

# A SURVEY OF COMMODITY MARKETS AND STRUCTURAL MODELS FOR ELECTRICITY PRICES

RENE CARMONA AND MICHAEL COULON

ABSTRACT. The goal of this survey is to review the major idiosyncrasies of the commodity markets and the methods which have been proposed to handle them in spot and forward price models. We devote special attention to the most idiosyncratic of all: electricity markets. Following a discussion of traded instruments, market features, historical perspectives, recent developments and various modeling approaches, we focus on the important role of other energy prices and fundamental factors in setting the power price. In doing so, we present a detailed analysis of the *structural* approach for electricity, arguing for its merits over traditional *reduced-form* models. Building on several recent articles, we advocate a broad and flexible structural framework for spot prices, incorporating demand, capacity and fuel prices in several ways, while calculating closed-form forward prices throughout.

## 1. Introduction

The non-storability of electricity and the wide availability of supply and demand data allow us to understand and analyze the relationship between prices and underlying drivers more easily than in most other markets. These characteristics naturally led to the development of a branch of literature which we refer to as *structural* models of electricity prices. Making use of similar mathematical tools to the reduced-form models, structural models dig one level deeper, by identifying at least some of the fundamental sources of randomness which appear simply as unobservable diffusion or jump processes in a typical reduced-form approach. In many cases, including such fundamental variables leads to new challenges, due to the very complicated nature of the price setting mechanism in power markets, and difficulty in piecing together the key components of the puzzle. Nonetheless, the extra insight on the causes of power price movements brings significant benefits, both in terms of adapting to changing market environments and different locations, as well as in capturing cross-commodity correlations and demand dependence which is crucial for accurate pricing of many common derivatives products and physical assets. Structural models stop short of fully replicating the intricacies of the price setting mechanism (as described by optimization-based stack models) in order to retain tractability and emphasise dominant relationships. Thus, a balance is typically struck between mathematical convenience and model realism. As such, a broad range of structural models exist, which differ both in the number of fundamental relationships they choose to capture and in the techniques used to capture them.

Electricity is a commodity and as a result, the electricity markets are most often introduced and studied within the broader framework of the commodity markets. Even though a significant amount of electricity is generated from renewable sources (e.g. wind and solar) or hydro or nuclear sources, the main production process remains the conversion of fossil fuels like coal, gas and oil. Since electricity is often traded on exchanges just a few hours before it is needed, the overall cost of production is essentially the cost of the fuels used in the production, even in markets with a substantial amount of hydro and nuclear production as these plants are hardly ever setting the price. In other words, since electricity is essentially

---

2010 *Mathematics Subject Classification*. Primary 91G20, 91B76.

Partially supported by NSF - DMS 0806591.

Partially supported by NSF - DMS-0739195.

not storable, it must be consumed as it is produced and the costs of production are an important part of the computation of the supply curve. For this reason, electricity price formation cannot be dissociated from the prices of the *fuels* used in its production.

In this context, it is clear that valuation should be done by equilibrium arguments matching supply and demand. This paper reviews some of the mathematical models used by academics and practitioners alike, and provides an introduction to the class of structural models which build on this idea. The present introduction gives a series of anecdotes illustrating the recurrent themes developed in the paper.

Electricity burst onto the financial scene with deregulation and the transition from a system where production, transportation, and distribution of electricity were vertically integrated under the monopoly of utilities, to a set of open competitive markets for production and retail, while the grid remained under control. This unbundling happened over a few years in several parts of the world, but was not equally successful. NordPool (Northern Europe), ERCOT (Texas), PJM (North East of the US) are generally regarded as successes but the California experience of the early 2000's was controversial and most of its original initiatives ended up being reversed in the long run. In any case, deregulation opened up new markets and a new price formation mechanism emerged based on constant supply - demand balance. While electricity shares equilibrium pricing with most other commodities, it stands out by the construction of the supply curve where the different modes of production (hydro, nuclear, solar, wind, coal, oil, gas, etc) are ordered (the resulting order being called *merit order*) in increasing order of costs of production, resulting in what is known as the production *stack*. Matching supply with demand leads to the concept of plant on the margin (or technology on the margin) which is fundamental in the understanding of price formation for electricity, and which is at the heart of the approach taken in this presentation.

The business of producing, delivering and retailing electricity is very complex. It requires capital intensive investments and long term financing. Financial mathematics and financial engineering have an important role to play, far beyond the traditional support of portfolio management. The first challenge has to do with a very different breed of data analysis: costs and prices are not always available, and when they are, the amount and the complexity of the data can be overwhelming. The multitude of locations (e.g. nodal pricing), the diverse nature of the electricity contracted (spot, day-ahead, on-peak, off-peak, firm, non-firm, forward, . . .), and the fact that, contrary to other commodities and financial products, electricity prices can be *negative*. And as if the challenges of the analysis of electricity price data was not enough, quants have to deal with a slew of derivative products with embedded features rarely seen on the traditional financial markets. They include features known as swings, recall / take-or-pay options, etc, and new derivatives intended to help market participants hedge some of the risks associated with physical factors impacting the bottom line (weather and emissions, tolling agreements, shipping and freight, gas storage, cross commodity derivatives, etc). While being a constant nightmare for regulators and managers, the complexity and the diversity of these derivatives became a *bonanza* for financial engineers and financial mathematicians who discovered a brand new source of challenging modeling and pricing problems. See for example [57, 32, 23] and the more recent articles [80, 14, 75] for a sample of mathematical and numerical developments prompted by the analysis of swing options.

Finally, the need to quantify the credit-worthiness of counter-parties and integrate this information in the valuation algorithms, became painfully obvious after the collapse of Enron and the ensuing rash of defaults in the industry. Ironically, Enron was one of the very first companies advocating the need to take counterparty credit-worthiness into account in any valuation exercise. The avalanche of bankruptcies and credit downgrades following Enron's collapse highlighted the need for a deep understanding of the statistics of credit migration, appropriate ways to include counter-party risk in the valuation of transactions, and possibly the enhancement of credit protection with specific derivative instruments. Unfortunately,

many of these derivatives depend upon industry indexes based on actual movements in the markets and these indexes have been proven to be easy targets of manipulation. Systematic reliance on clearing houses has been proposed as the ultimate solution to these uncertainties, and living with collateral requirements and margin calls is part of the every-day life of an energy trader. However, most transactions rely on tailor-made deals and it seems difficult to imagine that a minimal set of instruments could be designed in order to span all the energy contracts and make clearing a standard solution. We will not discuss these problems in this survey any further.

Under the influence of Enron, quants and academics alike embraced the real option approach to physical asset valuation, providing systematic ways to include the physical assets of a company (power plants, pipelines, barges, tankers, etc.) together with the financial instruments held at a given time, into a single portfolio. This innovative way to put together *apples and oranges* on the same book, opened the door to new forms of *hedging* the risks of financial positions using physical assets, or vice-versa. Undoubtedly, one of the most exciting challenges of the energy markets is the new breed of hedging imposed by the physical nature of the commodities underlying the financial contracts, and risk management of production and transportation facilities. Indeed, hedging the risks associated with mixtures of physical and financial assets is not part of the typical financial mathematics curriculum. While a necessity for electricity producers and retailers, it was perfected and developed into an art form by investment banks like Goldman Sachs, Morgan Stanley, JP Morgan and the like, which in order to optimize returns, have sited and leased power plants, and taken control of storage facilities and the transportation of goods via leasing of pipelines, tankers, etc. More than a decade later, these problems are still important drivers in academic research in commodity and energy market modeling, and the constant flow of academic publications on power plant and gas storage facility valuations is a case in point. While no obvious benchmark emerged, some of these methods are widely accepted and their use for marking to market purposes has become a common practice accepted by regulators. However, the physical nature of some of the assets in the portfolios of energy companies renders the computation of correlations and risk measures like Value at Risk (VaR) very much a challenge.

The simplest form of real option valuation of a power plant is to equate its value to a string of spread options, each option capturing the potential profit from the operation of the plant on a given day. In a nutshell this approach says that on any given day, if the difference between the price at which the electricity can be sold and the cost of the input fuels needed to produce it (plus the fixed costs of operation and maintenance of the plant) is positive, the plant should be run and this difference collected as a profit. While commonly relying on simple lognormal models for the prices of electricity and the input fuels (see for example [49, 25]), any pricing model with a new approach to capturing the dependence between electricity prices and the prices of the input fuels is likely to produce new plant valuation results. In this spirit, [31] suggests models integrating information about correlation contained in the prices of spread options traded on the market in the form of implied correlations. In this survey, we will review how the structural approach developed in [21] can provide valuations depending on future demand expectations and information contained in the forward curves of the input fuels. Viewing a power plant as a string of spread options is certainly not the only way to value power plants. More sophisticated methods use stochastic control techniques to take full advantage of the optionality of the plant. See for example [30, 3]. Moreover, some of these methods have been extended to value gas storage and we refer the interested reader to [28, 27, 50, 77]. and the references therein. However, as demonstrated in [21], the structural approach focuses more on energy price correlations and offers the flexibility of adapting to future scenarios for demand, capacity and input fuel forward prices.

The versatility and the adaptability of the structural approach is the main reason for our shameless attempt to promote it. As discussed further in Subsection 2, the commodity

and energy markets have seen dramatic changes in the last few years. The impact of some of these changes on electricity prices is rather subtle and cannot be easily captured by traditional reduced form models. The introduction of incentive programs favoring the use of renewable energy such as wind in Germany or solar in the US, the impact of mandatory regulations such as the European Union (EU) Emissions Trading Scheme (ETS) in Europe, the recent physical coupling of markets (e.g. France and Germany), the increase in correlation between stock and commodity prices due to index trading, the tightening of correlations between commodities included in these indexes, the dramatic drop in US natural gas prices following recent shale gas discoveries and large-scale development of fracking, etc. All of these changes are screaming for the use of flexible models which can accommodate these new relationships between the fundamental factors driving electricity prices. Historical prices may not be as relevant as forward-looking information and market knowledge: this gives structural approaches a big advantage over reduced-form models.

Excellent textbooks on mathematical models for the electricity (and other commodity) markets do exist, and we strongly recommend the reader to consult [13, 18, 38, 49, 52, 53, 67, 78] for the many aspects of the markets which we will not be able to cover in this survey.

We close this introduction with an outline of the contents of the paper. Section 2 gives a crash course on the commodity markets. The focus is mostly on energy and trading of the *fuels* entering the production of electricity. A discussion of the impact of index trading is included to emphasize, for better or worse, the growing socio-economic role played by commodity trading. The specific nature of the data needed to understand these markets is discussed and the importance of the forward markets is reflected in the construction of price models. The goal of Section 3 is to highlight how different electricity is from the other energy commodities. Its non-storability forces a difficult balancing act where supply and demand need to be matched in real time since electricity needs to be consumed as it is produced. Section 4 expands on the earlier discussions to introduce the building blocks of the structural models which we advocate in this survey. Section 5 then uses these ingredients to propose general classes of structural models for which closed-form prices of forward contracts can be found. We also discuss various issues related to model fitting and calibration, before concluding in Section 6.

## 2. Generalities on the Commodity Markets

As explained in the introduction, in order to understand the fundamentals of electricity prices, and especially the rationale for the structural models which we advocate, it is important to understand how electricity is produced, and the costs associated to the various fuels used in the process. This is the main reason for the need to understand the crude oil, coal and natural gas markets (before returning to electricity in the next section). Despite the fact that these represent only a small part of the commodity world, we discuss their main features as they pertain to commodity markets in general.

**2.1. Trading Commodities.** Commodities are considered as a separate asset class. Because of the physical nature of the interest underlying the contracts, their prices are determined by equilibrium arguments which involve matching supply and demand for the physical commodity itself. On the supply side, estimating and predicting inventories and quantifying the costs of storage and delivery are important factors which need to be taken into account. This is not always easy in the context of standard valuation methods which are mostly based on traditional finance theory (think for example of NPV which attempts to compute the present value of the flow of future dividends).

Whether they were called spot markets (when they involved the immediate delivery of the physical commodity), or forward markets (when delivery was scheduled at a later date),

commodity markets started as physical markets. Trading volume exploded with the appearance of financially settled contracts. While forward contracts are settled Over the Counter (OTC), and as such, carry the risk that the counterparty may default and not meet the terms of the contract, most of the financially settled contracts are exchange-traded futures for which the exchange acts as clearing house controlling default risk by a system of margin calls and attracting speculators to provide liquidity to the markets. While trading in physically and financially settled contracts were traditionally the two ways an investor could gain exposure to commodities, the creation of indexes and the increasing popularity of index tracking Exchange Traded Funds (ETFs) have offered a new way to gain exposure to commodities. Investing in commodities was promoted as the perfect portfolio diversification tool as they were *believed to be negatively correlated with stocks*. The exponential growth of this new form of investment in commodities which took place over the last decade may have been a self-defeating prophecy as recent econometric studies have shown that this form of index trading has created new correlations between commodities and stocks, and between the commodities included in the same index. Furthermore, Bouchouev [17] argues that the influence of investors has overturned Keynes' well-known 'theory of normal backwardation', causing a recent predominance of forward curves in contango, thus further weakening the attractiveness of investing in these markets.

One of the many convenient features of commodity trading is the specialization of the exchanges, leading to a simple correspondence between commodities and locations where they are traded. In other words, a given commodity is traded on *one or a small number* of specialized exchanges. The following table gives a few examples of some of these exchanges in the US and in Europe.

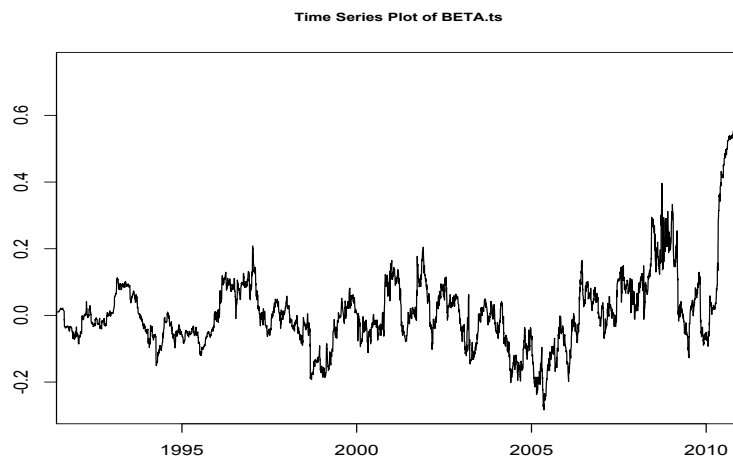
Exchange	Location	Contracts
Chicago Board of Trade (CBOT)	Chicago	Grains, Ethanol, Metals
Chicago Mercantile Exch. (CME)	Chicago, US	Meats, Currencies, Eurodollars
Intercontinental Exch. (ICE)	Atlanta, US	Energy, Emissions, Agricultural
Kansas City Board of Trade (KCBT)	Kansas City, US	Agricultural
New York Merc. Exch. (NYMEX)	New York, US	Energy, Prec. Metals, Indust. Metals
Climex (CLIMEX)	Amsterdam, NL.	Emissions
NYSE Liffe	Europe	Agricultural
European Climate Exch. (ECX)	Europe	Emissions
London Metal Exch. (LME)	London, UK	Industrial Metals, Plastics

There are several ways in which investors gain exposure to commodities.

1. The *old fashion way* to invest in commodities is to actually *purchase* the physical commodity itself. However most investors are not ready or equipped to deal with issues of transportation, delivery, storage and perishability. This form of involvement in commodities was created for and is essentially limited to the *naturals*, namely the hedgers who mitigate the financial risks associated with uncertainties in their production and delivery of these commodities.
2. Another way to gain exposure to commodities is to invest in stocks in commodity intensive businesses: for example buying shares of Exxon or Shell as a way to invest in oil. However, this type of investment offers at best an indirect exposure as shares of natural resource companies are not perfectly correlated with commodity prices.
3. A more direct form is straight investment in commodity futures and options. The exchanges offer transparency and integrity through clearing and relatively small initial investments are needed to take large positions through leverage. However, this convenience comes at a serious price as discovered by many *rookies* who ended up choking, unable to face the

margin calls triggered by adverse moves of the values of the interests underlying the futures contracts. Also, purely speculative investments of this type may need to be structured with a careful rolling forward of the contracts approaching maturity in order to avoid having to take physical delivery of the commodity: trading wheat futures can be done from the comfort of an office set up in a basement, but taking physical delivery of one lot (i.e. 5,000 bushels) of wheat requires a large backyard!

4. The final way to gain exposure to commodity which we discuss is investing directly in *Commodity Indexes* or in *Exchange Traded Funds* (ETFs) tracking these commodity indexes. Many ETFs simply invest in the nearest forward contract and automatically ‘roll’ the investment into the next month’s contract near maturity. This form of *passive investment* (after all there is no need for a *Commodity Trading Advisor* (CTA) for that), has become very popular as a way to diversify an investment portfolio with an exposure to commodities without having to deal with the gory details of all the convoluted idiosyncrasies of the relevant markets. Nevertheless, an understanding of forward curve dynamics and the effect of monthly rolls is still vital, as a recent investor in the natural gas ETF would undoubtedly agree: between June 2008 and March 2012 this ETF (called UNG) lost a shocking 96% of its value, with roughly half attributable to the spot price drop and half to the steep contango witnessed throughout this period.

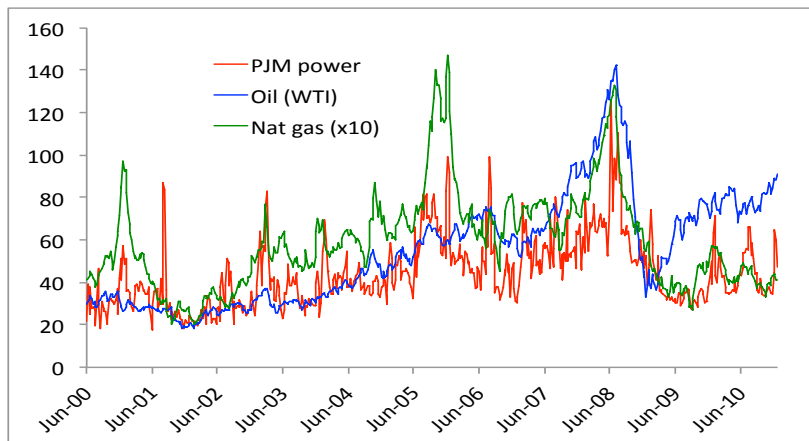


**Figure 1.** *Instantaneous Dependence ( $\beta$ ) of GSCI-TR returns upon S&P 500 returns.*

According to Barclays’ internal reports, in 2006 - 2007, index fund investment increased from 90 billion to 200 billion USD. Simultaneously, commodity prices increased 71% as measured by the CRB index. However, when prices declined dramatically from June 2008 through early 2009, many pointed to the large scale speculative buying by index funds, arguing that this created a *bubble* as futures prices far exceeded fundamental values. Some economists (including Nobel Prize winner P. Krugman, Pirrong, Sanders, Irwin, Hamilton and Kilian) remained skeptic about the “bubble theory” arguing that prices of commodities are set by supply and demand, and that rapid growth in emerging economies (e.g. China) increased demand and caused the 2008 surge in price. This did not stop commodity index investing from being under attack. Increased participation in futures markets by non-traditional investors was deemed *disruptive* and blamed for the 2007-2008 “Food Crisis” that is at the origin of the famous “*Casino of Hunger: How Wall Street Speculators Fueled the Global Food Crisis*”. A report from the U.S. Senate Permanent Subcommittee on Investigation “... finds that there is significant and persuasive evidence to conclude that these commodity index traders, in the aggregate, were one of the major causes of unwarranted increases in the

price of wheat futures contracts relative to the price of wheat in the cash market.....” . To add insult to injury, a group of 48 agriculture ministers meeting in Berlin said they were “... concerned that excessive price volatility and speculation on international agricultural markets might constitute a threat to food security...”, according to a joint statement handed out to reporters on Jan. 22, 2011. It is an empirical fact that return correlations are no longer what they used to be and now commonly accepted that commodity index trading *tightened correlations* between commodities [73]. However, many argue that this is a scale dependent phenomenon and it seems that high frequency traders do not see (and hence ignore) these correlation increases. Broadly speaking, the *financialization of commodities* should refer to the increased leverage and the exponential growth of financially settled contracts dwarfing their physically settled counterparts. More recently, this term has been used to refer to the significant impact of index trading on commodity prices, and even more narrowly speaking, to the increased correlations between the commodities included in the same index and also between equity returns and commodity index returns. This last fact is illustrated in Figure 1 which shows the time evolution as given by a Kalman filter, of the time-dependent “beta” of the least squares linear regression of the Goldman Sachs Commodity Index Total Return against the returns of the S&P 500 index.

