the underlying factors

such as fuel prices, especially in markets with complicated stack shapes and many generator types. Moreover, the model is still computationally tractable for energy trading purposes, such as forward calculations. In particular, it allows us to compute the electricity forward prices from the fuel forward prices to leverage their long forward curves. As another benefit of not relying on an explicit function, this model can readily be used for any bid-based electricity market, as long as a generation stack can be constructed from unit-specific data.

In this paper, we only discussed applying the model to compute electricity forward prices. This model can now be implemented to price a variety of different options and other derivative contracts, and for other trading purposes. Furthermore, the presented model can be applied in other power system planning problems. As the deregulated electricity markets continue to develop and provide new opportunities and challenges to both regulators, generators and consumers, it is clear that new modeling tools have a vital role to play in understanding and managing the many risks involved.

Acknowledgement(s)

This project is supported by the Public Service Electric and Gas Company (PSEG), through the Andlinger Center for Energy and the Environment at Princeton University. The authors would like to thank Belgacem Bouzaiene-Ayari and Rebecca Zhang for their contributions in this project.

References

Aid, R., L. Campi, A. N. Huu, and N. Touzi (2009). A structural risk-neutral model of electricity prices. *International Journal of Theoretical and Applied Finance* 12(7), 925–947.

Aid, R., L. Campi, and N. Langrene (2013). A structural risk-neutral model for pricing and hedging power derivatives. *Mathematical Finance* 23(3), 387–438.

Barlow, M. (2002). A diffusion model for electricity prices. *Mathematical Finance* 12(4), 287—298.

Barz, G. and B. Johnson (1999). *Energy Modelling and the management of uncerianty*. RISK books, 1 edition.

Benth, F. E., A. Cartea, and R. Kiesel (2008). 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.

Brass, H. and K. Petras (2011). *Quadrature Theory*. American Mathematical Society.

Burger, M., B. Klar, A. Muller, and G. Schindlmayr (2004). A spot market model for pricing derivatives in electricity markets. *Quantitative Finance* 4, 109—-122.

Carmona, R. and M. Coulon (2013). A survey of commodity markets and structural models for electricity prices. *In: Quantitative energy finance: modeling, pricing, and hedging in energy and commodity markets. Springer-Verlag, New York. .*

Carmona, R., M. Coulon, and D. Schwarz (2013). Electricity price modeling and asset valuation: A multi-fuel structural approach. *Mathematics and Financial Economics* 7(2), 167–202.

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

Cartea, A. and P. Villaplana (2008). Spot price modeling and the valuation of electricity forward contracts: the role of demand and capacity. *Journal of Banking and Finance* 32, 2501—2519.

Conejo, A. J., M. Carrion, and J. M. Morales (2010). *Decision Making under Uncertainty in Electricity Markets*. Springer, 1 edition.

Coulon, M. and S. Howison (2009). Stochastic behaviour of the electricity bid stack: From fundamental drivers to power price. *Journal of Energy Markets* 2(1), 29–69.

Coulon, M., W. B. Powell, and R. Sircar (2013). A model for hedging load and price risk in the texas electricity market. *Energy Economics* .

Davison, M., C. L. Anderson, B. Marcus, and K. Anderson (2002). Development of a hybrid model for electrical power spot prices. *IEEE Transactions on Power Systems* 17(2), 257–264.

de Maere d'Aertrycke, G. and Y. Smeers (2010). The valuation of power futures based on optimal dispatch. *The Journal of Energy Markets* 3(3), 27–50.

Deng, S. (2000). Stochastic models of energy commodity prices: Mean-reversion with jumps and spikes. *Program on Workable Energy Regulation (POWER) Report* pp. 1–45.

Eydeland, A. and K. Wolyniec (2002). *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Wiley Finance, 1 edition.

Golub, G. and G. Meurant (2010). *Matrices, moments, and quadrature with applications*. Princeton, N.J. ; Woodstock : Princeton University Press.

Johnson, B. and G. Barz (1999). *Selecting stochastic processes for modelling electricity prices*, chapter Energy Modelling and the Management of Uncertainty. Risk Publications, London.

Kanamura, T. and K. Ōhashi (2007). A structural model for electricity prices with spikes: Measurement of spike risk and optimal policies for hydropower plant operation. *Energy Economics* 29, 1010–1032.

Kleiber, C. and S. Kotz (2003). *Statistical SizeDistributions in Economics and Actuarial Sciences*. JohnWiley and Sons.

Longstaff, F. A. and A. W. Wang (2004). Electricity forward prices: A high-frequency empirical analysis. *The Journal of Finance* LIX(4).

Lyle, M. and R. Elliott (2009). A simple hybrid model for power derivatives. *Energy Economics* 31, 757—-767.

McDonald, R. L. (2013). *Derivatives Markets*. Pearson Series in Finance, 3rd edition edition.

Monitoring Analytics LLC Report (2005). 2004 state of the market: Market monitoring unit. Technical report.

Monitoring Analytics LLC Report (2013). 2012 state of the market report for PJM. volume 2: Detailed analysis. Technical report.