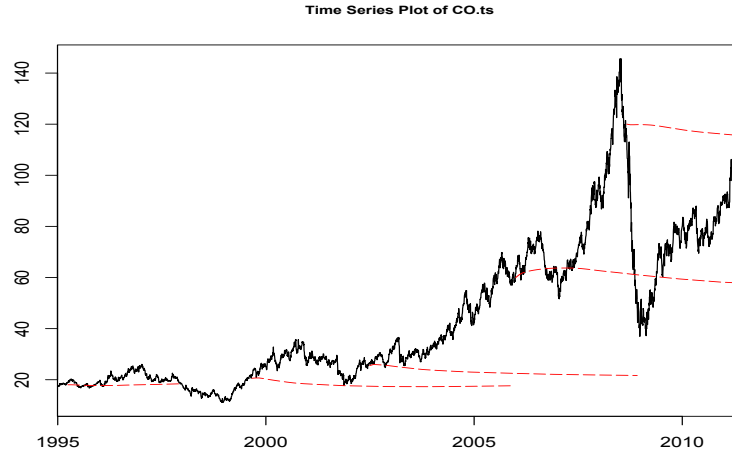
In this paper, our interest in commodities is mostly focused on the commodities used in the production of electricity, and in particular to crude oil and natural gas which are heavily represented in most commodity indexes. What we learn from the above discussion is that recent changes may affect their correlation and the correlation they have with the broader financial markets. Figure 2 shows weekly average spot (or nearest forward) prices for electricity, natural gas and crude oil, and illustrates the strong correlations between these energy commodities over a ten-year period. While the 2008 ‘bubble’ is most dramatic for crude oil, natural gas and power prices also rose sharply. In our search for electricity pricing models, it is important to bear in mind the diverse and changing factors affecting commodity markets in general.



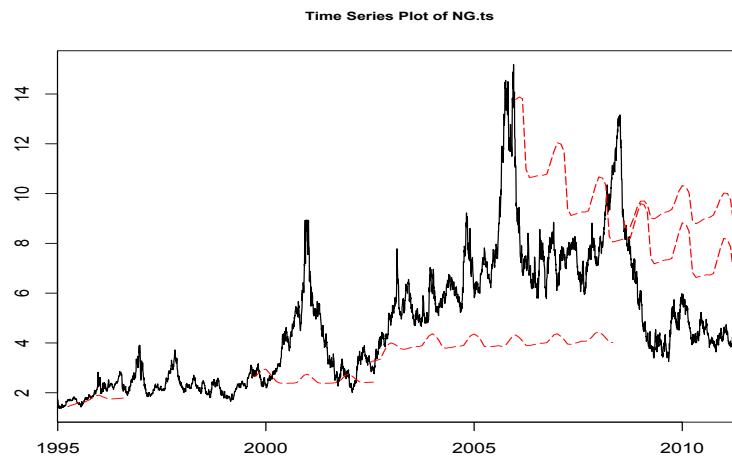
**Figure 2.** Weekly average prices for PJM (US electricity), Henry Hub (US natural gas) and WTI (US crude oil). Natural gas is multiplied by 10 to use the same axis.

**2.2. Spot and Forward Prices.** As we explained in the introduction, commodities are mostly traded through forward contracts and the first challenge of a quantitative analysis is the computation of the term structure of forward prices. Figures 3 and 4 give the time evolution of the price of the nearest forward contract of crude oil and natural gas respectively,

as used as a proxy for the spot price. In each case, we chose a few dates and superimpose the entire forward curve on these dates. We shall come back to these figures later in this section when we discuss the forward prices as expectations of future values of the spot price. In the case of crude oil or natural gas for which data are readily available, standard



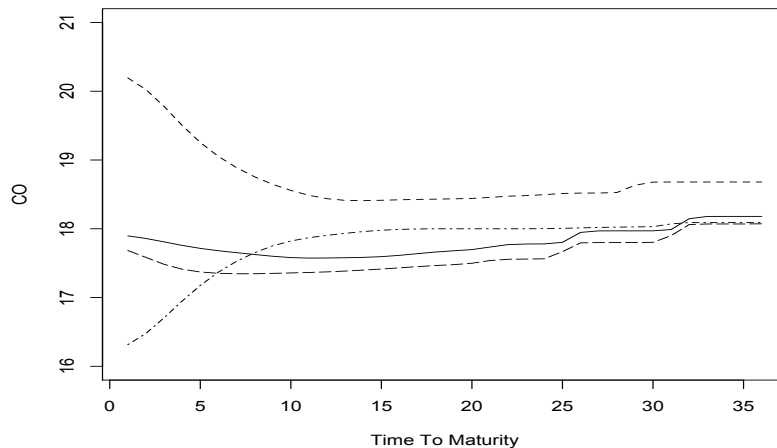
**Figure 3.** *Crude oil: time series of nearest forward prices and of a few forward curves*



**Figure 4.** *Natural gas: time series of nearest forward prices and of a few forward curves*

Principal Component Analysis (PCA) gives satisfactory results and shows that three factors are typically enough to explain over 95% of the variation in the daily changes in the forward curve. Like in the original analysis of the yield curve by Litterman and Scheinkman [63], the three factors are identified as parallel shift, tilt, and convexity. These account for the backwardation/contango duality illustrated in Figure 5 for the case of crude oil. However, the strong seasonality of natural gas (NG) forward curves makes an analysis in principal components more problematic and quite a significant “massaging” of the data is required for PCA to be used with any kind of success. Later in the section, we discuss a standard model for the time evolution of commodity forward curves. We use this theoretical model to give the fundamental rationale for PCA, and we explain within this framework, how seasonality





**Figure 5.** *Four different crude oil forward curves. Two are rather flat, one is in contango, the last one in backwardation.*

can be identified and accounted for. This leads to a procedure suggested first in [38] to perform PCA in the case of commodities with a strong seasonal component.

The analysis of electricity data can be more challenging as extreme complexity involving location, grade, peak/off peak, firm/non firm, interruptible, swings, etc, can muddy the water for the data analyst. Inconsistencies between different sources of information, illiquidity, wide bid-ask spreads, and delivery periods cascading from annual to quarterly to monthly as maturity approaches, etc, all require specific data manipulations which affect the outcome of the analysis. Nevertheless, Koekebakker and Ollmar [62] used PCA to show that 75% of the forward price variation can be explained by two factors, while this number is closer to 95% in other markets such as interest rates. Evidence from the Nordpool market indicates that long term forwards appear to be driven by different factors from short term forwards. Based on a similar PCA, Audet *et al* [5] propose a forward price structure with decreasing correlation as difference between maturity increases.

Throughout the paper, we shall use the notation  $F(t, T)$  for the value at time  $t$  of a forward contract with maturity  $T$ , and  $S(t)$  (or  $S_t$ ) for the spot price at time  $t$ . We use the term maturity by analogy with the fixed income markets, although delivery date is a term better suited to commodities. Forward contracts include the exact grade of the commodity to be delivered, and the terms of the delivery. For some commodities (for example natural gas and electricity), the date of delivery is not really a date, but a period over which the delivery is taking place. So a more appropriate notation would be  $F(t, T_1, T_2)$  if the delivery is spread uniformly over the time interval between  $T_1$  and  $T_2$ . In the case of electricity prices, we refer the interested reader to [13] for a lucid mathematical treatment of this issue, explanations of how to relate  $F(t, T)$  to  $F(t, T_1, T_2)$ , and modelling approaches geared towards handling delivery periods.

For commodities, the term spot price means the price of the commodity for immediate delivery. Mathematically, this would mean that  $S(t) = \lim_{T \searrow t} F(t, T)$ . In practice, immediate delivery is highly unrealistic, and different time lags before delivery exist for different markets. In many cases, we use the price of the nearest contract as a proxy for the spot price. This is in analogy with the use of the three month T-bill as a proxy for the instantaneous (short) interest rate in many studies. However, it is important to keep in mind the differences between commodities. Using the price  $F(t, T)$  of the nearest contract (i.e. the maturity date  $T$  closest to  $t$ ) gives a time-to-maturity lag  $T - t$  which varies from a few days to almost one

month as  $t$  varies. So when we use such a proxy for the spot price of crude oil or natural gas, the resulting approximation evolves over time as an accordion. However, when treating the price for next day delivery as the spot price of electricity, this phenomenon is not present since  $T - t$  remains constant and equal to one day!

One of the most useful concepts for the analysis of instruments traded on a forward basis is the *Spot - Forward* relationship which was proven to be a powerful tool in the analysis of financial markets where one can hold positions at no cost and easily take short positions. In that case, a simple arbitrage argument shows that

$$F(t, T) = S(t)e^{r(T-t)},$$

where  $r$  denotes the short interest rate. Since we are using a deterministic interest rate, forward and futures prices coincide in our mathematical treatment. Furthermore

$$F(t, T) = \mathbb{E}_t[S(T)]$$

where the above expectation is a risk neutral expectation conditioned by all the information available at time  $t$ . However, when bought with the intent to be sold later, a physical commodity needs to be transported and stored, adding to the cost of the financing of the purchase. On the other hand, holding the physical commodity can be advantageous, particularly in times of market stress or during supply shocks. The theory of storage was developed with the intention of explaining normal *backwardation* arguing that  $F(t, T)$  is a *downward biased* estimate of  $S(T)$ , namely the spot price exceeds the forward prices. (See Figures 3 and 4 for examples.) In order to translate this into a reduced-form expression and explain the different relationships between spot and forward prices, the notion of *convenience yield* was introduced. Next we review some of the quirks of this theory, even though we should keep in mind that while it can be applied to crude oil, natural gas, coal (fuels used in electricity production), it does not apply to electricity prices since for all practical purposes, electricity is not a storable commodity.

*The Case of Storable Commodities.* The argument above leads to the formula

$$F(t, T) = S(t)e^{(r-\delta)(T-t)}$$

for some quantity  $\delta \geq 0$  which is called the *convenience yield*. If we decompose this quantity in the form  $\delta = \delta_1 - c$  with  $\delta_1$  modeling the benefit from owning the physical commodity and  $c$  the costs of storage, then

$$F(t, T) = e^{r(T-t)}e^{-\delta_1(T-t)}e^{-c(T-t)}$$

where  $e^{r(T-t)}$  represents the *cost of financing* the purchase,  $e^{c(T-t)}$  the *cost of storage*, and  $e^{-\delta_1(T-t)}$  the sheer *benefit from owning* the physical commodity. The advantage of this representation, as artificial as it may be, is to explain the backwardation / contango duality within the proposed framework. Indeed, backwardation which occurs when the curve  $T \leftrightarrow F(t, T) = S(t)e^{(r+c-\delta_1)(T-t)}$  is decreasing holds when  $r + c < \delta_1$ , namely when benefits of holding the commodity outweigh interest rates and storage costs. On the other hand, the forward curve is in contango, namely the curve  $T \leftrightarrow F(t, T) = S(t)e^{(r+c-\delta_1)(T-t)}$  is increasing, when  $r + c \geq \delta_1$ . Empirical evidence shows that the convenience yield changes over time, and that it is related to several economic indicators, and in particular, inversely related to inventory levels. So it is natural, as we do next, to include it as a stochastic factor in a pricing model.

**2.3. Convenience Yield Models.** For quite a long time, the standard model has been the Gibson-Schwartz [55] two-factor model with factors given by the commodity spot price  $S_t$  and the convenience yield  $\delta_t$ . It posits risk neutral dynamics of the form

$$(1) \quad \begin{cases} dS_t &= (r_t - \delta_t)S_t dt + \sigma S_t dW_t^1, \\ d\delta_t &= \kappa(\theta - \delta_t)dt + \sigma_\delta dW_t^2. \end{cases}$$

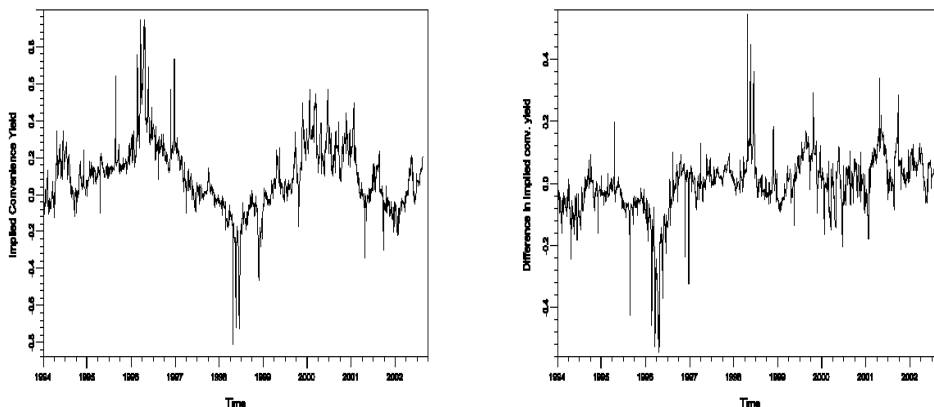
One of the major attraction of the model is that, being a particular case of the so-called exponential affine models, explicit formulas are available for many derivatives. In particular the prices of the forward contracts are given by

$$F(t, T) = S_t e^{\int_t^T r_s ds} e^{B(t, T)\delta_t + A(t, T)}$$

where

$$\begin{aligned} B(t, T) &= \frac{e^{-\kappa(T-t)} - 1}{\kappa}, \\ A(t, T) &= \frac{\kappa\theta + \rho\sigma_s\gamma}{\kappa^2} (1 - e^{-\kappa(T-t)} - \kappa(T-t)) + \\ &\quad + \frac{\gamma^2}{\kappa^3} (2\kappa(T-t) - 3 + 4e^{-\kappa(T-t)} - e^{-2\kappa(T-t)}). \end{aligned}$$

However, as demonstrated in [29], this strength of the model comes at a price. For any given maturity  $T$ , one can follow the time evolution of the forward price  $F(t, T)$  from market quotes, and from the above formulae, one can infer for each day  $t$ , the value of the convenience yield  $\delta_t$ . Internal consistency of the model requires that this implied convenience yield is independent of the choice of the particular contract maturity  $T$ . However, Figure 6 borrowed from [29] shows that this is not the case. Instabilities and inconsistencies in the implied  $\delta_t$  demonstrate that the two factor model ignores significant maturity specific effects.



**Figure 6.** *Crude oil convenience yield implied by a 3 month futures contract (left); Difference in implied convenience yields between 3 and 12 month contracts. (right)*

As suggested in [29], one possible way out of this quandary is to model directly the historical dynamics, for each fixed maturity  $T_0$ , of the forward price  $F_t = F(t, T_0)$  instead of the spot  $S_t$ , assuming that

$$\begin{aligned} dF_t &= (\mu_t - \delta_t)F_t dt + \sigma F_t dW_t^1, \\ d\delta_t &= \kappa(\theta - \delta_t)dt + \sigma_\delta dW_t^2 \end{aligned}$$

or more generally

$$d\delta_t = b(\delta_t, F_t)dt + \sigma_\delta(\delta_t, F_t)dW_t^2.$$

One can still compute the values of the convenience yield implied by the model. Indeed the assumption that  $F_t$  is tradable and observable while the forward convenience yield  $\delta_t$  is not, sets up a standard filtering problem which can be solved to construct a convenience yield

for each maturity. See [29] for details. There are other approaches to modeling the term structure of convenience yield, and the reader may want to consult [15] for a risk neutral approach à la Heath-Jarrow-Morton (HJM) which bares to the Gibson-Schwartz model (1) the same relationship as the classical HJM models to the standard short rate models.

**2.4. Dynamic Model for the Forward Curves.** In this subsection, we follow [38] to describe a standard HJM-like  $n$ -factor forward curve model which we use to derive the dynamics of the spot commodity model, and prepare for the explanations given in the next subsection on how to calibrate the model to price data using PCA, even when strong seasonal effects spoil a direct and naive application of the method. We start with a model under the historical measure

$$(2) \quad \frac{dF(t, T)}{F(t, T)} = \mu(t, T)dt + \sum_{k=1}^n \sigma_k(t, T)dW_k(t) \quad t \leq T$$

where  $\mathbf{W} = (W_1, \dots, W_n)$  is a  $n$ -dimensional standard Brownian motion, and the drift  $\mu$  and the volatilities  $\sigma_k$  are deterministic functions of  $t$  and the time-of-maturity  $T$ . Notice that  $\mu(t, T)$  will be set to zero for pricing purposes. In general,  $\mu(t, T)$  is calibrated to historical data for risk management applications. By the simplicity of this lognormal model, explicit solutions exist for the forward prices:

$$F(t, T) = F(0, T) \exp \left[ \int_0^t \left[ \mu(s, T) - \frac{1}{2} \sum_{k=1}^n \sigma_k(s, T)^2 \right] ds + \sum_{k=1}^n \int_0^t \sigma_k(s, T) dW_k(s) \right]$$

and the forward prices are *log-normal* random variables of the form

$$F(t, T) = \alpha e^{\beta X - \beta^2/2}$$

with  $X \sim N(0, 1)$  and

$$\alpha = F(0, T) \exp \left[ \int_0^t \mu(s, T) ds \right], \quad \text{and} \quad \beta = \sqrt{\sum_{k=1}^n \int_0^t \sigma_k(s, T)^2 ds}$$

From these, we can derive an expression for the spot price  $S(t) = F(t, t)$  defined as the left hand point of the forward curve:

$$S(t) = F(0, t) \exp \left[ \int_0^t \left[ \mu(s, t) - \frac{1}{2} \sum_{k=1}^n \sigma_k(s, t)^2 \right] ds + \sum_{k=1}^n \int_0^t \sigma_k(s, t) dW_k(s) \right]$$

and differentiating both sides we get an equation for its dynamics:

$$dS(t) = S(t) \left[ \left( \frac{1}{F(0, t)} \frac{\partial F(0, t)}{\partial t} + \mu(t, t) + \int_0^t \frac{\partial \mu(s, t)}{\partial t} ds - \frac{1}{2} \sigma_S(t)^2 - \sum_{k=1}^n \int_0^t \sigma_k(s, t) \frac{\partial \sigma_k(s, t)}{\partial t} ds + \sum_{k=1}^n \int_0^t \frac{\partial \sigma_k(s, t)}{\partial t} dW_k(s) \right) dt + \sum_{k=1}^n \sigma_k(t, t) dW_k(t) \right]$$

from which we can identify the spot volatility

$$(3) \quad \sigma_S(t)^2 = \sum_{k=1}^n \sigma_k(t, t)^2.$$

Hence, if we define the Wiener process  $\tilde{W}_t$  by  $\tilde{W}_t = \sigma_S(t)^{-1} \sum_{k=1}^n \sigma_k(t, t) dW_k(t)$ , then the dynamics of the spot can be rewritten in the form:

$$\frac{dS(t)}{S(t)} = \left[ \frac{\partial \log F(0, t)}{\partial t} + d(t) \right] dt + \sigma_S(t) d\tilde{W}_t$$

provided we define the drift component  $d(t)$  by:

$$\begin{aligned} d(t) = & \mu(t, t) - \frac{1}{2}\sigma_S(t)^2 + \int_0^t \frac{\partial\mu(s, t)}{\partial t} ds - \sum_{k=1}^n \int_0^t \sigma_k(s, t) \frac{\partial\sigma_k(s, t)}{\partial t} ds \\ & + \sum_{k=1}^n \int_0^t \frac{\partial\sigma_k(s, t)}{\partial t} dW_k(s). \end{aligned}$$

Looking more closely at the expression giving the drift we notice that, in a risk-neutral setting, the logarithmic derivative of the forward can be interpreted as a discount rate, while  $d(t)$  can be interpreted as a convenience yield. We also notice that the drift is generally *not Markovian*. However, in the particular case of a single factor, when  $\mu(t, T) \equiv 0$ , and  $\sigma_S(t) = \sigma_1(t, T) = \sigma e^{-\lambda(T-t)}$  which is consistent with what is known as the *Samuelson's effect*, we have

$$d(t) = \lambda[\log F(0, t) - \log S(t)] + \frac{\sigma^2}{4\lambda}(1 - e^{-2\lambda t}),$$

and the dynamics of the spot become

$$\frac{dS(t)}{S(t)} = [\mu(t) - \lambda \log S(t)]dt + \sigma dW(t),$$

which shows that in this case, the spot price is an exponential Ornstein Uhlenbeck process, an instance of the formal equivalence between mean reversion and the exponential decay of the forward volatility away from maturity.

**2.5. Rationale for PCA.** For data analysis and computational purposes, it is convenient to change variable from the time-of-maturity  $T$  to the time-to-maturity  $\tau$ . This changes the dependence upon  $t$  in several formulae. To be specific, if we set

$$t \mapsto F(t, T) = F(t, t + \tau) = \tilde{F}(t, \tau)$$

for pricing purposes, it is important to keep in mind that for  $T$  fixed,  $\{F(t, T)\}_{0 \leq t \leq T}$  is a **martingale** while for  $\tau$  fixed,  $\{\tilde{F}(t, \tau)\}_{0 \leq t}$  is **NOT**! The dynamics of the forward prices in this parameterization (known as *Musiela parameterization*) become

$$d\tilde{F}(t, \tau) = \tilde{F}(t, \tau) \left[ \left( \tilde{\mu}(t, \tau) + \frac{\partial}{\partial \tau} \log \tilde{F}(t, \tau) \right) dt + \sum_{k=1}^n \tilde{\sigma}_k(t, \tau) dW_k(t) \right], \quad \tau \geq 0$$

if we set:

$$\tilde{\mu}(t, \tau) = \mu(t, t + \tau), \quad \text{and} \quad \tilde{\sigma}_k(t, \tau) = \sigma_k(t, t + \tau).$$

We use the above model for the evolution of the forward curves to justify PCA, and in so doing, we explain how to handle seasonal effects (as seen in the case of natural gas). Our fundamental assumption is that the volatilities appearing in (2) are of the form

$$\sigma_k(t, T) = \sigma(t)\sigma_k(T - t) = \sigma(t)\sigma_k(\tau)$$

for some function  $t \mapsto \sigma(t)$ . Then, the spot volatility  $\sigma_S(t)$  defined in (3) becomes

$$\sigma_S(t) = \tilde{\sigma}(0)\sigma(t)$$

provided we set

$$\tilde{\sigma}(\tau) = \sqrt{\sum_{k=1}^n \sigma_k(\tau)^2},$$

and as a consequence,  $t \mapsto \sigma(t)$  is (up to a constant) the instantaneous spot volatility. This simple remark provides us with a rationale for a new form of PCA which we now describe. First we fix times-to-maturity  $\tau_1, \tau_2, \dots, \tau_N$  and we assume that on each day  $t$ , quotes for the forward prices with times-of-maturity  $T_1 = t + \tau_1, T_2 = t + \tau_2, \dots, T_N = t + \tau_N$  are

available (some smoothing is required beforehand as these exact maturity dates are typically not available). From the model we know that

$$\frac{d\tilde{F}(t, \tau_i)}{\tilde{F}(t, \tau_i)} = \left( \tilde{\mu}(t, \tau_i) + \frac{\partial}{\partial \tau} \log \tilde{F}(t, \tau_i) \right) dt + \sigma(t) \sum_{k=1}^n \sigma_k(\tau_i) dW_k(t) \quad i = 1, \dots, N.$$

So if we define the matrix  $\mathbf{F}$  by  $\mathbf{F} = [\sigma_k(\tau_i)]_{i=1, \dots, N, k=1, \dots, n}$ , the instantaneous variance/covariance matrix  $\{M(t); t \geq 0\}$  defined by

$$M_{i,j}(t)dt = d[\log \tilde{F}(\cdot, \tau_i), \log \tilde{F}(\cdot, \tau_j)]_t$$

and satisfies

$$M(t) = \sigma(t)^2 \left( \sum_{k=1}^n \sigma_k(\tau_i) \sigma_k(\tau_j) \right) = \sigma(t)^2 \mathbf{F} \mathbf{F}^*.$$

We summarize the successive steps of the procedure in the following way:

- Estimate the instantaneous volatility  $\sigma(t)$  (e.g. in a rolling window);
- Estimate  $\mathbf{F} \mathbf{F}^*$  from historical data as the empirical auto-covariance of  $\ln(F(t, \cdot)) - \ln(F(t-1, \cdot))$  after normalization by  $\sigma(t)$ ;
- Perform a Singular Value Decomposition (SVD) of the auto-covariance matrix and extract the eigenvectors  $\tau \mapsto \sigma_k(\tau)$ ;
- Choose the order  $n$  of the model according to the rate of decay of the corresponding eigenvalues.

**2.6. New Commodity Markets.** While several new markets were introduced in the recent past, including for example freight trading, we limit this review to a short discussion of the two markets with relevance to electricity.

**2.6.1. The Weather Markets.** As will be emphasized once more in the next section, temperature is typically the dominant variable determining demand for electricity. This is certainly true in countries like the US, where air conditioning is the major source of demand in the summer, and heating is often a significant factor during the winter. In order to mitigate some of the risks associated with unpredictable fluctuations in demand, electricity producers and merchants have been the major driving force behind the design and the development of the weather markets. *“Weather is not just an environmental issue; it is a major economic factor. At least 1 trillion USD of our economy is weather-sensitive.”* (William Daley, 1998, US Commerce Secretary). It is estimated that 20% of the world economy is directly affected by weather, with the energy sector being concerned the most, followed by the entertainment and tourism industries. While we are not discussing these markets further for fear of distracting the reader from the main thrust of this review article, we refer the interested reader to a sample of papers addressing valuation issues [20, 69], risk transfer mechanism [10], the comprehensive book [6], and to the web site of the Weather Risk Management Association (WRMA) for more information about these markets. While temperature is the deepest and most liquid of the weather markets, other meteorological variables such as humidity and precipitation have also been shown to have significant correlations with electricity demand, while rainfall, cloud cover and wind speed clearly also affect electricity supply from hydro, solar and wind energy. Coupled with the impact of these variables on the revenues of businesses such as amusement parks or road construction, separate instruments were introduced, though not with the appeal and the success of temperature options. See for instance [24] for an example of rainfall option pricing.

2.6.2. *The Emissions Markets.* As equilibrium pricing of commodities is based on matching demand with supply, the latter being directly affected by the costs of production of electricity, any regulation changing these costs will have a significant impact on the price of electricity. Modelled after the successful cap-and-trade schemes used in the US acid rain program to control  $\text{SO}_x$  and  $\text{NO}_x$  emissions, the mandatory Emission Trading Scheme (ETS) created by the European Union (EU) for the purpose of meeting its  $\text{CO}_2$  emissions commitments within the framework of the Kyoto protocol, has demonstrated that for pricing purposes, the cost of emissions must be included in the costs of production. So for all practical purposes,  $\text{CO}_2$  emissions can be considered as an additional fuel and carbon allowance price as an additional factor driving electricity price. While early incarnations of allowance redemption for the purpose of emission offsetting was mostly done on a voluntary basis in the US, RGGI (Regional Green House Gas Initiative) covering 10 states in the North East of the country and the recently adopted California legislation have prompted electricity producers and merchants to include, like their European counterparts, the price of  $\text{CO}_2$  emissions in the price of electricity. We shall not dwell on this issue in this survey paper, but the reader may wish to consult [26, 56] for more on the link between power markets and equilibrium emissions allowance prices, as well as [22], where the structural approach in this paper is extended to include the cost of  $\text{CO}_2$  emissions when pricing spreads for the purpose of power plant valuation.

### 3. What is so Special About Electricity?

Given the material reviewed earlier, the obvious answer which first comes to mind is the fact that one cannot store the physical commodity (economically, in any meaningful quantity). But there are many other features which distinguish electricity from other commodities and this section is attempting to review how they impact electricity price formation.

The services provided by power traders include physical delivery of electricity as well as financial obligations. The delivery may be firm or non-firm, short or long term, one-time or stretching over time. In order for mathematical models for electricity prices to be tractable, they often ignore the diversity of these conditions and concentrate on easier to capture features. Like with other commodities, trading is mostly done on a forward basis but the nature of the delivery as well as the spectrum of delivery dates common in the electricity markets is quite peculiar. Typically what we mean by spot market is in fact a *day-ahead* market, so what we shall mean by forward market is a market on which contracts with deliveries beyond one day are traded. For longer term contracts, the delivery of the power as specified in the indenture of the contract has to take place **over** a period  $[T_1, T_2]$  as opposed to a fixed date as assumed by most mathematical models. Delivery periods are often monthly, but restricted to certain times of the day or week (e.g. on-peak or off-peak), and these should be treated differently because of significant differences in price levels and volatilities. Here, for the sake of simplicity, we shall only deal with contracts with fixed maturity dates and also avoid differentiating between deliveries at different times of the day or any other contract variations. While voluminous, electricity forward data are still sparse because of the large number of locations and flavors of deliveries, and despite the encouragements of the Committee of Chief Risk Officers of energy companies and their upbeat white papers, prices still lack transparency and poor reporting (or lack thereof) still hinder the development of healthy electricity markets.

Given the complexities of forward curve data, it is perhaps not surprising that the spot price often serves as the preferred starting point for modellers. Figure 8a gives a time series plot of the daily average spot price of electricity on the PJM exchange between 2000 and 2010. What we call the *spot price* is the market clearing price set for each hour in the day-ahead auction. The system operator needs advanced notice to make sure that the schedule

is feasible and transmission constraints are met. Hence a ‘day-ahead’ price is determined via a large optimization problem, but as rebalancing of supply and demand is required up until actual delivery, a ‘real-time’ price also exists (and is sometimes referred to as the spot price). In any case, the type of time evolution shown in Figure 8a has nothing in common with equity prices or even other commodity prices. The most obvious difference is the high frequency of sudden spikes when the price jumps up very quickly before dropping down to near its previous level in a very short amount of time. As a result, the volatility of the ‘returns’ (a questionable term for a non-storable commodity!) is excessively high, say in a range from 50% up to 200% which is very different from the volatility of other financial products.

In line with the structural approach described in the next section, we made a definite choice to answer the question “Which spot price should one use?”. But mathematical models could also be developed for the real-time price, the price on the balancing market, the balance-of-the-week price, the balance-of-the-month price, etc. For all these mathematical models to be consistent, the diversity of candidates begs the question: can a complete forward curve be constructed (for all  $T$ ) and does the forward price then converge to spot as the time to maturity goes to zero? If this is indeed the case, it would make sense to define the “mathematical spot price” as

$$S(t) = \lim_{T \downarrow t} F(t, T)$$

as we did in Subsection 2.4, and expect that its statistical properties will coincide with those of the day-ahead price chosen as a proxy.

**3.1. More Data Peculiarities.** Beyond the issues already mentioned (e.g. integrity, sparsity, etc), one of the most surprising features of electricity prices is the fact that some of them are frequently **negative**. If we consider for example the case of the PJM (Pennsylvania, New Jersey, Maryland) region in the North East of the US, every single day, real-time and day-ahead prices as well as hour by hour load prediction for the following day are published for over 3,000 nodes in the transmission network, and many negative prices can be found. For example, in 2003 over 100,000 such hourly instances occurred across the grid. They come in geographic clusters, at special times of the year (shoulder months) and times of the day (night and early morning). The first suspects are obviously errors in predictions of the load, and high temperature volatilities. More sophisticated explanations involve network transmission and congestion, causing an oversupply in one location and an undersupply in another. While we do not want to dwell on the issue of negative prices, it is a useful example to highlight the fact that electricity pricing cannot be done by mere application of techniques and results developed for the financial markets, and that the physical nature of the commodity, its demand patterns and the idiosyncrasies of its production and transmission need to be taken into account.

**3.2. Modeling the Demand: the Load / Temperature Relationship.** As explained earlier, demand for electricity in the US is in great part driven by weather conditions and especially temperature. Figure 7 illustrates this fact by showing that a simple regression can be used to predict the demand for electricity as a function of the temperature. As a result, weather dynamics need to be included in pricing and this adds another source of incompleteness to the mathematical models.

**3.3. Reduced-Form Models.** By nature, reduced-form models try to identify stylized properties of electricity prices, and capture them in simple relationships from which derivative prices can be obtained, preferably through analytic formulae. So instead of modeling the fundamentals of supply and demand and having prices appear as the result of equilibrium considerations, reduced-form models strive for tractability, and for this reason, they usually involve a small number of factors and parameters. The source of their popularity is the



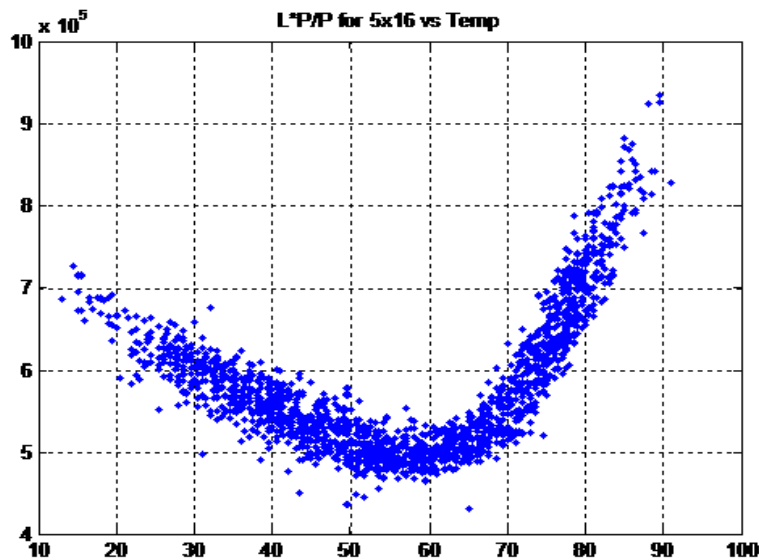


Figure 7. Daily Load versus Daily Temperature (PJM)

fact that their fairly simple formulation often leads to theoretical developments which can be tested against empirical evidence. In this spirit, the term structure of forward prices is most often derived from simple reduced-form models for the spot price via the spot-forward relationship discussed earlier.

An early spot price model by Lucia and Schwartz [64] proposed a two-factor diffusion model to capture the different short and long-term dynamics of power prices. Building on ideas in [71], this model was based on an ansatz of the form  $S_t = \exp(f(t) + X_t + Y_t)$  where  $f(t)$  is a seasonality function and the two factors  $X_t$  and  $Y_t$  satisfy

$$(4) \quad \begin{cases} dX_t = -\kappa X_t dt + \sigma_X dW_t \\ dY_t = \mu dt + \sigma_Y d\tilde{W}_t \end{cases}$$

and the two Brownian Motions  $W_t$  and  $\tilde{W}_t$  can be correlated. The initial success of the model can be attributed to the fact that spot and forward prices are lognormal in this model and Black-Scholes like formulae can be derived for option prices. However, the importance of electricity spikes prompted many authors to add jumps to the mix, leading to the popularity of jump-diffusion processes (cf. [34, 61]). As noticed in the analysis of credit models, including jumps does not necessarily mean giving up on closed-form formulae for forwards and options. Indeed working in the affine jump-diffusion framework promoted in [46] by Duffie, Pan and Singleton, still leads to convenient formulas for derivative prices. Indeed, if we assume that  $X_t \in \mathbb{R}^n$  is a vector of state variables,  $W_t$  a standard  $n$ -dimensional Wiener process, and  $Z_t$  a pure jump process, the times of jump forming a point process on  $[0, \infty)$  with intensity  $\lambda(X_t)$ , the jumps sizes being independent and identically distributed in  $\mathbb{R}^n$  with common distribution  $\nu$ , and if they satisfy

$$(5) \quad dX_t = \mu(X_t)dt + \sigma(X_t)dW_t + dZ_t$$

with

$$\mu(X_t) = A_1 + A_2 X_t \quad \sigma(X_t)\sigma(X_t)^\dagger = A_3 + A_4 X_t \quad \text{and} \quad \lambda(X_t) = A_5 + A_6 X_t,$$

where  $A_1 \in \mathbb{R}^n$ ,  $A_2, A_3 \in \mathbb{R}^{n \times n}$ ,  $A_4 \in \mathbb{R}^{n \times n \times n}$ ,  $A_5 \in \mathbb{R}^m$ , and  $A_6 \in \mathbb{R}^{m \times n}$  and we use the notation  $^\dagger$  to denote the transpose of a vector or a matrix, then the conditional characteristic

function of  $X_t$  has the form

$$\psi(u) = E_t[e^{u^\dagger X_T}] = e^{\alpha(t) + \beta(t)^\dagger X_t} \quad \text{for } t \leq T$$

for any  $u \in \mathbb{C}$ , where  $\alpha(t) \in \mathbb{R}$  and  $\beta(t) \in \mathbb{R}^n$  satisfy the Riccati ordinary differential equations

$$\begin{aligned} \frac{d}{dt}\alpha(t) &= -A_1^\dagger\beta(t) - \frac{1}{2}\beta(t)^\dagger A_3\beta(t) - A_5^\dagger[\zeta(\beta(t)) - 1] \\ \frac{d}{dt}\beta(t) &= -A_2^\dagger\beta(t) - \frac{1}{2}\beta(t)^\dagger A_4\beta(t) - A_6^\dagger[\zeta(\beta(t)) - 1] \end{aligned}$$

with  $\alpha(t) = 0$  and  $\beta(t) = u$ , and where  $\zeta(c) = \int_{\mathbb{R}^n} e^{c^\dagger z} dv(z)$ .

Deng [45] considers three cases of two-factor affine jump-diffusions, including deterministic volatility, stochastic volatility and regime-switching jumps. Exploiting the results above, derivative prices are calculated throughout, including cross-commodity spread options and locational spread options. Although Deng incorporates fuel prices in his models, correlation with power is achieved only through the matrix  $\sigma(X_t)$ , as opposed to the power price actually being a function of fuel prices (as we shall see later).