Ott, A. L. (2003). Experience with PJM market operation, system design, and implementation. *IEEE Transactions On Power Systems* 18(2), 528–534.

Pirrong, C. (2012). *Commodity Price Dynamics: A Structural Approach*. Cambridge University Press.

Pirrong, C. and M. Jermakyan (2008). The price of power: The valuation of power and weather derivatives. *Journal of Banking and Finance* 68(32), 2520—-2529.

Schweppe, F. C., M. C. Caramanis, R. D. Tabors, and R. E. Bohn (1988). *Spot Pricing of Electricity*. Power Electronics and Power Systems. Kluwer Academic Publishers, 1 edition.

Seydel, R. U. (2006). *Tools for Computational Finance*. Springer, 3 edition.

Skantze, P., A. Gubina, and M. Ilic (2000). Bid-based stochastic model for electricity prices: the impact of fundamental drivers on market dynamics. *MIT E-lab report* .

Skantze, P., M. Ilic, and A. Gubina (2004). Modelling locational price spreads in competitive electricity markets; applications for transmission rights valuation and replication. *IMA Journal of Management Mathematics* 15, 291—319.

Skantze, P. L. and M. D. Ilic (2001). *Valuation, Hedging and Speculation in Competitive Electricity Markets: A Fundamental Approach*. Kluwer Academic Publishers Group.

Swindle, G. (2014). *Valuation and Risk Management in Energy Markets*. Cambridge University Press.

Thompson, M., M. Davidson, and H. Rasmussen (2004). Valuation and optimal operation of electrical power plants in deregulated markets. *Operations Research* 52, 546–562.

Ventyx Inc. (2012). Unit capacity blocks & ramp rates.

Veraart, A. E. D. and L. A. M. Veraart (2013). Risk premia in energy markets. *working paper, Department of Economics and Business, Aarhus University* pp. 1–29.

Villaplana, P. (2005). Valuation of electricity forward contracts: The role of demand and capacity. *working paper, Department of Economics and Business, Universitat Pompeu Fabra* pp. 1–36.

Weron, R. (2014). Electricity price forecasting: A review of the state-of-the-art with a look into the future. *International Journal of Forecasting* 30(4), 1030–1081.

Wood, A. J. and B. F. Wollenberg (1996). *Power Generation, Operation, and Control*. Wiley-Interscience, 2 edition.

A Proof of Proposition 5.1

Here, we provide a proof of Proposition 5.1 presented in Section 5.

Proof. Under the equivalent $\mathbb{Q}$-martingale measure,

$$\begin{aligned} d\log P_t^{\mathbf{G}} &= \kappa_G \left(\hat{m}^{\mathbf{G}}(t) - \log P_t^{\mathbf{G}}\right) dt + \eta_G \; dW_t^{\mathbf{G}}, \\ d\log P_t^{\mathbf{C}} &= \kappa_C \left(\hat{m}^{\mathbf{C}}(t) - \log P_t^{\mathbf{C}}\right) dt + \eta_C \; dW_t^{\mathbf{C}}. \end{aligned}$$

Solving these SDEs by standard techniques for Ornstein-Uhlenbeck processes, we find that under the risk-neutral measure $\mathbb{Q}$, fuel prices are jointly lognormal, with

$$\begin{pmatrix} \log P_t^{\mathbf{G}} \\ \log P_t^{\mathbf{C}} \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} \mu_G(t,T) \\ \mu_C(t,T) \end{pmatrix}, \Sigma_{\mathbf{GC}}(t,T) \right),$$

where $\mu_G(t,T)$, $\mu_C(t,T)$, and $\Sigma_{\mathbf{GC}}(t,T)$ are defined as follows:

$$\mu_G(t,T) \stackrel{\text{def}}{=} P_t^{\mathbf{G}} e^{-\kappa_G(T-t)} + \kappa_G \int_t^T \bar{m}^{\mathbf{G}}(u) e^{-\kappa_G(T-u)} du, \tag{29}$$

$$\mu_C(t,T) \stackrel{\text{def}}{=} P_t^{\mathbf{C}} e^{-\kappa_C(T-t)} + \kappa_C \int_t^T \bar{m}^{\mathbf{C}}(u) e^{-\kappa_C(T-u)} du, \tag{30}$$

$$\Sigma(t,T) \stackrel{\text{def}}{=} \begin{pmatrix} \frac{\eta_G^2}{2\kappa_G}\left(1 - e^{-2\kappa_G(T-t)}\right) & \frac{\rho_{\mathbf{CG}}\,\eta_{\mathbf{G}}\,\eta_{\mathbf{C}}}{\kappa_{\mathbf{G}}+\kappa_{\mathbf{C}}}\left(1 - e^{(-(\kappa_{\mathbf{G}}+\kappa_{\mathbf{C}})(T-t))}\right) \\ \frac{\rho_{\mathbf{CG}}\,\eta_{\mathbf{G}}\,\eta_{\mathbf{C}}}{\kappa_{\mathbf{G}}+\kappa_{\mathbf{C}}}\left(1 - e^{(-(\kappa_{\mathbf{G}}+\kappa_{\mathbf{C}})(T-t))}\right) & \frac{\eta_C^2}{2\kappa_C}\left(1 - e^{-2\kappa_C(T-t)}\right) \end{pmatrix}. \tag{31}$$