As another example, Culot *et al* [41] apply affine jump-diffusion models to the Amsterdam Power Exchange. The authors propose a three-factor mean-reverting component  $X_t$ , (different reversion speeds), combined with an independent three-factor jump component  $\tilde{X}_t$ . With spot price  $S_t = \exp(\gamma^\dagger X_t + \tilde{\gamma}^\dagger \tilde{X}_t)$ , this approach allows log forward prices to be affine functions of the state variables, and hence the Kalman Filter can easily be implemented for calibration. Derivative prices are calculated using a Fourier transform technique based on the work of Carr and Madan [33]. The jump (or spike) component involves regime-switching ideas, as  $\tilde{\gamma}^\dagger \tilde{X}_t$  can only equal zero or one of three possible spike levels, so jump sizes are fixed and a Markov chain transition matrix governs the intensities of all the possible jumps.

Benth *et al* [11, 13] have suggested several alternative jump-based models, using Ornstein-Uhlenbeck processes driven by Levy processes instead of Brownian Motions. In particular, they suggest approaches of the form

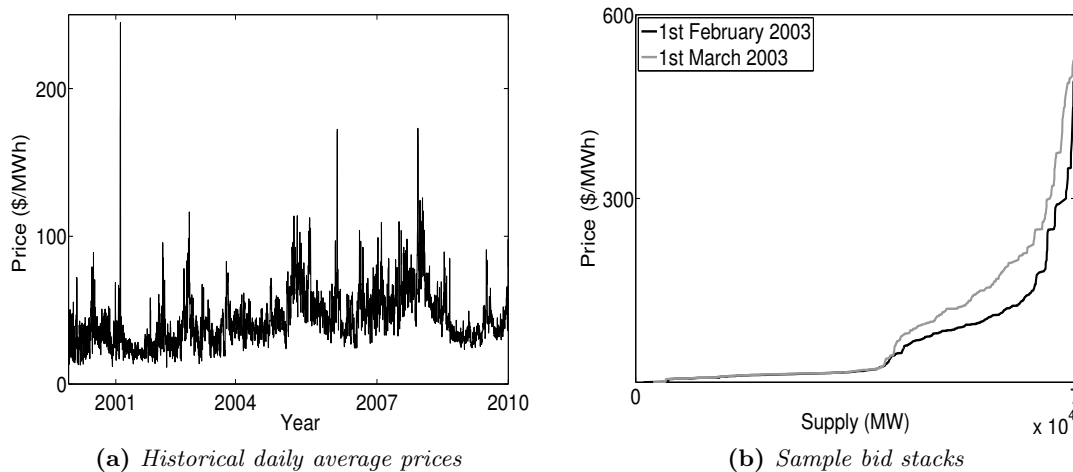
$$S_t = \sum_{i=1}^n w_i Y_t^i, \quad \text{where } dY_t^i = -\lambda_i Y_t^i dt + \sigma_t^i dL_t^i$$

where  $L_t^i$  are increasing pure jump processes, used to capture both small variations in the price (for certain  $i$ ), and the spikes (for other  $i$ ). By avoiding diffusion processes while maintaining an additive structure (instead of the more common exponential structure), the authors are able to find explicit formulas for forward prices without ignoring or approximating delivery periods. We recommend the book [13] for an exposition of various related approaches and extensions of this framework, including capturing cross-commodity correlation. More recently, Barndorff-Nielsen *et al* [9, 8] propose a new approach for both spot and forward prices using ambit fields, and in particular Levy semi-stationary processes.

In a reduced-form model, at least partial separation of jumps (or spikes) from more ‘normal’ diffusion factors is needed due to the large difference in spike recovery speed relative to other mean-reverting behavior. Possible approaches include the use of multiple factors with many speeds of mean reversion, regime switching jumps (which lead to downwards jumps to recover from spikes) or pure regime switching models. The last of these has been studied for example by De Jong and Huisman [43] and Weron *et al* [79], where independent dynamics are given for the ‘spike’ and ‘non-spike’ regime. Kholodnyi [60] retains a closer connection between

the two regimes in his model, instead suggesting that the price jumps from  $X_t$  to  $\lambda X_t$  for some constant  $\lambda$  when there is a regime switch. Regime switching models benefit from the fact that high prices can last for several time periods (typically just a few hours, so one should not interpret the terminology ‘regime’ to mean a lasting paradigm shift), reflecting for example periods of generator outages. The recovery from an outage can be as sudden as the outage itself, a characteristic difficult to mimic with mean-reverting jump-diffusions. A variation proposed by Geman and Roncoroni [54] is a jump-diffusion model which forces jumps to be downwards when prices are above a certain threshold.

While many of the models discussed above produce useful results and realistic price dynamics, they often face calibration challenges due to the need for multiple unobservable factors, an inability to adapt to changing market conditions, or to the complication of identifying historical spikes (or regimes). In addition, and perhaps more importantly from an industry perspective when managing complex portfolios of assets, they typically fail to capture the important correlations between power prices, other energy prices and power demand.



**Figure 8.** *Historical prices and bids from the PJM market in the North East US*

**3.4. A First Structural Model for Spot Prices.** For electricity as for all other commodities, the balancing act between supply and demand in the price formation leads to mean reversion of prices towards costs of production. Furthermore, the relationships between underlying supply and demand factors in electricity markets are more observable and better understood than in other markets. This has naturally led to the development of so-called structural models. In this category, the first real proposal for a tractable spot pricing model based on a supply/demand argument is due to Martin Barlow [7] and we review briefly the main components of his pricing model. Motivated by observed auction data, Barlow proposed to use a vertical demand curve (reminiscent of the inelasticity of the demand for electricity) and a supply curve given by a nonlinear function of a simple diffusion process:

$$S(t) = \begin{cases} f_\alpha(X_t) & 1 + \alpha X_t > \epsilon_0 \\ \epsilon_0^{1/\alpha} & 1 + \alpha X_t \leq \epsilon_0 \end{cases}$$

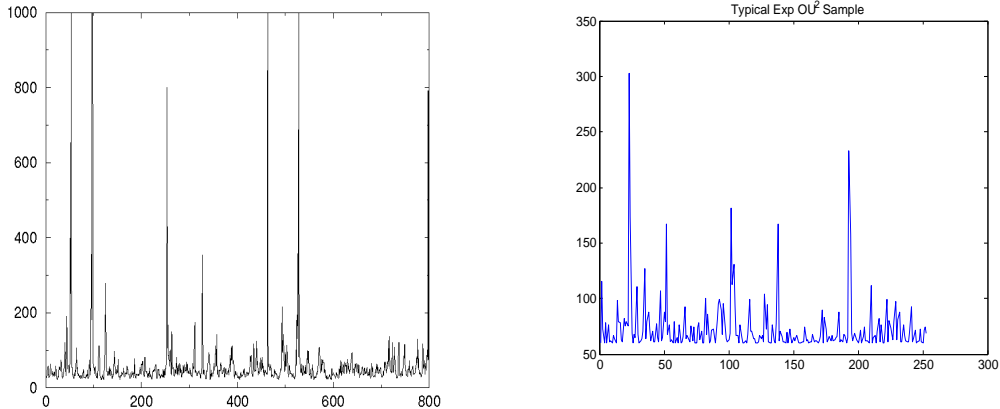
for the **non-linear** function

$$f_\alpha(x) = \begin{cases} (1 + \alpha x)^{1/\alpha}, & \alpha \neq 0 \\ e^x & \alpha = 0 \end{cases}$$

of an **Ornstein-Uhlenbeck** diffusion (representing demand)

$$dX_t = -\lambda(X_t - \bar{x})dt + \sigma dW_t$$

By varying the choice of  $\alpha$ , one can clearly vary the steepness of the supply stack. In par-



**Figure 9.** Monte Carlo sample from Barlow’s spot model (left) “cheap” alternative from the exponential of an Ornstein – Uhlenbeck squared (right).

ticular, any  $\alpha < 0$  corresponds to a function steeper than the exponential, while the special cases of  $\alpha = 0$  and  $\alpha = 1$  are linear and exponential respectively. Given the ‘spiky’ price data used to calibrate the model, Barlow finds negative values of  $\alpha$  for both Canadian and US markets. Barlow’s simple model is a natural starting point for understanding the structural approach, as demand is the one random factor and the transformation is described by a simple one-parameter function. An Ornstein-Uhlenbeck process is a common choice to capture the mean-reverting behaviour of demand, driven by temperature fluctuations. While most demand models include a deterministic seasonal function, Barlow omits this for simplicity as his data shows relatively little seasonality. Power demand typically includes deterministic components for both annual and intra-day periodicities, as well as weekly patterns to capture weekend and holiday effects.

Even in such a simple model, we can begin to see benefits of the structural approach. With an appropriate parameter  $\alpha$ , Barlow’s model can capture extreme spikes with a one-factor pure diffusion process, and without excessively large parameters  $\kappa$  or  $\sigma$  (see Figure 9 for a simulated price path). In contrast, a reduced-form one-factor jump diffusion price process might still capture the extreme spikes, but at the expense of a very high  $\kappa$ , dampening the volatility of prices at other times. On the other hand, Barlow’s approach also highlights an important challenge for structural models, namely capturing accurately the top of the bid stack function, which determines the range of spike levels attained in the market. In order to avoid unreasonably high values, Barlow suggests to cap the price at a maximum level, corresponding to the event that demand reaches maximum capacity. This is a reasonable assumption, especially as maximum bid levels exist in most markets (eg, \$3000 in ERCOT, €3000 in EEX, \$1000 in PJM). However one should be mindful of possible limitations. If the tail of the demand distribution and the shape of the stack combine to create a very thin tail for the price distribution, model simulations may reveal a rather high proportion of spikes ending up at the price cap, instead of more evenly spread below the cap.

#### 4. BUILDING BLOCKS OF STRUCTURAL MODELING

A broad range of structural models exists, ranging from Barlow’s simple approach above to complicated multi-fuel approaches, which attempt to get ever closer to the true price

setting mechanism of the power market auction, all the while retaining a certain level of mathematical elegance and tractability. In this section, we discuss the key relationships between spot prices and factors, while reviewing existing approaches in this branch of the literature, and piecing together the important components of a successful structural model for electricity.

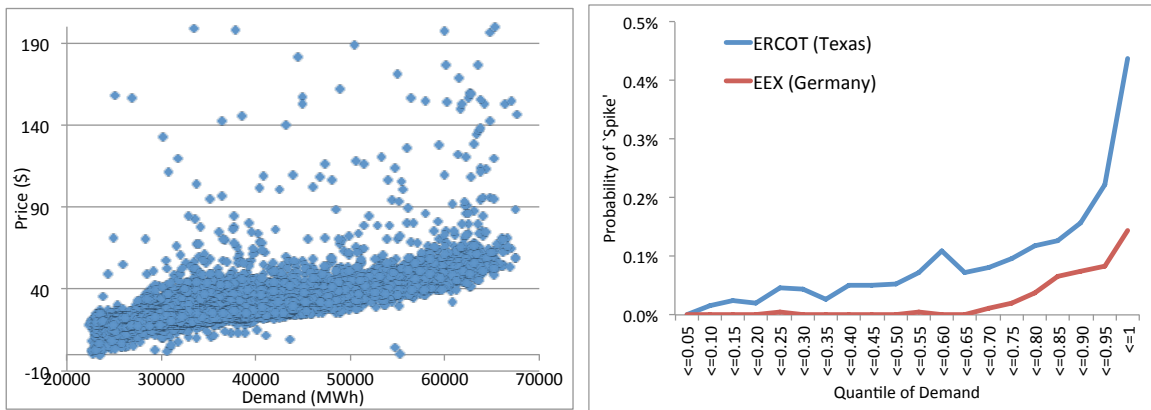
**4.1. Price Relationship with Demand.** The most striking characteristic of wholesale electricity demand is arguably its degree of price-inelasticity, perhaps unmatched among all commodity markets. As end-users typically do not feel the impact of short term price fluctuations (paying instead slow-moving retail prices), and black-outs are understandably rather frowned upon, utilities are often faced with buying last-minute power in the spot market to satisfy their obligations, no matter what the cost! Coupled with the lack of any inventories to help guard against supply or demand shocks, this extreme inelasticity of demand to price is directly responsible for the well-known and dramatic spikes in power prices. Moreover, historical price and load data provides compelling evidence for the important role of demand in driving prices both during spikes and in quieter times (as shown in Figure 10a). In most markets, detailed historical load (demand) data is readily available, and thanks to the inelasticity described above, no rough estimation is needed to produce a reasonable inverse demand curve - it's hard to go wrong with a vertical line! It is therefore not surprising that all structural models (including Barlow's described above) are built first and foremost on a process for demand, and a function to capture the link with price. This function can be described in traditional economic terms as the inverse supply curve, or in terms more specific to power markets, the *bid stack*.

The bid stack is a concept closely linked to the production stack discussed earlier, as both are driven by the merit order of fuels. The bid stack is constructed by the market administrator using daily auction data, whereby generators submit price and quantity pairs describing how much power they are willing to sell at a certain price. Thus, if the market is competitive and generators bid at or near cost, then the bid stack and production stack are very similar, and move in close tandem. (see [49] for more discussion on the relationship between the two.) Figure 8b shows sample bids from PJM for two dates in February and March 2003, between which the price of natural gas increased rapidly. Note that in reality both supply and demand side bids (sometimes called offers and bids) are submitted, but in many markets the demand side bids are predominantly made at the maximum price level (price cap) due to the inelasticity discussed above. Notable exceptions are markets (such as EEX in Europe) where only a fraction of actual load is traded on the market, implying that if market prices are low, companies may choose to buy from the market in order to satisfy off-market commitments, while switching off their regular generators. Such behaviour leads to significant demand side elasticity in bids, even if overall demand is still inelastic, due to the interplay between market and off-market dynamics. Nonetheless, the relationship between price and load can still be approximated by a bid stack approach, even if the bidding behaviour itself is more complicated.

**4.2. Price Relationship with Capacity or Margin.** Barlow's key contribution was the basic idea of a parametric relationship between  $S_t$  and an underlying demand process  $D_t$ , which can be adapted to local market conditions, for example the 'spikyness' of a given market. Another similar approach by Kanamura and Ohashi [59] proposes an alternative parametric form, with price piecewise quadratic in demand. However, while it is clear that demand is a key driver of spot prices, it is also clear that they are not perfectly correlated, as illustrated by Figure 10. The first plot shows the price to load relationship in the Texas market (ERCOT) over the year 2011, for the price interval [\$0,\$200]. This plot does not show the very high spikes, but more clearly shows the price to load dependence in the normal price region. Note that ERCOT is a particularly 'spiky' market, and that such extreme values

can be observed even for low values of load, although the probability of a spike certainly increases with demand. This is illustrated in Figure 10b, which also compares with EEX, a market with some but fewer spikes than ERCOT.

Many authors have built on Barlow’s seminal contribution, extending the tight link between price and demand to a more sophisticated model, capable of replicating the typical price-load scatter plots shown in Figure 10a. A common remedy is the inclusion of a stochastic process for the availability of generation capacity. Indeed, generator outages can be common occurrences in some markets, while seasonal maintenance patterns also serve to shift supply.



(a) Price against load (ERCOT in 2011)

(b) 'Spike' probability against load

**Figure 10.** Left plot illustrates the price to load relationship in ERCOT (all hours included) in the year 2011. Price axis is cut at \$200 but values up to \$3000 occurred. Right plot illustrates the relative frequency of a 'spike' in both ERCOT (2005-11 data) and EEX (2007-09 data). Here a spike is crudely defined as an hourly price greater than 3 times the average monthly value for the period.

In a detailed stack model, the removal of a generating unit due to an outage can be simulated by explicitly removing a particular section of the stack function, and shifting the remainder to the left (see chapter on hybrid models in [49] for further discussion of this style of model). However, calibration requires detailed market data, and it is difficult to adopt this approach while retaining a convenient mathematical function for the stack at all times. Instead, several authors (cf. [72, 36, 65]) have proposed writing prices directly as a function of both demand  $D_t$  and total market capacity  $C_t$  using an exponential form such as

$$(6) \quad S_t = \exp(aD_t + bC_t), \quad \text{where } a > 0, b < 0.$$

Skantze *et al* [72] proposed an early model of this form, with both demand  $D_t$  and 'supply'  $C_t$  driven by two-factor Ornstein-Uhlenbeck plus arithmetic Brownian motion models, with additional outage effects for  $C_t$ . In their model, the supply factor  $C_t$  is not assumed to be observed but instead calibrated as a residual of the model's fit to price and load. Cartea and Villaplana [36] also suggest the form (6), but instead estimate 'generation capacity'  $C_t$  directly either from hydro reservoir levels for the Nordpool market, or available / installed capacity data for England and Wales, and for PJM. They model  $D_t$  and  $C_t$  as correlated Ornstein-Uhlenbeck process with seasonality and seasonal volatility. As power prices are then lognormal, forward prices are easy to calculate, and the authors investigate the model's implications for risk premia in the forward curve. Option prices can also be found in this

lognormal special case, as discussed for example by Lyle and Elliot [65].

While the simplicity of equation (6) is attractive, it raises several questions.

- Firstly, is  $C_t$  really an observable variable, or simply a noise term which approximates the shifts in the stack which distort the price-load relationship? If capacity data is available, will it be enough to explain the price variations as suggested by the model? In practice, prices may spike not because of a lack of total capacity in the market, but because of difficulty in matching the capacity with the demand, due to either transmission constraints through the grid or operational constraints such as ramp-up times.
- Secondly, should decreases in  $C_t$  lead to parallel shifts in the bid stack, as suggested by (6)? If all generating units are equally likely to be removed from the stack, then the effect should be multiplicative, not additive, making power price a function of  $D_t/C_t$ , not  $D_t - C_t$ . Parallel shifts suggest that capacity is being primarily removed from the far left of the stack, and therefore not steepening the relationship with demand.
- Thirdly, should the event  $D_t \leq C_t$  be guaranteed by the model, implying that demand never exceeds available capacity? If so, how should this be achieved mathematically, as all processes for  $D_t$  mentioned above have unbounded support?

A large variety of models exist which take various approaches to the three interrelated issues raised above. Broadly speaking we can categorize structural models into two groups depending on their treatment of the supply-driven process  $C_t$ . If  $C_t$  is modelled strictly as the available capacity in the market, then the treatment of the second and third issues above is more important, as there is a clear benefit to using the ratio  $D_t/C_t$ , and ensuring that it always remains between 0 and 1, either through direct capping or something more sophisticated. We shall call this interpretation of  $C_t$  Version A. However, in practice it may be beneficial to treat  $C_t$  as an unobserved residual noise process, backed-out from prices and implicitly capturing a range of other ‘capacity-related’ effects, including outages, reserves, maintenance, market constraints, and imports or exports. In this case, which we shall call Version B, both the ordering of  $D_t$  and  $C_t$  and the distinction between parallel and non-parallel shifts in the stack is less important, suggesting that the form of (6) can suffice. Some authors have proposed models which straddle both of these categories, as in [19]. Here the authors introduce a non-parametric ‘price-load curve’  $f(t, D_t/c(t))$  to represent the inverse supply curve, where  $c(t)$  tracks the seasonal level of capacity available, driven by weather and maintenance patterns. However, they also add additional noise terms  $X_t$  and  $Y_t$ , and define

$$S_t = \exp \left\{ f \left( t, \frac{D_t}{c(t)} \right) + X_t + Y_t \right\}, \quad t = 0, 1, \dots$$

where  $X_t$  and  $Y_t$  are unobservable short term and long term factors both attributed primarily to “psychological aspects of the behaviour of speculators and other influences”. Using  $C_t$  as a noise term, or adding other unobservable factors to capture all residual effects is an effective way of ensuring that the model reflects the high volatility witnessed in power markets, without worrying about the exact source.

On the other hand, evidence suggests that the level of demand relative to available capacity is crucial. If  $C_t$  truly tracks total capacity, then times when  $D_t$  approaches  $C_t$  should intuitively be those which lead to price spikes. Some authors have given special attention to such effects by directly modeling the behaviour of the ‘reserve margin’  $C_t - D_t$  (or the percentage reserve margin  $1 - D_t/C_t$ ), emphasizing the advantage of capturing both demand and capacity movements in a single variable. Boogert and Dupont [16] analyse the relationship

between margin and spot price as well as margin and spike probability, and suggest a non-parametric approach. Cartea *et al* [35] advocate using forward-looking margin information as an indicator of when a spike is likely to occur, and defining a separate price regime when a threshold level of margin is reached. Similarly, Mount *et al* [66] and Anderson and Davison [4] propose regime-switching frameworks whereby either mean price levels or transition probabilities between regimes are allowed to depend on the margin level. In [40], Coulon *et al* make use of the exponential form in (6), but with a second exponential for a spike regime, whose probability is linear in the quantile of demand. Note that while these models can still be thought of as structural in spirit, some do not necessarily rely on the notion of a supply curve mapping demand to price, since demand (or margin) may instead be used to determine which of two or more spot price processes is most likely to apply at a given time.

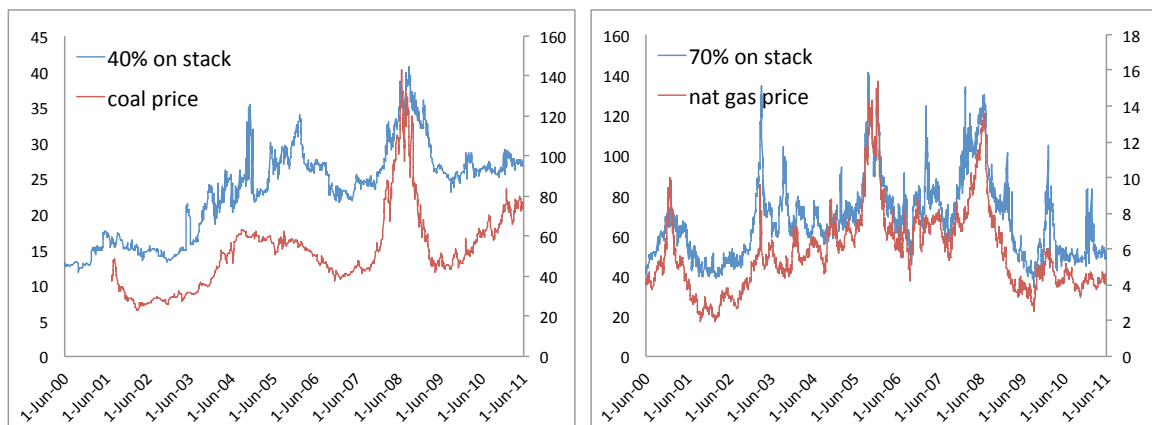
**4.3. Price Relationship with a Single Marginal Fuel.** Figure 10 confirms that while demand and capacity are very important drivers of power prices over short time horizons, Figure 2 (and 8a) shows that the long term levels of prices tend to match closely with costs of production. This is particularly striking during the period of record highs in almost all commodity prices in 2008, as discussed in detail in Section 2. Hence, any structural model to be used for medium to long-term purposes must incorporate the risk of movements in the fuel prices appropriate for that market, and preferably also the information contained in fuel forward curves. One could argue that these factors essentially inherit the role of the longer term (non-stationary) factor in the classical reduced-form model of Schwartz and Smith presented in [71]. The challenge is how to incorporate fuel price movements into supply curve movements, particularly in markets with multiple production technologies and complicated merit orders. Hydropower, renewables and nuclear all require slightly different considerations as well, since the quantity of power generated from these sources is driven not by fuel price movements but instead by resource availability (in the case of hydro and renewables) and the need to avoid shut-down costs (for nuclear).

Pirrong and Jermakyan [68, 67] stress the importance of writing power as a function of marginal fuel, and propose a useful model for a heavily gas-based market. They assume that power prices are driven by two factors, both observable: fuel prices (natural gas specifically) and demand. Demand  $D_t$  is assumed to be driven by an exponential Ornstein-Uhlenbeck process with seasonality, while gas prices  $G_t$  follow a Geometric Brownian Motion. The authors assume that the inverse supply curve is multiplicative in fuel price, meaning that  $S_t = G_t \phi(\ln D_t)$ , or more generally  $S_t = G_t^\gamma \phi(\ln D_t)$ , for  $\gamma \geq 0$ . They then suggest several methods for determining the function  $\phi(x)$  (and possibly  $\gamma$  too, noting  $\gamma = 1$  is a natural choice, for which  $\phi$  can be thought of as a ‘heat rate’ curve). One option is to use specific data on marginal costs of power production to construct the ‘generation stack’, a second option is assuming a parametric form for  $\phi$ , and a third (which they favour) is to directly use historical bid data from one year earlier, rescaled by the change in gas price. Various other authors discuss the need to include marginal fuel prices when modeling power. Eydeland and Geman [48] suggest multiplying an exponential function of demand by the marginal fuel price in the market, while Coulon *et al* [40] multiply two exponential functions (for two regimes) by natural gas price in the gas-dominated ERCOT market.

**4.4. Price Relationship with Multiple Fuels.** In some electricity markets (particularly those dominated by natural gas generators), a single fuel factor combined with demand and/or capacity effects is sufficient to describe very well the dynamics of power prices. However, in other cases, one fuel is simply not enough, leading various authors to propose models which incorporate two or more different fuel prices. In the reduced-form world, some authors have suggested modeling power and fuels as cointegrated processes (cf. [47, 18, 44, 74]), while others have suggested multi-commodity Lévy based models with various ways of capturing correlations between jumps and/or diffusion components. (cf. [45, 58, 51]). However, these



approaches fail to capture the intricate dependencies between fuel prices, demand and capacity, which lead to state-dependent correlations. For example, at times of low demand, power prices are correlated more closely with fuel prices of cheaper technologies, while at times of high demand, more expensive fuels tend to set the power price and produce a stronger correlation. This can perhaps be most easily illustrated by looking at actual bid data from PJM, as shown in Figure 11. Here we see that over more than 10 years of historical data, the overall pattern of bid movements lower in the stack (at 40% of total capacity) tends to follow trends in coal prices, while the higher portion of the PJM stack (at 70% of total capacity) has a remarkably strong link with natural gas. However, it is important to note that the relatively stable historical PJM merit order is particularly susceptible to merit order changes today, as US natural gas prices have fallen to record lows of under \$2 in 2012. An increasing number of gas generators are displacing coal generators in the stack, and impacting electricity price correlations in the process.



(a) Coal vs. 40% point on PJM stack

(b) Gas vs. 70% point on PJM stack

**Figure 11.** Illustration of correlation between bid stack dynamics and fuel prices (with left axes used for stack level, and right axes used for fuel prices, all in \$)

In the context of structural models, key modeling questions include whether to impose a strict ordering of fuel types by demand, whether to allow regions of overlap between fuels, and how to reconcile the fuel price dependence with other features such as spikes. Coulon and Howison [39] proposed an innovative approach to handling such merit order changes, constructing the stack by approximating the distribution of the clusters of bids from each technology. Hence they write the bid stack as the inverse cumulative distribution function of a mixture distribution for bids, and model demand and margin as correlated exponential Ornstein-Uhlenbeck processes, with jumps in margin added. In this approach, all regions of the bid stack are technically driven by all fuels (since the bid clusters have unbounded support), but to very different degrees at times when the cluster means are far apart. In the work of Aïd *et al* [1], the authors simplify the stack construction by assuming only one bid price per fuel type, corresponding to a constant heat rate per technology. Hence there is no region of overlap between fuels, and the marginal fuel type changes at a series of demand thresholds corresponding to capacity per technology. This provides much more convenient formulas for pricing derivatives, but at the expense of a major oversimplification of spot price dynamics. In an extension of the earlier model, Aïd *et al* [2] extended this approach to improve spot price dynamics and capture spikes, by multiplying a ‘scarcity function’ (of margin) by the heat rate and fuel price of the marginal fuel. They choose this function to be a power law of the reserve margin (with a cap), arguing this to be more effective than

the common choice of exponential. While the choice of marginal fuel is still determined by demand, they sacrifice the possibility of merit order changes, by assuming the ordering of fuels is fixed initially, arguing that this is reasonable over short horizons. Carmona *et al* [21] propose instead a framework based on different exponential bid curves for each fuel (corresponding intuitively to the range of heat rates per technology), and combine these precisely as the merit order dictates to produce a piecewise exponential function for the market as a whole. Hence demand thresholds exist where marginal fuel type(s) changes, but these are highly dynamic, with overlap regions appearing and disappearing in the stack, and the merit order changing as fuel prices move. The model is particularly convenient for the two-fuel case as closed-form expressions exist for forwards and spread options. However, it is fair to say that for three or more fuels the calculations become unmanageable. In the following section, we will present a broad multi-fuel structural framework which builds perhaps most closely on the last of these models, but draws on ideas from all existing work discussed here. Note that for simplicity we do not include carbon emissions prices into our structural framework here, and instead refer the interested reader to [22, 42, 56] for stack-based models which include carbon emissions prices as additional production costs, typically in conjunction with multiple fuel types with different emissions rates.

## 5. FORWARD PRICING IN A STRUCTURAL APPROACH

While reduced-form models are often designed specifically to facilitate derivative pricing (including those mentioned in Section 3.3), structural models often face a choice between staying true to the market's structure and cutting some corners to price forwards or options efficiently. Ideally, a model should capture the structural relationships accurately while retaining convenient expressions for derivatives, but in some markets this may be very challenging indeed. As forward contracts are by far the most widely and liquidly traded contracts, allowing for rapid calibration to the observed forward curve is typically a top priority for any model, while convenient option pricing results are a welcome bonus but secondary concern. In this section, we discuss the challenges of derivative pricing in structural models, and suggest general frameworks which allow for the explicit calculation of at least forward prices. (In some special cases, options and other derivatives can be priced, as discussed for power plant valuation in [21], but we do not investigate this here.)

As several authors have discussed (cf. [21, 1, 39]), one advantage of using structural price models for pricing forwards is to capture the dependence of electricity forwards on fuel forwards in a manner which is consistent with their stack-based relationship in the spot market. Another advantage is to capture forward-looking information about upcoming market changes, such as a changing generation mix (e.g. increased renewables, technological developments, the nuclear moratorium in Germany, etc.), or the introduction of new regulation (e.g. emissions markets, market coupling). Finally, one might wish to include a view on load growth or upcoming maintenance schedules. No matter the motivation, the core goal here is to choose a flexible and realistic functional relationship between price and its underlying drivers, such that expectations of future spot prices can be explicitly calculated, and hence forward prices found.

Let  $F^p(t, T)$  denote the forward price of electricity at time  $t$  for delivery at time  $T$ . Recall that  $F^p(t, T) = \mathbb{E}_t[S_T]$  where  $\mathbb{E}_t[\cdot]$  denotes the conditional expectation given time  $t$  information with respect to a risk neutral probability measure. (We assume this measure is given, and relegate a discussion of risk premia to Section 5.3.) The same equation holds for fuel prices, so for example we write  $F^g(t, T) = \mathbb{E}_t[S_T^g]$  for the forward price of gas.

Our aim is to build a general framework that draws on techniques introduced in a number of different papers, highlighting the assumptions needed to provide closed-form expressions for forward prices. We will make use of the following ingredients:

- *Lognormal fuel spot prices* - This assumption is a very common and natural choice for modeling energy (non-power) prices. Geometric Brownian Motion (GBM) with constant convenience yield, the classical exponential Ornstein-Uhlenbeck model of Schwartz [70], and its two-factor extensions [71, 55] all satisfy the log-normality assumption, as does the general forward curve model given in (2).
- *Power price multiplicative in marginal fuel* - This assumption is made by many authors (cf. [48, 68, 21, 2] among others), and reflects the fact that fuel costs are the dominant drivers of power bids, and large compared to other operational costs. The power price can be thought of as a product of the marginal fuel cost and a ‘heat rate function’, describing the heat rate of the marginal generator at the appropriate demand level. The more possible marginal fuels in a market, the greater the challenge in building a structural model!
- *Gaussian demand (and possibly ‘capacity’)* - Power demand is often assumed to be Gaussian and modelled as an Ornstein-Uhlenbeck process (cf. [7, 36, 65, 59] among others). This is consistent with the (piecewise) linear dependence we highlighted in the discussion surrounding Figure 7, and the fact that temperature is reasonably well modelled by a Gaussian autoregressive model with strong seasonal components. Depending on the role and treatment of capacity, the process may need to be strictly capped at the top and bottom of the stack, or alternatively an additional Gaussian noise term may be added to represent capacity changes.
- *Exponential heat rate functions* - The relationship between price and load is typically convex and often modelled with an exponential function, as discussed in Subsection 4.1. Coupled with Gaussian demand, this set-up can provide convenient flexibility to produce specialized results.
- *Multiple spot price regimes* - The assumptions listed above are typically not sufficient to capture the heavy-tailed nature of spot prices, with both positive and negative spikes possible. Various authors have proposed regime-switching models (cf. [43, 79, 66, 4] among others) to handle this primary feature of power prices. While some of these are pure reduced-form approaches, others merge regime-switching with a structural framework.

**5.1. Single Fuel Markets.** We begin with a single fuel model, suitable for markets in which the marginal generator is almost always of the same fuel type. Note that generators which always bid at very low price levels (or simply at zero) can be incorporated into this framework most easily by replacing the demand process  $D_t$  by the residual demand process after subtracting their capacity. This is particularly relevant for generation types such as nuclear and renewables. Note that this does not mean that these generators are simply ignored, as the adjustment to model residual demand may require some care, as discussed for wind and solar power in Germany in [76]. For example the volatility of wind availability may mean that the residual demand distribution has a significantly higher volatility than the original demand distribution. Hence, while it is assumed these units don’t set the power price, they may well influence the power price. For simplicity, we shall call the unique marginal fuel natural gas, with spot price denoted  $S_t^g$ . We now divide our framework into two types of models, which differ in their treatment of capacity  $C_t$ . Version A will treat  $C_t$  as strictly the available generation capacity in the market, while Version B will treat  $C_t$  more loosely as a stochastic perturbation driven by capacity changes.

Throughout, we consider one particular maturity of interest,  $T$ , and specify the conditional distributions (given time  $t$  information) of gas price as lognormal, and both demand and capacity as normal and possibly correlated. We assume the fuel price to be independent of demand and capacity. This is a reasonable assumption as power demand is typically driven predominantly by temperature, which fluctuates at a faster time scale and depends more on local or regional conditions than fuel prices. In summary, the random factors determining  $S_T$  are

$$(7) \quad \begin{bmatrix} D_T | D_t \\ C_T | C_t \end{bmatrix} \sim N \left( \begin{bmatrix} \mu_d \\ \mu_c \end{bmatrix}, \begin{bmatrix} \sigma_d^2 & \rho\sigma_d\sigma_c \\ \rho\sigma_d\sigma_c & \sigma_c^2 \end{bmatrix} \right), \quad \log S_T^g | S_t^g \sim N(\mu_g, \sigma_g^2)$$

Indeed, we stress that we are interested in making minimal assumptions on the behaviour of these factors, as different markets may have different characteristics, and different authors may favour different Gaussian processes, seasonal patterns and other variations. Our emphasis is instead primarily on the stack-based mapping to power prices.

### Useful Notation and Results:

The calculation of forward prices throughout this section relies on the computation of various integrals over the multivariate Gaussian density, and thus repeatedly makes use of the following standard result:

$$(8) \quad \int_{-\infty}^h e^{cx} \Phi \left( \frac{a+bx}{d} \right) \frac{e^{-\frac{1}{2}x^2}}{\sqrt{2\pi}} dx = e^{\frac{1}{2}c^2} \Phi_2 \left( h - c, \frac{a+bc}{\sqrt{b^2+d^2}}; \frac{-b}{\sqrt{b^2+d^2}} \right)$$

where  $a, b, c, d, h$  are constants (with  $h = \infty$  in some cases), and  $\Phi(\cdot)$  and  $\Phi_2(\cdot, \cdot; \rho)$  the cumulative distribution functions of the univariate and bivariate (correlation  $\rho$ ) standard (i.e. mean zero and variance one) Gaussian distributions respectively. Note that the constant  $d$  is redundant in the expression above, but in practice it is convenient to use this form.

In addition, in some multi-fuel cases, we may require integrating over a bivariate Gaussian distribution function, in which case the following related result is used:

$$(9) \quad \begin{aligned} & \int_{-\infty}^h e^{cx} \Phi_2 \left( \frac{a_1+b_1x}{d_1}, \frac{a_2+b_2x}{d_2}; \lambda \right) \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} dx \\ &= e^{\frac{1}{2}c^2} \Phi_3 \left( h - c, \frac{a_1+b_1c}{\sqrt{b_1^2+d_1^2}}, \frac{a_2+b_2c}{\sqrt{b_2^2+d_2^2}}; \begin{bmatrix} 1 & \rho_{12} & \rho_{13} \\ \rho_{12} & 1 & \rho_{23} \\ \rho_{13} & \rho_{23} & 1 \end{bmatrix} \right) \end{aligned}$$

where  $a_1, a_2, b_1, b_2, c, d_1, d_2, h, \lambda$  are constants (with  $h = \infty$  in some cases),  $\Phi_3(\cdot, \cdot, \cdot; \Sigma)$  is the standard trivariate Gaussian cumulative distribution function with correlation matrix  $\Sigma$ , and

$$\rho_{12} = \frac{-b_1}{\sqrt{b_1^2+d_1^2}}, \quad \rho_{13} = \frac{-b_2}{\sqrt{b_2^2+d_2^2}}, \quad \rho_{23} = \frac{b_1b_2 + \lambda d_1d_2}{\sqrt{(b_1^2+d_1^2)(b_2^2+d_2^2)}}.$$

Finally, given the frequency of integrating between two finite limits and obtaining a difference between Gaussian cumulative distribution functions, we introduce the following useful shorthand notation:

$$\Phi_2 \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y; \rho \right) := \Phi_2(x_1, y; \rho) - \Phi_2(x_2, y; \rho).$$