Therefore,

$$\begin{aligned} F^{\mathbf{G}}(t,T_s) &= \mathbb{E}_t^{\mathbb{Q}}[P_{T_s}^{\mathbf{G}}] \\ &= \exp\left(\mathbb{E}_t^{\mathbb{Q}}[\log P_{T_s}^{\mathbf{G}}] + \frac{1}{2}\mathrm{Var}_t^{\mathbb{Q}}[\log P_{T_s}^{\mathbf{G}}]\right) \\ &= \exp\left(\mu_G(t,T_s) + \frac{\eta_G^2}{4\kappa_G}\left(1 - e^{-2\kappa_G(T_s-t)}\right)\right) \\ &= (P_t^{\mathbf{G}})^{\exp(-\kappa_G(T_s-t))} \exp\left(\kappa_G \int_t^{T_s} \hat{m}_G(u) \exp(-\kappa_G(T_s-u)) du + \frac{\eta_G^2}{4\kappa_G}\left(1 - e^{-2\kappa_G(T_s-t)}\right)\right) \\ &= (P_t^{\mathbf{G}})^{\exp(-\kappa_G(T_s-t))} \exp\left(\sum_{i=1}^{s} \hat{m}_i^{\mathbf{G}}\left(e^{-\kappa_G(T_s-T_i)} - e^{-\kappa_G(T_s-T_{i-1})}\right) + \frac{\eta_G^2}{4\kappa_G}\left(1 - e^{-2\kappa_G(T_s-t)}\right)\right). \end{aligned}$$

Thus, now we have

$$\sum_{i=1}^{s} \hat{m}_i^{\mathbf{G}}\left(e^{-\kappa_G(T_s-T_i)} - e^{-\kappa_G(T_s-T_{i-1})}\right) = \log(F^{\mathbf{G}}(t,T_s)) - e^{-\kappa_G(T_s-t)}\log(P_t^{\mathbf{G}}) - \frac{\eta_G^2}{4\kappa_G}\left(1 - e^{-2\kappa_G(T_s-t)}\right).$$

This gives us a system of N linear equations in N unknowns and N equations, which is straightforward to solve starting with the shortest maturity, T_1, and progressing through to the the longest, T_N. The result is the set of equations for $\hat{m}_s^{\mathbf{G}}$ and $\hat{m}_s^{\mathbf{C}}$ given in the proposition. □

B Estimated Parameters for the PJM Load Seasonal Component $S_t^{\mathbf{L}}$

Historical Data Used for Estimation: $1-\text{Jan}-2008$ to $31-\text{May}-2013$						
Hour	a_{1,h_t}	a_{2,h_t}	a_{3,h_t}	a_{4,h_t}	a_{5,h_t}	a_{6,h_t}
1	70234.08	2786.39	3998.90	−10934.75	3998.71	1658.52
2	66901.92	1290.26	3998.90	−10448.35	4004.97	1124.91
3	66072.50	−1533.93	4000.72	9509.84	4001.92	−1188.24
4	65254.86	−1650.49	3999.73	8992.93	4001.87	−867.23
5	66189.35	−2590.16	3999.52	8688.05	4001.88	−1379.34
6	69986.77	−3773.01	3999.43	8303.63	4001.87	−2291.68
7	76889.58	−5660.94	3999.43	7391.99	4001.83	−4123.50
8	82278.53	−5390.69	3999.49	7754.40	4001.91	−5221.06
9	84541.42	−2620.26	3999.81	8916.51	4001.92	−3802.88
10	86079.60	2110.27	4004.61	10229.56	4001.88	−2036.66
11	87511.62	5264.35	3998.83	−11612.58	3998.70	−697.23
12	88316.57	8558.22	3998.93	12874.04	4008.10	417.60
13	88717.48	11361.95	3998.97	13888.18	4001.80	1290.54
14	89057.36	13726.05	3998.98	−14676.11	4004.92	1907.44
15	88912.62	−15494.96	4008.41	−15352.39	4004.91	2620.62
16	88852.28	−16419.83	3989.56	−15931.74	4004.92	3160.35
17	89268.91	−15703.87	3976.99	−16693.30	4004.96	4297.49
18	90590.09	12361.09	3998.91	−17873.63	4005.03	5140.57
19	92084.67	9660.72	3998.79	−17542.24	4005.04	2583.45
20	91773.89	8015.88	3998.71	−15312.54	4004.99	1570.87
21	90386.12	7102.83	3998.88	13460.78	4008.07	2628.41
22	87426.45	7337.98	3999.02	13263.25	4008.12	2834.46
23	81740.82	6219.28	3999.03	−12760.69	4005.01	2585.68
24	75532.76	4561.42	3998.96	−11864.82	3998.72	1796.26

Table 1: Parameters relating to $S_t^{\mathbf{L}}$.