5.1.1. *Version A:* In a model for which  $C_t$  is strictly the maximum capacity in the market, we require  $0 \leq D_t \leq C_t$ , and define a functional form for the bid stack over this range. However, allowing for the possibility of multiple, say  $N$ , price regimes (e.g., a normal regime, spike regime, negative price regime), we define multiple functional forms and attach probabilities  $p_i$  (for  $i = 1, \dots, N$ ) to being in each regime. As evidence suggests that the likelihood of a

spike is load-dependent but spikes do occasionally occur even for low load (see Figure 10), we allow also for load dependent probabilities  $p_i(D_t)$  as suggested in [40]. Since electricity spot prices are discrete time processes (typically taking 24 values per day, one for each hour, and with relatively weak links between neighboring hours due to non-storability), we do not necessarily need to define a continuous time Markov chain to drive transitions between regimes. However, for modeling purposes, one can choose to interpret electricity prices as the discrete time observation of a hidden continuous time process (see [13] for more discussion) in which case the probabilities here could be the result of a rapidly moving continuous time Markov Chain, which approximately reaches its stationary distribution in less than an hour. More generally, we could also allow for  $p_i$  to depend on the current regime (e.g. for spikes which last several hours) but we do not consider this complication here, as capturing the timing of clusters of spike values is not our priority. In our current model, the spot power price  $S_t$  at any  $t$  is given by

$$(10) \quad S_t = (-1)^{\tilde{\delta}_i} (S_t^g)^{\delta_i} \exp\left(\alpha_i + \beta_i \hat{D}_t\right) \quad \text{with probability } p_i(D_t)$$

where  $\hat{D}_t = \max(0, \min(C_t, D_t))$  is capped demand. The parameters  $\delta_i, \tilde{\delta}_i \in \{0, 1\}$  allow for switching on and off fuel price dependence and negative prices respectively, and

$$(11) \quad p_i(D_t) = \underline{p}_i + \bar{p}_i \Phi\left(\frac{D_t - \mu_d}{\sigma_d} (-1)^{\tilde{\delta}_i}\right) \quad \text{for } i = 1, \dots, N-1,$$

$$\text{and } p_N(D_t) = 1 - \sum_{i=1}^{N-1} p_i(D_t).$$

For the  $N$ -th regime (with is most intuitively thought of as the ‘normal’ regime when no extreme events occur), we can write its probability in the same form as all other regimes:

$$(12) \quad p_N(D_t) = \underline{p}_N + \bar{p}_N \Phi\left(\frac{D_t - \mu_d}{\sigma_d} (-1)^{\tilde{\delta}_N}\right)$$

where, defining the sets  $I = \{i : \tilde{\delta}_i = 0\}$  and  $J = \{1, \dots, N-1\} \setminus I$ , we have

$$\underline{p}_N = 1 - \sum_{i=1}^{N-1} \underline{p}_i - \sum_{i \in J} \bar{p}_i$$

$$\bar{p}_N = \sum_{i \in J} \bar{p}_i - \sum_{i \in I} \bar{p}_i$$

$$\tilde{\delta}_N = 0$$

We require  $p_i(D_t) \in (0, 1)$  for all  $D_t$  and all  $i = 1, \dots, N$ , which is guaranteed if  $\sum_{i=1}^{N-1} (\underline{p}_i + \bar{p}_i) < 1$ . Note that the probabilities  $p_i(D_t)$  are chosen to be linear functions of the *quantile* of the demand at time  $t$  (or of the quantile of  $-D_t$  if  $\tilde{\delta}_i = 1$ ). Intuitively, for a ‘spike’ regime with  $\tilde{\delta}_i = 0$  (and relatively high  $\alpha_i$  and/or  $\beta_i$ ), the likelihood of being in such a regime increases gradually with load, from  $\underline{p}_i$  up to a maximum of  $\underline{p}_i + \bar{p}_i$ . On the other hand, for a negative price regime with  $\tilde{\delta}_i = 1$ , the likelihood decreases steadily with load. The use of the quantile of load is both convenient mathematically and supported by empirical evidence (see Figure 10b), although for some markets a piecewise linear function of the quantile seems more appropriate (and still leads to closed-form formulas, just a little messier!). Finally, note that we expect  $\delta_i = 1$  for the ‘normal’ regime(s) where the price is most typically set, but may prefer to set  $\delta_i = 0$  for other regimes such as the negative price regime since the size of a downwards spike is unlikely to depend on the current gas price.

Consider first the case that generation capacity  $C_T = \bar{\xi}$  is constant (or known in advance), so that  $\sigma_c = 0$ . The forward price for time  $T$  delivery in this case is given by:

$$\begin{aligned} F^p(t, T) &= \mathbb{E}_t[S_T] \\ &= \mathbb{E}_t[\mathbb{E}_t[S_T|D_T]] \\ &= \mathbb{E}_t \left[ \sum_{i=1}^N (-1)^{\tilde{\delta}_i} \mathbb{E}_t[(S_T^g)^{\delta_i}] \exp(\alpha_i + \beta_i \max(0, \min(\bar{\xi}, D_T))) p_i(D_T) \right] \\ &= \sum_{i=1}^N (-1)^{\tilde{\delta}_i} (F^g(t, T))^{\delta_i} \mathbb{E}_t \left[ p_i(D_T) \left( e^{\alpha_i} \mathbb{I}_{D_t \leq 0} + e^{\alpha_i + \beta_i D_t} \mathbb{I}_{0 \leq D_t \leq \bar{\xi}} + e^{\alpha_i + \beta_i \bar{\xi}} \mathbb{I}_{D_t \geq \bar{\xi}} \right) \right] \end{aligned}$$

Given the form of  $p_i(D_T)$  in (11), the approach of first conditioning on  $D_T$ , allows us to use (8) for each term above. We obtain

$$(13) \quad F^p(t, T) = \sum_{i=1}^N (-1)^{\tilde{\delta}_i} (F^g(t, T))^{\delta_i} f(\mu_d, \sigma_d, \bar{\xi}, \underline{p}_i, \bar{p}_i, \alpha_i, \beta_i, \tilde{\delta}_i)$$

where the function  $f(\mu_d, \sigma_d, \bar{\xi}, \underline{p}_i, \bar{p}_i, \alpha_i, \beta_i, \tilde{\delta}_i)$  is given by

$$(14) \quad f(\mu_d, \sigma_d, \bar{\xi}, \underline{p}_i, \bar{p}_i, \alpha_i, \beta_i, \tilde{\delta}_i) = \left\{ e^{\alpha_i} \left( \underline{p}_i \Phi \left( \frac{-\mu_d}{\sigma_d} \right) + \bar{p}_i \Phi_2 \left( \frac{-\mu_d}{\sigma_d}, 0; \frac{-1}{\sqrt{2}} \right) \right) \right. \\ \left. + e^{\alpha_i + \beta_i \mu_d} \left( \underline{p}_i \Phi \left( \left[ \begin{array}{c} (\bar{\xi} - \mu_d - \beta_i \sigma_d^2)/\sigma_d \\ (-\mu_d - \beta_i \sigma_d^2)/\sigma_d \end{array} \right] \right) + \bar{p}_i \Phi_2 \left( \left[ \begin{array}{c} (\bar{\xi} - \mu_d - \beta_i \sigma_d^2)/\sigma_d \\ (-\mu_d - \beta_i \sigma_d^2)/\sigma_d \end{array} \right], \frac{\beta_i \sigma_d (-1)^{\tilde{\delta}_i}}{\sqrt{2}}; -\frac{(-1)^{\tilde{\delta}_i}}{\sqrt{2}} \right) \right) \right. \\ \left. + e^{\alpha_i + \beta_i \bar{\xi}} \left( \underline{p}_i \Phi \left( -\frac{\bar{\xi} - \mu_d}{\sigma_d} \right) + \bar{p}_i \Phi_2 \left( -\frac{\bar{\xi} - \mu_d}{\sigma_d}, 0; -\frac{(-1)^{\tilde{\delta}_i}}{\sqrt{2}} \right) \right) \right\}.$$

While the expression above may appear involved, this is only because of the truncation of demand. The terms are readily identifiable, as the first line corresponds to the event of hitting the bottom of the stack ( $D_t \leq 0$ ) for each regime, second line the middle of the stack ( $D_t \in (0, \bar{\xi})$ ) and third line the top ( $D_t \geq \bar{\xi}$ ). Typically we would expect parameters  $\mu_d, \sigma_d$  to be such that the first and third lines play very little role, but are of course still necessary.

## Treatment of Capacity

For fixed capacity  $C_T = \bar{\xi}$ , the version of the model presented above captures several of the key structural relationships discussed in Section 3, and allows for load-dependent spikes, both upwards and downwards. However, without any randomness in  $C_T$ , it may not be able to reproduce the high intra-day price volatility often observed in electricity markets, as both  $D_T$  and  $G_T$  are relatively slow moving from hour to hour. Moreover, maintenance schedules often lead to seasonal patterns in available capacity. Adding time-dependence and randomness to  $C_T$  is a natural remedy, but unfortunately not necessarily an easy one. In particular, in (10) in its current form, a decrease in capacity can only lower the spot price  $S_t$ , since the price will be capped at a lower level when  $D_t = C_t$ . In other words, all the capacity is removed from the top of the bid stack, causing it to end at a lower level. Several alternative formulations are possible:

- *Deterministic Capacity:* Firstly, if we are interested primarily in capturing deterministic changes in capacity (e.g. maintenance schedules), then we choose a deterministic function  $c(t)$  representing the percentage of installed capacity  $\bar{\xi}$  available at time  $t$ . Next, we assume that capacity is removed evenly throughout the stack. In other words, the range of market heat rates implied by the model should remain fixed. Hence set  $\beta_i(t) = \bar{\xi}/c(t)$ , such that the time dependence

in  $\beta_i(t)$  exactly offsets  $c(t)$ , ensuring that (in regime  $i$ ) the highest price set is  $S_t^g \exp(\alpha_i + \beta_i \bar{\xi})$ , for any value of  $c(t)$ .

- *Demand Over Capacity:* The approach above is equivalent to writing the stack as a function of  $D_t/C(t)$  directly (as suggested for example in [19, 39]). Extending this idea, one could think of modeling  $D_t/C_t$  directly as a Gaussian process, without disentangling the role of demand and capacity changes. We then still have only two random variables (including gas), but are likely to have a more volatile demand process, as it incorporates additional supply-related uncertainty. It may be convenient to treat  $D_t$  and  $C_t$  jointly if we want the dynamics of the process to be specified to match observed forward or option prices, as we shall discuss briefly in Section 5.3.
- *Stochastic Capacity:* Finally, we note that the full version of the model with Gaussian  $C_T$ , correlated with  $D_T$  (as in (7)), is also possible. We suggest in this case a slight modification to the expression (10), replacing capped demand  $\hat{D}_t = \max(0, \min(C_t, D_t))$  with capped margin  $\hat{M}_t = C_t - \hat{D}_t$ . In this case  $\exp(\alpha_i)$  needs to be interpreted as the highest, not lowest, heat rate, and  $\beta_i < 0$ . Effectively, an outage (a decrease in  $C_t$ ) then removes capacity from the bottom of the stack instead of the top. However, the resulting expression for forward prices  $F^p(t, T)$  is significantly more complicated, both because of the additional random variable and because of the need for additional caps or floors on this variable (e.g. at  $C_t = 0$ ), producing additional terms.

5.1.2. *Version B:* If we choose instead to treat  $C_t$  as an additional noise term which moves the bid stack left or right in parallel shifts, but is not interpreted strictly as the maximum capacity available, then we avoid many of the complications discussed above. In particular, we do not impose the restriction that  $D_t = 0$  and  $D_t = C_t$  correspond to the lowest and highest power prices possible (for a given fuel price), and hence we do not introduce a capped demand process  $\hat{D}_t$ . In all other respects, the model retains the features of Version A. We define the power spot price by

$$(15) \quad S_t = (-1)^{\tilde{\delta}_i} (S_t^g)^{\delta_i} \exp(\alpha_i + \beta_i D_t - \gamma_i C_t) \quad \text{with probability } p_i(D_t, C_t)$$

where

$$(16) \quad \begin{aligned} p_i(D_t, C_t) &= \underline{p}_i + \bar{p}_i \Phi((\zeta_i + \eta_i D_T + \theta_i C_t)) \quad \text{for } i = 1, \dots, N-1, \\ \text{and } p_N(D_t, C_t) &= 1 - \sum_{i=1}^{N-1} p_i(D_t, C_t). \end{aligned}$$

Notice that this time we let the regime probabilities be more general than in (11), allowing dependence on both  $D_t$  and  $C_t$ . However, in practice, we might prefer to return to the earlier special case where  $p_i$  is linear in the quantile of demand (as in [40]) by simply setting

$$(17) \quad \zeta_i = -\frac{\mu_d}{\sigma_d}, \quad \eta_i = \frac{1}{\sigma_d}, \quad \theta_i = 0.$$

On the other hand, if we prefer a linear function of the quantile of capacity (with  $p_i$  decreasing in  $C_t$ , say) then we could set

$$\zeta_i = \frac{\mu_c}{\sigma_c}, \quad \eta_i = 0, \quad \theta_i = -\frac{1}{\sigma_c}.$$

Assuming the stack model in (15) and (16), along with distributions given by (7) and  $p_i$  by (17), the forward power price for time  $T$  delivery can be found (again using (8)) to be

$$F^p(t, T) = \sum_{i=1}^N (-1)^{\tilde{\delta}_i} h_i (F^g(t, T))^{\delta_i} \left[ p_i + \bar{p}_i \Phi \left( \frac{(-1)^{\tilde{\delta}_i} (\beta_i \sigma_d - \rho \gamma_i \sigma_c)}{\sqrt{(1 - \rho^2)(2 - \rho^2)}} \right) \right]$$

where the constant  $h_i = \mathbb{E}_t [\exp(\alpha_i + \beta_i D_t - \gamma_i C_t)]$  is given by

$$(18) \quad h_i = \frac{1}{\sqrt{1 - \rho^2}} \exp \left( \alpha_i + \beta_i \mu_d - \gamma_i \mu_c + \frac{\beta_i^2 \sigma_d^2 - 2\rho \beta_i \gamma_i \sigma_d \sigma_c + \gamma_i^2 \sigma_c^2}{2(1 - \rho^2)} \right).$$

The general framework introduced above essentially models the electricity price as a mixture of lognormal random variables (and/or the negative of a lognormal when  $\tilde{\delta}_i = 1$ ), with mixing probabilities which can be state dependent. Given the distributions in (7), this characterization is accurate for Version B of the framework, and approximate for Version A where demand is truncated at the top and bottom of the stack. Recall that in practice we are likely to have only two or three regimes at most, corresponding to ‘normal prices’, unusually high prices, and possibly unusually low or negative prices. However, the very general framework above allows for the possible subdivision of spikes into low, medium or high spikes, as is sometimes suggested (cf. [41]). If we were instead in the case  $N = 1$ , (e.g. in a spike-free market), then in Version B the spot price  $S_t$  becomes lognormal and we return to the special case of some early models discussed in Section 4, as in (6). Although the multiple regimes depart somewhat from the strictest definition of a bid stack model, it is well-known that during times of extreme market stress, the price can be set at levels which depart wildly from the typical stack prediction, and hence we argue that allowing for multiple exponential curves is well-justified as a form of hybrid structural approach.

**5.2. Multi-Fuel Markets.** In many electricity markets, two or more fuel types may be present and set the power price at different times, depending both on demand and the relative prices of the fuels. In particular, the ‘merit order’ determines the sequence in which different fuels become marginal as demand increases. While an easy concept to explain and understand, this provides a big challenge for structural models, particularly in markets driven by several correlated fuels which can overlap and also swap places in the bid stack. As discussed in Section 3, only a few existing papers fully address the multi-fuel case via a structural approach.

In this section, we build on the framework introduced for a single fuel market above, and again price power forwards for a given maturity  $T$ , where the distributions of the underlying factors are lognormal or normal. We now include  $n$  correlated fuel spot prices  $S_t^1, S_t^2, \dots, S_t^n$ :

$$(19) \quad \begin{bmatrix} D_T | D_t \\ C_T | C_t \end{bmatrix} \sim N \left( \begin{bmatrix} \mu_d \\ \mu_c \end{bmatrix}, \begin{bmatrix} \sigma_d^2 & \rho \sigma_d \sigma_c \\ \rho \sigma_d \sigma_c & \sigma_c^2 \end{bmatrix} \right), \quad \begin{bmatrix} \log S_T^1 | S_t^1 \\ \vdots \\ \log S_T^n | S_t^n \end{bmatrix} \sim N \left( \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_n \end{bmatrix}, \Sigma^{(S)} \right)$$

where  $\Sigma^{(S)}$  is the covariance matrix with  $j, k$  entry  $\Sigma_{j,k}^{(S)} = \rho_{jk} \sigma_j \sigma_k$  corresponding to covariance between fuels  $j$  and  $k$ . As before, we split our approach into Version A and Version B depending on the treatment of capacity and the capping of demand.

**5.2.1. Version A.** In the single fuel case, Version A provided us with relatively few advantages over Version B, except perhaps a clearer intuition regarding the meaning of  $C_t$ , and possible avoidance of unrealistic prices thanks to bounded demand. In contrast, in the multi-fuel setting, treating capacity truly in terms of installed or available quantity gives us a natural way to capture the relative chance of each technology being marginal. Hence let  $\xi^1, \dots, \xi^n$  represent available capacity from fuel types  $1, \dots, n$ , and  $\bar{\xi} = \sum_{j=1}^n \xi^j$ . We assume these are



known with certainty, and hence set  $C_T = \bar{\xi}$  (so  $\sigma_c = 0$ ). As discussed in the single fuel case, stochastic capacity greatly increases the complexity of the computation of forward prices in Version A, and thus alternatives such as deterministic capacity trends and the treatment of  $D_t/C_t$  directly as a single random factor are advisable. The priority in the multi-fuel case is typically the relationship between the various energy prices.

We aim to build on the model developed by Carmona *et al* [21], by including multiple regimes with demand-dependent probabilities. Note that in [21], spikes and negative prices are incorporated as well, but only at the top and bottom of the stack, thus triggered by the events  $D_t \leq 0$  and  $D_t \geq \bar{\xi}$ . Instead, here we allow for the possibility of spikes even for lower levels of demand, as can sometimes occur in practice. Furthermore, via the regime-switching set-up, we obtain closed-form forward prices for a market with more than two fuels, by considering the interaction of bids from only two fuel types *within* each regime. First, for each fuel type  $j = 1, \dots, n$ , we define a fuel bid curve as a function of  $D_t$  and  $S_t^j$  with the usual form:

$$(20) \quad b_i(D_t, S_t^j) := S_t^j \exp(\alpha_j + \beta_j D_t), \quad \text{for } D_t \in [0, \xi^j]$$

Note that parameters  $\alpha_1, \alpha_2, \dots$  and  $\beta_1, \beta_2, \dots$  now correspond to fuels, not regimes. These will be used in the  $N_1$  ‘normal’ spot price regime(s), where the usual merit order rules will apply to combine the fuel bid curves, producing a piecewise exponential function as in [21]. For each regime  $i \in \{1, \dots, N_1\}$ , the spot price is driven by two fuels (say  $i_+, i_- \in \{1, \dots, n\}$ ). Let  $\tilde{D}_t^i$  represent capped demand renormalized to the capacities of regime  $i$  fuels. Hence for  $i \in \{1, \dots, N_1\}$ , set

$$\tilde{D}_t^i = \left( \frac{\xi_{i_+} + \xi_{i_-}}{\bar{\xi}} \right) \hat{D}_t, \quad \text{and} \quad \tilde{\mu}_d^i = \left( \frac{\xi_{i_+} + \xi_{i_-}}{\bar{\xi}} \right) \mu_d, \quad \tilde{\sigma}_d^i = \left( \frac{\xi_{i_+} + \xi_{i_-}}{\bar{\xi}} \right) \sigma_d$$

Then we define the spot price (for regime  $i \in \{1, \dots, N_1\}$ ) by

$$S_t = \begin{cases} S_t^{i_+} \exp(\alpha_{i_+} + \beta_{i_+} \tilde{D}_t^i) & \text{if } b_{i_+}(\tilde{D}_t^i, S_t^{i_+}) \leq b_{i_-}(0, S_t^{i_-}) \\ S_t^{i_-} \exp(\alpha_{i_-} + \beta_{i_-} \tilde{D}_t^i) & \text{if } b_{i_-}(\tilde{D}_t^i, S_t^{i_-}) \leq b_{i_+}(0, S_t^{i_+}) \\ S_t^{i_+} \exp(\alpha_{i_+} + \beta_{i_+} (\tilde{D}_t^i - \tilde{\xi}^{i_-})) & \text{if } b_{i_+}(\tilde{D}_t^i - \tilde{\xi}^{i_-}, S_t^{i_+}) > b_{i_-}(\tilde{\xi}^{i_-}, S_t^{i_-}) \\ S_t^{i_-} \exp(\alpha_{i_-} + \beta_{i_-} (\tilde{D}_t^i - \tilde{\xi}^{i_+})) & \text{if } b_{i_-}(\tilde{D}_t^i - \tilde{\xi}^{i_+}, S_t^{i_-}) > b_{i_+}(\tilde{\xi}^{i_+}, S_t^{i_+}) \\ b_{i_+, i_-}(\tilde{D}_t^i, S_t^{i_+}, S_t^{i_-}) & \text{otherwise,} \end{cases}$$

where  $b_{i_+, i_-}(\tilde{D}_t^i, S_t^{i_+}, S_t^{i_-}) = (S_t^{i_+})^{\gamma_0^+} (S_t^{i_-})^{\gamma_0^-} \exp(\gamma_1 + \gamma_2 \tilde{D}_t^i)$  and

$$\gamma_0^+ = \frac{\beta_{i_-}}{\beta_{i_+} + \beta_{i_-}}, \quad \gamma_0^- = 1 - \gamma_0^+, \quad \gamma_1 = \frac{\alpha_{i_+} \beta_{i_-} + \alpha_{i_-} \beta_{i_+}}{\beta_{i_+} + \beta_{i_-}}, \quad \gamma_2 = \frac{\beta_{i_+} \beta_{i_-}}{\beta_{i_+} + \beta_{i_-}}$$

and with probability  $p_i(D_t)$  (for regime  $i \in \{1, \dots, N_1\}$ ) as given in (11). Note that the five cases above have a straightforward interpretation as follows: only one fuel is marginal and the other unused since  $D_t$  is low (cases 1-2), only one fuel is marginal and the other is used to capacity since  $D_t$  is high (cases 3-4), or both fuels are jointly marginal (case 5).

In addition, we define  $N_2$  ‘spike’ regimes (including negative spikes), where the price will be set by a single exponential function of demand, with the choice of fuel price dependence (all or none) and negative prices similarly to earlier. For regimes  $i = N_1 + 1, \dots, N_1 + N_2$ ,

$$S_t = (-1)^{\tilde{\delta}_i} \prod_{j=1}^n \left( S_t^j \exp(\alpha_j) \right)^{\tilde{\delta}_i} \exp(\alpha_i + \beta_i \hat{D}_t)$$

again with probability  $p_i(D_t)$  as given in (11).

We let  $J_i = \{i_+, i_-\}$  represent the set of two fuels (with prices  $\mathbf{S}_t^{J_i} = (S_t^{i_+}, S_t^{i_-})$ ) driving power prices in ‘normal’ regime  $i$ , for  $i \in \{1, \dots, N_1\}$  and hence use notation  $b_{J_i}(D_t, \mathbf{S}_t^{J_i})$  for the case of jointly margin fuels (case 5 above). Without loss of generality, we assign  $i_+$  and  $i_-$  such that  $\xi^{i_+} \geq \xi^{i_-}$ . Finally, as all terms have approximately the same form, we introduce one more piece of useful notation:

$$\Phi_3^{p_i} \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y, z; \rho \right) := p_i \Phi_2 \left( \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, y; \rho \right) + \bar{p}_i \left( \Phi_3(x_1, y, z_2; \Sigma) - \Phi_3(x_2, y, z_2; \Sigma) \right)$$

where

$$\Sigma = \begin{pmatrix} 1 & \rho & -1/\sqrt{2} \\ \rho & 1 & -\rho/\sqrt{2} \\ -1/\sqrt{2} & -\rho/\sqrt{2} & 1 \end{pmatrix}.$$

Then (using (9)) the forward power price  $F_t^p$  for time  $T$  delivery can be written as follows in terms of the forward fuel prices  $F_t^1, \dots, F_t^n$  (notation shortened for convenience):

$$\begin{aligned} F_t^p &= \sum_{i=1}^{N_1} \left\{ \sum_{j \in J_i} e^{\frac{\beta_j^2 (\tilde{\sigma}_d^i)^2}{2}} \left\{ b_j \left( \tilde{\mu}_d^i, F_t^j \right) \Phi_3^{p_i} \left( \begin{bmatrix} \frac{\tilde{\xi}^j - \tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \beta_j \tilde{\sigma}_d^i \\ \frac{-\tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \beta_j \tilde{\sigma}_d^i \end{bmatrix}, \frac{R_{jk}(\tilde{\mu}_d^i, 0) - \beta_j^2 (\tilde{\sigma}_d^i)^2}{\sigma_{j,d}}, \frac{\beta_j \tilde{\sigma}_d^i}{\sqrt{2}}, \frac{\beta_j \tilde{\sigma}_d^i}{\sigma_{j,d}} \right) \right. \right. \\ &+ b_j \left( \tilde{\mu}_d^i - \tilde{\xi}^k, F_t^j \right) \Phi_3^{p_i} \left( \begin{bmatrix} \frac{\tilde{\xi} - \tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \beta_j \tilde{\sigma}_d^i \\ \frac{\tilde{\xi}^k - \tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \beta_j \tilde{\sigma}_d^i \end{bmatrix}, \frac{-R_{jk}(\tilde{\mu}_d^i - \tilde{\xi}^k, \tilde{\xi}^k) + \beta_j^2 (\tilde{\sigma}_d^i)^2}{\sigma_{j,d}}, \frac{\beta_j \tilde{\sigma}_d^i}{\sqrt{2}}, \frac{-\beta_j \tilde{\sigma}_d^i}{\sigma_{j,d}} \right) \left. \right\} \\ &+ \sum_{j \in J_i} \hat{\delta}_j e^{\eta} b_{J_i}(\tilde{\mu}_d^i, \mathbf{F}_t^{J_i}) \left\{ -\Phi_3^{p_i} \left( \begin{bmatrix} \frac{\tilde{\xi}^j - \tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \gamma_2 \tilde{\sigma}_d^i \\ \frac{-\tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \gamma_2 \tilde{\sigma}_d^i \end{bmatrix}, \frac{R_{jk}(\tilde{\mu}_d^i, 0) + \gamma_0^k \sigma^2 - \gamma_2 \beta_j (\tilde{\sigma}_d^i)^2}{\hat{\delta}_j \sigma_{j,d}}, \frac{\gamma_2 \tilde{\sigma}_d^i}{\sqrt{2}}, \frac{\beta_j \tilde{\sigma}_d^i}{\hat{\delta}_j \sigma_{j,d}} \right) \right. \\ &+ \Phi_3^{p_i} \left( \begin{bmatrix} \frac{\tilde{\xi} - \tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \gamma_2 \tilde{\sigma}_d^i \\ \frac{\tilde{\xi}^k - \tilde{\mu}_d^i}{\tilde{\sigma}_d^i} - \gamma_2 \tilde{\sigma}_d^i \end{bmatrix}, \frac{R_{jk}(\tilde{\mu}_d^i - \tilde{\xi}^k, \tilde{\xi}^k) + \gamma_0^k \sigma^2 - \gamma_2 \beta_j (\tilde{\sigma}_d^i)^2}{\hat{\delta}_j \sigma_{j,d}}, \frac{\gamma_2 \tilde{\sigma}_d^i}{\sqrt{2}}, \frac{\beta_j \tilde{\sigma}_d^i}{\hat{\delta}_j \sigma_{j,d}} \right) \left. \right\} \\ &+ \left[ p_i \Phi \left( \frac{-\tilde{\mu}_d^i}{\tilde{\sigma}_d^i} \right) + \bar{p}_i \Phi_2 \left( \frac{-\tilde{\mu}_d^i}{\tilde{\sigma}_d^i}, 0; \frac{-1}{\sqrt{2}} \right) \right] \sum_{j \in J_i} b_j \left( 0, F_t^j \right) \Phi \left( \frac{R_{jk}(0, 0)}{\sigma_{J_i}} \right) \\ &+ \left[ p_i \Phi \left( \frac{\tilde{\mu}_d^i - \tilde{\xi}}{\tilde{\sigma}_d^i} \right) + \bar{p}_i \Phi_2 \left( \frac{\tilde{\mu}_d^i - \tilde{\xi}}{\tilde{\sigma}_d^i}, 0; \frac{-1}{\sqrt{2}} \right) \right] \sum_{i \in I} b_j \left( \tilde{\xi}^j, F_t^j \right) \Phi \left( \frac{-R_{jk}(\tilde{\xi}^j, \tilde{\xi}^k)}{\sigma_{J_i}} \right) \left. \right\} \\ &+ \sum_{i=N_1+1}^{N_2} (-1)^{\hat{\delta}_i} \left( \prod_{j=1}^n F_t^j \right)^{\hat{\delta}_i} \exp \left( \hat{\delta}_i \sum_{j=1}^n \left( \alpha_j - \frac{1}{2} \sigma_j^2 + \sum_{k=1}^n \Sigma_{j,k}^{(S)} \right) \right) f(\mu_d, \sigma_d, \bar{\xi}, p_i, \bar{p}_i, \alpha_i, \beta_i, \tilde{\delta}_i) \end{aligned}$$

where  $k = J_i \setminus \{j\}$ ,  $\hat{\delta}_j = (-1)^{\mathbb{1}_{\{j=i_+\}}}$  (for  $j \in J_i$ , where  $i \in \{1, \dots, N_1\}$ ), with  $f(\cdot)$  as defined in (14) and

$$\begin{aligned} \sigma_{J_i}^2 &:= \sigma_{i_+}^2 - 2\rho_{i_+i_-} \sigma_{i_+} \sigma_{i_-} + \sigma_{i_-}^2, \\ \sigma_{j,d}^2 &:= \beta_j^2 (\tilde{\sigma}_d^i)^2 + \sigma_{J_i}^2, \\ \eta &:= \frac{\gamma_2^2 (\tilde{\sigma}_d^i)^2 - \alpha_c \alpha_g \sigma^2}{2} \end{aligned}$$

$$R_{jk}(\xi_j, \xi_k) := \alpha_k + \beta_k \xi_k - \alpha_j - \beta_j \xi_j + \log \left( F_t^k \right) - \log \left( F_t^j \right) - \frac{1}{2} \sigma_{J_i}^2.$$

Although the formula above appears remarkably involved, note that for a two fuel market (e.g. gas and coal) we are likely to only have one ‘normal’ regime ( $N_1 = 1$ ) and one or two other regimes, reducing the complexity of the formula. If  $N_2 = 0$ , we return to the result of [21]. However, the additional generality allows extra flexibility. For example, in a market

with three fuels but very little chance of overlap between the highest and lowest, we might set  $N_1 = 2$ , prescribing one regime where the highest fuel mixes with the middle fuel, and one where the middle fuel mixes with the lowest.

**5.2.2. Version B.** In the second version of the multi-fuel framework, we do not define capacities  $\xi^1, \dots, \xi^n$ . Hence it is less clear how to capture changes in the merit order, because there is no concept of a threshold capacity level where the marginal fuel type changes when demand crosses that threshold. Instead, the only way to approximate the subtle interplay between demand and marginal fuel type is to use the idea of regimes instead. For example, for an  $n$  fuel market, one might define  $n + 1$  regimes, one driven by each underlying fuel, and one for spikes. Although this does not incorporate an overlap regime, the simplification to a single marginal fuel is intuitively appealing and similar in spirit to the work of Aid *et al* [1, 2]. Instead of their strict capacity-driven thresholds, we then use our demand-dependent regime probabilities to ensure that fuels higher in the merit order are more likely to be used when demand is high. A big obstacle to either approach is that the merit order may change, particularly over medium to long time horizons. Indeed, in [2], to retain mathematical tractability in the  $n$  fuel case, the authors assume that the initial merit order is fixed and enforce this by modeling the spreads between neighbouring fuels as Geometric Brownian Motions, a departure from commonly-used models for fuel prices. In this section, we present another variation in order to retain the chance of future merit order changes, following more closely the original setup of [1].

Similarly to (20), we define an exponential curve for each fuel type, but now treating  $C_t$  as an additional stochastic factor

$$b_i(D_t, C_t, S_t^i) := S_t^i \exp(\alpha_i + \beta_i D_t - \gamma_i C_t), \quad \text{for } i = 1, \dots, n.$$

Regimes  $1, \dots, n$  (the ‘normal’ regimes) will be driven by fuels  $1, \dots, n$  only. Note that it is certainly possible to incorporate ‘overlap’ regimes in this framework, for example by choosing the function  $(S_t^1)^\epsilon (S_t^2)^{1-\epsilon} \exp(\alpha_{1,2} + \beta_{1,2} D_t - \gamma_{1,2} C_t)$  for a regime driven jointly by fuels 1 and 2. While explicit forward curves can still be found, this adds unnecessary complications for our illustrative purposes here. Regimes  $n + 1, \dots, N$  (the ‘spike’ regimes) will be driven by either no fuels or all fuels jointly. Thus,

$$b_i(D_t, C_t, S_t^i) := (-1)^{\delta_i} \prod_{j=1}^n \left( S_t^j \exp(\alpha_j) \right)^{\delta_j} \exp(\alpha_i + \beta_i D_t - \gamma_i C_t) \quad \text{for } i = n + 1, \dots, N$$

Next, before defining the power price  $S_t$ , we define a permutation  $\{\pi_t(1), \dots, \pi_t(n)\}$  over the set  $\{1, \dots, n\}$  of fuels, such that

$$S_t^{\pi_t(1)} \exp(\alpha_{\pi_t(1)} + \beta_{\pi_t(1)} \mu_d) \geq \dots \geq S_t^{\pi_t(n)} \exp(\alpha_{\pi_t(n)} + \beta_{\pi_t(n)} \mu_d).$$

This is similar to the approach followed in [1], but without the restriction of a single heat rate per fuel type (i.e. a step function bid stack). We then define the spot price as

$$(21) \quad S_t = b_{\pi_t(i)}(D_t, C_t, S_t^{\pi_t(i)}) \quad \text{with probability } p_i(D_t, C_t)$$

where again (in the most general form)

$$p_i(D_t, C_t) = \underline{p}_i + \bar{p}_i \Phi((\zeta_i + \eta_i D_t + \theta_i C_t)) \quad \text{for } i = 1, \dots, N - 1,$$

$$\text{and } p_N(D_t, C_t) = 1 - \sum_{i=1}^{N-1} p_i(D_t, C_t).$$

In other words, the idea is that  $\pi_t$  approximately captures the ordering of fuel types (the merit order), using the average demand level in the market. Then the regime probabilities  $p_i$  (increasing in  $D_t$  for lower values of  $i$ , decreasing in  $D_t$  for higher values of  $i$ ) can do

the work of linking the more expensive fuel types to higher demand states, and the cheaper ones to lower demand states. Unlike in Version A, determining which fuel price sets the power prices does not depend on a function of both  $D_t$  and  $S_t^1, \dots, S_t^n$  jointly, thus easing the computation of forward prices. While the connection between demand and price is looser than in a strict stack model, this is not necessarily unrealistic for a market with significant noise from  $C_t$ .

Assuming the model in (21), distributions in (19) and  $p_i$  given by (17) for simplicity, in the case of two fuels ( $n = 2$ ) the forward power price  $F_t^p$  for time  $T$  delivery is given by

$$F_t^p = \sum_{i=1}^2 \left\{ \sum_{j=1}^2 h_j F_t^j \Phi \left( \frac{(-1)^i R_{jk}(\mu_d, \mu_d)}{\sqrt{\sigma_1^2 - 2\rho_{12}\sigma_1\sigma_2 + \sigma_2^2}} \right) \left[ p_i + \bar{p}_i \Phi \left( \frac{(\beta_j\sigma_d - \rho\gamma_j\sigma_c)}{\sqrt{(1-\rho^2)(2-\rho^2)}} \right) \right] \right\} \\ + \sum_{i=3}^N (-1)^{\delta_i} h_i \left( F_t^1 F_t^2 \exp \left( 2\rho_{12}\sigma_1\sigma_2 + \sum_{j=1}^2 \left( \alpha_j + \frac{1}{2}\sigma_j^2 \right) \right) \right)^{\delta_i} \left[ p_i + \bar{p}_i \Phi \left( \frac{(-1)^{\delta_i}(\beta_i\sigma_d - \rho\gamma_i\sigma_c)}{\sqrt{(1-\rho^2)(2-\rho^2)}} \right) \right]$$

where  $k = \{1, 2\} \setminus j$  and as before,  $h_i = \mathbb{E}_t[\exp(\alpha_i + \beta_i D_t - \gamma_i C_t)]$  as given (for  $i = 1, \dots, N$ ) in (18) and  $R_{jk}$  also as defined earlier.

At the expense of a weaker link to the merit order, this result is clearly much simpler than the Version A forward price, due to the lack of any indicator functions involving demand. Nonetheless the complexity increases rapidly for more than two fuels. The three fuel case ( $n = 3$ ) is still realistic to write out on paper (using the trivariate relationship in (9) and some determination!), but for  $n > 3$  the increasing dimensionality of the multivariate Gaussian and the increasing number of permutations of fuels renders a closed-form solution nearly infeasible, although numerical implementation is still straightforward.

**5.3. Parameter Estimation and Forward Curve Calibration.** The choice of structural model clearly depends on both the electricity market in question and the goals of the model. The framework above and its many versions were intended to emphasize the variety of tools available for modeling the many features we observe in price dynamics, while retaining a common core to the model and a reasonable level of mathematical tractability. No matter which specific model is ultimately chosen, an important next step is a reliable and robust method for parameter estimation and forward curve calibration. These two issues are most often tackled in stages, first estimating some parameters from history, and then selecting others to match forward looking market quotes. In all cases, the explicit formulas above for  $F^p(t, T)$  (and their explicit dependence on fuel forwards) provide a valuable computational benefit, as an optimal fit to observed forward curves in a high dimensional model quickly becomes unmanageable if limited to Monte Carlo simulation. In this subsection we discuss briefly the main challenges involved in fitting a structural model to data.

- *Observable vs Unobservable Factors:* As discussed throughout this paper, many of the underlying factors (e.g. demand, fuel prices) in electricity markets are easily observable and exogenously modelled, meaning that their parameters can be estimated independently of the power price model itself, by standard techniques such as maximum likelihood. However, the sheer complexity of the market typically means that some factors are either truly unobservable or their treatment in the model approximates several effects, making them effectively unobservable for modeling purposes. In our framework, ‘capacity’  $C_t$  typically falls in this category, particularly when treated as a catch-all noise process (i.e. Version B). In such cases the model-implied history of the process can be backed out from spot price histories as a model residual (as suggested in [72, 68, 39] among others),

typically producing a closer fit to historical price dynamics. However this benefit must be weighed against the risk that structural or regulatory changes can make market price histories unreliable, a weakness not suffered by models driven only by observed factors.

- *Stack Parameters:* As discussed by [68] and [49] among others, the estimation of stack parameters ( $\alpha$ 's and  $\beta$ 's in our expressions) can be achieved in several ways. In summary, (i) from costs, (ii) from bids, (iii) from prices. The first of these relies somewhat on the assumption of competitive markets and limited strategy bidding (ie, generators bidding their production costs), but has the significant advantage of avoiding messy estimation and fixing some key parameters to well understood market variables. For example, in the multi-fuel case (Version A), we may be able to easily approximate the range of heat rates (efficiencies) in the market for each technology  $j \in \{1, \dots, n\}$ , which can then be equated to the range  $[\exp(\alpha_j), \exp(\alpha_j + \beta_j \xi^j)]$  for the 'normal' regimes. On the other hand, if historical auction data is available, then parameters  $\alpha$  and  $\beta$  can be fitted directly to the bid stack (as in [39]), though this is less suitable if there are multiple 'normal' regimes. Finally, the use of only prices to fit the exponential bid curves is disadvantageous since we only observe one point on the curve each hour. Hence regions of the curves which only rarely set the price may be harder to accurately estimate than when using the historical stack data. However, as the occurrence of 'spike' regimes in history is unobservable, price data may be particularly useful for filtering out these extreme points, and fitting their parameters separately. In practice, good judgement is needed from market to market when deciding how best to tackle stack calibration, and a combination of the above options may be preferable.
- *History vs Market Quotes:* The choice between using history and using current market quotes revolves around several key questions, including the availability and reliability of data, the treatment of risk premia, and the desired model inputs. Variables such as temperature (which can be mapped to demand, as discussed in Section 3.2) have long and reliable histories, and are hence one of the arguments in favor of structural models over reduced-form. However, fitting all parameters to history and assuming a constant market price of risk when needed will of course fail to reproduce the market forward prices, a typical first step in any modeling problem. Therefore, a balance must be struck between parameters matched to history and parameters matched to future risk-neutral dynamics (i.e. to observed prices). The simplest approach in the structural framework described above is to first fit everything to history, before allowing the mean level of demand  $\mu_d$  (or capacity  $\mu_c$ ) to be time-dependent, chosen precisely to reproduce each market forward price  $F^p(t, T)$  (see, [37] for a similar approach in reduced-form). Solving for  $\mu_d$  numerically is straightforward given the expressions above, giving an exact calibration to the forward curve. In other cases, we may be interested in allowing more parameters to be free, in order to match other input prices from the market, such as at-the-money options if they are sufficiently liquid.
- *Risk premia:* A key advantage of the expressions above is that power forwards  $F^p(t, T)$  are written directly in terms of fuel forwards (say  $F^g(t, T)$ ), which can be treated as observed market prices. Hence, for the fuel component of our model, no assumption regarding risk premia is needed, as the risk-neutral drift is implicitly specified by observed forwards. Moreover, no calibration technique is needed for the fuel forward curves. On the other hand, the dynamics of other factors such as demand and capacity can only be estimated under the physical measure via history (unless the careful use of weather derivatives can provide information for demand via temperature). Hence, rather arbitrary assumptions about the form

of the market price(s) of risk may be required. However, as discussed in the previous paragraph, we typically desire exact calibration to the observed power forward curve, through which the market price of risk can be absorbed into the mean level  $\mu_d$  (or  $\mu_c$ ) needed to match forwards. While this may simply sweep the issue under the carpet, for many practical applications it should suffice as a reasonable assumption and claims of incompleteness can be politely ignored by pointing to a liquid forward curve for all maturities.

- *Delivery Periods:* As we gently sweep one issue under the carpet, another crops up around the corner! While it may be true that liquid forward prices exist covering several years from the current time, they certainly do not exist for every specific maturity  $T$ , simply because of delivery periods. As mentioned in Section 2, the convention in all electricity markets is that forward contracts specify electricity delivery over a period of time, often one month, but sometimes even a quarter or a year for longer contracts. Hence the price should be written  $F(t, T_1, T_2)$  and correspond to an average of the expected spot price for all hours in the month (or delivery period). While some authors have approximated this in continuous time as an integral over the delivery period and designed models for which the integral simplifies (cf. [13, 11]), in reality a sum is arguably more appropriate, since  $S_t$  is indeed a discrete time process. Adding an outer summation to our expressions above is a simple adjustment, but in this case we should note that fuel forward prices then also require single hour maturities, which is unfortunately not the case, as these are typically also monthly. Various remedies are possible, including the smoothing of observed forward curves to obtain prices for all  $T$  (cf. [12, 61] for smoothing electricity forward quotes), the assumption of piecewise constant fuel forwards, or the choice of a representative single date per delivery period (reasonable for longer maturities). Unfortunately, there is no clear answer, and such implementation challenges exist no matter what price model we use!
- *Hour of Day Considerations:* Finally, we note that observed power forward curves often exist for delivery over different hours of the day (day vs. night) and days of the week (weekday vs. weekend), categorized as peak, off-peak or base-load contracts. Hence, the calibration to observed forward quotes may require multiple calibrations per delivery month. In the simplest case, one might simply adjust the mean demand  $\mu_d$  (or  $\mu_c$ ) by a different amount for peak and off-peak hours. A well-fitted model for load should first capture the well-known hourly patterns across the day, as well as seasonal periodicities which may vary significantly for different hours of the day. While some authors have chosen to treat each hour of the day as a separate (but correlated) stochastic process, others treat only the deterministic component of demand differently by hour. Having 24 separate processes may introduce too many parameters, particularly as the historical sample size drops significantly, making it hard for example to stably model the tail of the price distribution.

In conclusion, the fitting procedure for structural models is typically a fine art, combining different approaches for different components of the overall framework. As with all models for electricity, approximations must be made. Structural models in particular may have very many parameters to estimate, but in exchange can have much data available to help.

## 6. CONCLUSION

In this survey, we have attempted to give the reader a flavour of many of the interesting and unique characteristics of energy (and commodity) markets, and in particular the most unusual of all, electricity. While many different price modeling approaches now exist, the

topic still provides many avenues for important new research, both building on current work and addressing new questions as they arise. For example, how will electricity grids manage the rapid growth of renewable energy with large supply variability, and how will prices be affected? Will the emissions markets grow in importance globally and produce more dramatic changes in the merit order? Will the smart grid and growth of electric vehicles cause a structural change, with both the demand inelasticity and non-storability assumptions under threat? Will new storage technology bring electricity price dynamics closer in line with other commodities? What about the ‘financialization’ of electricity, if power forwards some day begin to appear in commodity indices? How global can electricity markets become (e.g. with solar panels in the Sahara powering much of Europe and Africa)? Some of these thoughts may be a long way off, but others could be just around the corner! We do not promote the structural approach discussed in detail here as an answer to such intriguing speculative questions, but we do recommend thinking beyond the historical price series, especially at times of fundamental market change. We have presented and discussed structural models which meet this criteria by directly incorporating demand, capacity and fuel prices, and without necessarily sacrificing the mathematical benefits traditionally reserved for reduced-form approaches. We hope that the flexible, intuitive and practical framework we advocate can play a useful role in understanding and tackling the many risks ahead in the fascinating next chapter of the global energy markets.

## REFERENCES

- [1] R. Aïd, L. Campi, Adrien N. Huu, and N. Touzi. A structural risk-neutral model of electricity prices. *International Journal of Theoretical and Applied Finance*, 12:925–947, 2009.
- [2] R. Aïd, L. Campi, and N. Langrené. A structural risk-neutral model for pricing and hedging power derivatives. *Mathematical Finance*. Published on Line Feb. 13, 2012.
- [3] R. Aïd, A. Chemla, A. Porchet, and N. Touzi. Hedging and vertical integration in electricity markets. *Management Science*, 57:1438–1452, 2011.
- [4] C.L. Anderson and M. Davison. A hybrid system-econometric model for electricity spot prices: Considering spike sensitivity to forced outage distributions. *IEEE Transactions on Power Systems*, 23:927–937, 2008.
- [5] N. Audet, P. Heiskanen, J. Keppo, and I. Vehvilainen. Modeling electricity forward curve dynamics in the nordic market. In D. Bunn, editor, *Modelling Prices in Competitive Electricity Markets*, pages 251–265. Wiley, 2004.
- [6] E. Banks. *Weather Risk Management: Markets, Products and Applications*. Palgrave, 2002.
- [7] M. Barlow. A diffusion model for electricity prices. *Mathematical Finance*, 12(4):287–298, 2002.
- [8] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Modelling electricity forward prices by ambit fields. Technical report, 2011. Preprint.
- [9] O.E. Barndorff-Nielsen, F.E. Benth, and A.E.D. Veraart. Modelling energy spot prices by Lévy semistationary processes. Technical report, 2011. Preprint.
- [10] P. Barrieu and N. El Karoui. Pricing, hedging and designing derivatives with risk measures. In R. Carmona, editor, *Indifference Pricing*, pages 77–146. Princeton University Press, Princeton, NJ, 2009.
- [11] F. E. Benth, J. Kallsen, and T. Meyer-Brandis. A non-gaussian ornstein-uhlenbeck process for electricity spot price modeling and derivatives pricing. *Applied Mathematical Finance*, 14:153–169, 2007.
- [12] F. E. Benth, S. Koekebakker, and F. Ollmar. Extracting and applying smooth forward curves from average-based commodity contracts with seasonal variation. *Journal of Derivatives*, Fall 15:52–66, 2007.
- [13] F.E. Benth, J.S. Benth, and S. Koekebakker. *Stochastic Modeling of Electricity and Related Markets*. World Scientific.
- [14] M. Bernhart, H. Pham, P. Tankov, and X. Warin. Swing options valuation: A bsde with constrained jumps approach. In R. Carmona et al., editor, *Numerical methods in finance*, Springer Proceedings in Mathematics. Springer Verlag, 2012.
- [15] T. Björk and C. Landén. On the term structure of futures and forward prices. In *Eds.*, *Mathematical Finance-Bachelier Congress 2000*, pages 111–149. Springer, 2002.
- [16] A. Boogert and D. Dupont. When supply meets demand: the case of hourly spot electricity prices. *IEEE Transactions on Power Systems*, 23(2):389–398, 2008.
- [17] I. Bouchouev. The inconvenience yield or the theory of normal contango. *Energy Risk Magazine*, September 2011.
- [18] M. Burger, B. Graeber, and G. Schindlmayr. *Managing Energy Risk: An Integrated View on Power and Other Energy Markets*. Wiley.

- [19] M. Burger, B. Klar, A. Müller, and G. Schindlmayr. A spot market model for pricing derivatives in electricity markets. *Quantitative Finance*, 4:109–122, 2004.
- [20] R. Carmona. Applications to weather derivatives and energy contracts. In R. Carmona, editor, *Indifference Pricing*, pages 241–264. Princeton University Press, Princeton, NJ, 2009.
- [21] R. Carmona, M. Coulon, and D. Schwarz. Electricity price modeling and asset valuation: a multi-fuel structural approach. Technical report, 2012.
- [22] R. Carmona, M. Coulon, and D. Schwarz. The valuation of clean spark spread options: Linking electricity, emissions and fuels. Technical report, 2012.
- [23] R. Carmona and S. Dayanik. Optimal multiple stopping of linear diffusions. *Mathematics of Operations Research*, 33:446–460, 2008.
- [24] R. Carmona and P. Diko. Pricing precipitation based derivatives. *International Journal of Theoretical and Applied Finance*, 7:959–988, 2005.
- [25] R. Carmona and V. Durrleman. Pricing and hedging spread options. *SIAM Rev.*, 45(4):627–685, 2003.
- [26] R. Carmona, F. Fehr, J. Hinz, and A. Porchet. Market designs for emissions trading schemes. *SIAM Review*, 52:403–452, 2010.
- [27] R. Carmona and J. Hinz. Least Squares Monte Carlo approach to convex control problems. Technical report, Princeton University, 2011.
- [28] R. Carmona and M. Ludkovski. Valuation of energy storage: an optimal switching approach. *Quantitative Finance*, pages 359–374.
- [29] R. Carmona and M. Ludkovski. Spot convenience yield models for the energy markets. In *Mathematics of finance*, volume 351 of *Contemp. Math.*, pages 65–79. Amer. Math. Soc., Providence, RI, 2004.
- [30] R. Carmona and M. Ludkovski. Pricing asset scheduling flexibility using optimal switching. *Applied Mathematical Finance*, 15(5):405–447, 2008.
- [31] R. Carmona and Y. Sun. Implied and local correlations from spread options. *Applied Mathematical Finance*, 2012. (to appear).
- [32] R. Carmona and N. Touzi. Optimal multiple stopping and valuation of swing options, 2008.
- [33] P. Carr and D. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(5):61 – 73, 1998.
- [34] A. Cartea and M. Figueroa. Pricing in electricity markets: A mean reverting jump diffusion model with seasonality. *Applied Mathematical Finance*, 12(4):313–335.
- [35] A. Cartea, M. Figueroa, and H. Geman. Modelling electricity prices with forward looking capacity constraints. *Applied Mathematical Finance*, 32:2501–2519, 2008.
- [36] A. Cartea and P. Villaplana. Spot price modeling and the valuation of electricity forward contracts: the role of demand and capacity. *Journal of Banking and Finance*, 32:2501–2519, 2008.
- [37] L. Clewlow and C. Strickland. Valuing energy options in a one factor model fitted to forward prices. working paper, 1999.
- [38] L. Clewlow and C. Strickland. *Energy Derivatives: Pricing and Risk Management*. Lacima Productions, 2000.
- [39] M. Coulon and S. Howison. Stochastic behaviour of the electricity bid stack: from fundamental drivers to power prices. *Journal of Energy Markets*, 2:29–69, 2009.
- [40] M. Coulon, W. Powell, and R. Sircar. A model for hedging load and price risk in the Texas electricity market. Technical report, 2012.
- [41] M. Culot, V. Goffin, S. Lawford, S. de Menten, and Y. Smeers. An affine jump diffusion model for electricity. working paper, Catholic University of Leuven.
- [42] G.d.M. D’Aertrycke and Y. Smeers. The valuation of power futures based on optimal dispatch. *Journal of Energy Markets*, 3(3):27–50, 2010.
- [43] C. De Jong and R. Huisman. Option pricing for power prices with spikes. *Energy Power Risk Management*, 7:12–16, 2003.
- [44] C. De Jong and S. Schneider. Cointegration between gas and power spot prices. *Journal of Energy Markets*, 2(3):27–46, 2009.
- [45] S. Deng. Stochastic models of energy commodity prices and their applications: Mean-reversion with jumps and spikes. Technical report, University of California Energy Institute, 1999.
- [46] D. Duffie, J. Pan, and K. Singleton. Transform analysis and asset pricing for affine jump-diffusions. *Econometrica*, 68(6):1343–76, 2000.
- [47] G. W. Emery and Q. Liu. An analysis of the relationship between electricity and natural-gas futures prices. *The Journal of Futures Markets*, 22(2):95–122, 2002.
- [48] A. Eydeland and H. Geman. Pricing power derivatives. *Risk Magazine*, 10:71–73, 1998.
- [49] A. Eydeland and K. Wolyniec. *Energy and Power Risk Management: New Developments in Modeling, Pricing and Hedging*. Finance. Wiley, 2003.
- [50] P. Forsythe. A semi-lagrangian approach for natural gas valuation and optimal operation. *SIAM Journal on Scientific Computing*, 30:339–368, 2007.



- [51] N. Frikha and V. Lemaire. Joint modelling of gas and electricity spot prices. *Applied Mathematical Finance*, 2012. to appear.
- [52] H. Geman. *Commodities and commodity derivatives: Modeling and Pricing of Agriculturals, Metals and Energy*. Finance. Wiley, 2005.
- [53] H. Geman. *Risk Management in Commodity Markets: From Shipping to Agriculturals and Energy*. Finance. Wiley, 2008.
- [54] H. Geman and A. Roncoroni. Understanding the fine structure of electricity prices. *Journal of Business*, 79, 2006.
- [55] R. Gibson and E. S. Schwartz. Stochastic convenience yield and the pricing of oil contingent claims. *Journal of Finance*, XLV(3):959–976, 1990.
- [56] S. Howison and D. Schwarz. Structural modelling of carbon emission markets. Technical report, 2012. submitted.
- [57] P. Jaillet, E. Ronn, and S. Tompaidis. Valuation of commodity-based swing options. *Management Science*, 50:909921, 2004.
- [58] S. Jaimungal and V. Surkov. Lévy based cross-commodity models and derivative valuation. *SIAM Journal of Financial Mathematics*, 2:464 –, 2011.
- [59] T. Kanamura and K. Ohashi. A structural model for electricity prices with spikes: measurement of spike risk and optimal policies for hydropower plant operation. *Energy Economics*, 29(5):1010–1032, 2007.
- [60] V. Kholodnyi. A non-markovian process for power prices with spikes and valuation of european contingent claims on power. Technical report, Erasmus Energy Library, 2001.
- [61] T. Kluge. *Pricing Swing Options and Other Electricity Derivatives*. PhD thesis, 2006. DPhil Thesis, University of Oxford.
- [62] S. Koekebakker and F. Ollmar. Forward curve dynamics in the nordic electricity market. *Managerial Finance*, 31:74–95, 2005.
- [63] R. Litterman and J. Scheinkman. Common factors affecting bond returns. *Journal of Fixed Income*, 1:49–53, 1991.
- [64] J. Lucia and E. Schwartz. Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research*, 2002.
- [65] M. Lyle and R. Elliott. A "simple" hybrid model for power derivatives. *Energy Economics*, 31:757–767, 2009.
- [66] T.D. Mount, Y. Ning, and X. Cai. Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics*, 28:62–80, 2006.
- [67] C. Pirrong. *Commodity Price Dynamics: A Structural Approach*. Cambridge University Press, 2012.
- [68] C. Pirrong and M. Jermakyan. The price of power: The valuation of power and weather derivatives. *Journal of Banking and Finance*, 32:2520–2529, 2008.
- [69] J. Porter. Evolution of the global weather derivatives market. In P. Field, editor, *Modern Risk Management: a History*. Risk Books, 2004.
- [70] E. Schwartz. The stochastic behaviour of commodity prices: Implications for valuation and hedging. *The Journal of Finance*, 3:923–973, 1997.
- [71] E. Schwartz and J. Smith. Short-term variations and long-term dynamics in commodity prices. *Management Science*, 46:893–911, 2000.
- [72] P. Skantze, A. Gubina, and M. Ilic. Bid-based stochastic model for electricity prices: the impact of fundamental drivers on market dynamics. MIT E-lab report, November 2000.
- [73] K. Tang and W. Xiong. Index investment and financialization of commodities. Technical report, 2010. NBER Working Paper 16385.
- [74] S. Thoenes. Understanding the determinants of electricity prices and the impact of the German Nuclear Moratorium in 2011. Technical report, 2011. EWI Working Paper.
- [75] F. Turbault and Y. Youlal. Swing option pricing by optimal exercise boundary estimation. In R. Carmona et al., editor, *Numerical methods in finance*, Springer Proceedings in Mathematics. Springer Verlag, 2012.
- [76] A. Wagner. Residual demand modelling and application to electricity pricing. Technical report, Technische Universitaet Kaiserslautern, 2012.
- [77] X. Warin. Gas storage hedging. In R. Carmona et al., editor, *Numerical methods in finance*, Springer Proceedings in Mathematics. Springer Verlag, 2012.
- [78] R. Weron. *Modeling and Forecasting Electricity Loads and Prices: a statistical approach*. Wiley.
- [79] R. Weron, M. Bierbrauer, and S. Truck. Modeling electricity prices: jump diffusion and regime switching. *Physica A: Statistical Methods and It's Applications*, 336:39–48, 2004.
- [80] K. Wiebauer. A practical view on valuation of multi-exercise american style options in gas and electricity markets. In R. Carmona et al., editor, *Numerical methods in finance*, Springer Proceedings in Mathematics. Springer Verlag, 2012.

BENDHEIM CENTER FOR FINANCE, DEPT. ORFE, PRINCETON UNIVERSITY, PRINCETON NJ 08544, USA

*E-mail address:* `rcarmona@princeton.edu`

ORFE, UNIVERSITY OF PRINCETON, PRINCETON NJ 08544, USA

*E-mail address:* `mcoulon@princeton.edu`