

# Analysis and Forecasting of Electricity Price Risks with Quantile Factor Models

*Derek Bunn\*, Arne Andresen\*\*, Dipeng Chen\*\*\*, and Sjur Westgaard\*\**

---

## ABSTRACT

Forecasting quantile and value-at-risk levels for commodity prices is methodologically challenging because of the distinctive stochastic properties of the price density functions, volatility clustering and the importance of exogenous factors. Despite this, accurate risk measures have considerable value in trading and risk management with the topic being actively researched for better techniques. We approach the problem by using a multifactor, dynamic, quantile regression formulation, extended to include GARCH properties, and applied to both in-sample estimation and out-of-sample forecasting of traded electricity prices. This captures the specification effects of mean reversion, spikes, time varying volatility and demonstrates how the prices of gas, coal and carbon, forecasts of demand and reserve margin in addition to price volatility influence the electricity price quantiles. We show how the price coefficients for these factors vary substantially across the quantiles and offer a new, useful synthesis of GARCH effects within quantile regression. We also show that a linear quantile regression model outperforms skewed GARCH-t and CAViaR models, as specified on the shocks to conditional expectations, regarding the accuracy of out-of-sample forecasts of value-at-risk.

**Keywords:** Electricity, Prices, Forecasting, Risk

<http://dx.doi.org/10.5547/01956574.37.1.dbun>

## 1. INTRODUCTION

For managers involved in risk and operations, as well as for regulators concerned with market surveillance, modelling and forecasting the tails of price distributions in traded markets may often be a more crucial activity than formulating central expectations. As a task, it is certainly the more methodologically challenging. Both the relative sparseness of the data in the tails and the extreme sensitivity of the results to misspecification in the functional form of the distribution create severe difficulties. Thus, robust parametric methods for specifying predictive distributions (eg Guermat et al, (2001), (2002)), as well as semiparametric formulations for estimating specific quantiles (eg Engle et al (2004), Gerlach et al., (2011)), have characterised recent research. Ad hoc representations of skewed and fat-tailed properties have become particularly necessary in practice (Kuester et al. (2006)). In the context of forecasting volatility, or Value-at-Risk (VaR) calculations (eg

\* Corresponding author. E-mail: dbunn@london.edu. Department of Management Science and Operations, London Business School, Sussex Place, Regent's Park, NW1 4SA, London, UK.

\*\* Department of Industrial Economics and Technology Management, Norwegian University of Science and Technology, Alfred Getz vei 3, 7041 Trondheim, Norway.

\*\*\* Department of Management Science and Operations, London Business School, Sussex Place, Regent's Park, NW1 4SA, London, UK.

Hong et al, 2009; Glasserman et al, 2000) the increasing range of pragmatic approaches to this problem has inevitably led to hybrid and combining proposals in attempts to improve out-of-sample predictive performances (Taylor et al, 1998, 1999; Jeon et al, 2013). Evidently there is a need in practice to develop more accurate forecasts of tail risk rather than relying upon what the stylised theoretical models of the price formation processes generally provide.

Quantile regression, following Koenker et al (1978), has promised several attractive features in this respect. Firstly, it offers a semiparametric formulation of the predictive distribution, the quantiles of which can be efficiently estimated with distinct regressions. It is possible therefore to have a fine resolution of the tail characteristics in terms of empirical estimates for the required quantiles. Fundamental factors can be specified in the quantile regressions, which may exhibit different coefficients according to the quantile levels. This feature offers greater predictive insights and accuracy. Secondly, the conditional nature of the regressions is valuable for the explicit representations of varying dependencies in scenario tree creation (as may be required, for example, in stochastic optimisation models for risk management, eg Alexander et al, 2004). Thirdly, the non-parametric way in which a tail distribution can be developed from separate quantile models offers an alternative perspective to multi-process modelling in the contexts where modellers may be tempted to use mixtures or regime-switching to capture different price formation process for normal and extreme events. Instead of an unobserved latent variable, such as in Markov regime switching (eg, Karakatsani et al., (2008)) and a restriction in the number of regimes, such as in smooth transition logistic regressions (eg Chen et al (2010)), quantile methods inherently associate a separate regime with each quantile.

One of the most critical areas for the use of well estimated tail probabilities is in VaR calculations, which are specified as quantiles, and have therefore motivated substantial research in finding effective quantile forecasting methods (eg Taylor (2008); Füss, (2010)). The methodological benchmark in theory appears to be the conditional autoregressive value at risk (CAViaR) approach inspired by applying GARCH ideas to quantiles, whereby the quantiles are modelled conditionally as autoregressive processes (see Engle et al (2004), Taylor (2005)). These have proven to be very successful in predicting VaR in research studies, but (unlike GARCH models) may be difficult to apply in practice and appear to have acquired only limited implementation in financial risk management. Furthermore, in situations where potential exogenous factors are expected to be significantly distinct in their effects across the quantiles, the CAViaR methodology, unlike quantile regression, has lacked this extra specification, and exogenous factors have only been formulated to estimate the conditional mean.

In this respect, spot prices in wholesale electricity markets present one of the most demanding application areas for comparative methodological research on this theme, and offer a rich set of known exogenous factors that may have different effects across the quantiles. Power prices are characterised by high volatility, positive skewness, substantial volatility clustering and large spikes. Furthermore, evidence for the substantial impact of exogenous fundamental drivers, and their nonlinear response functions with respect to demand, excess capacity and fuel prices, is well established (Chen et al (2010)). Forecasting tail probabilities and VaR, however, despite its widespread appeal and use in energy trading, remains under-researched and retains many open questions. Whilst the nonlinear properties of exogenous variables have motivated interest in regime switching models, their out-of sample performances have not matched their intuitive and highly significant in-sample fitting (eg Kosater et al (2006), Misorek et al., (2006)). Evidence on the value of regime-switching models for forecasting is mixed and seems to suggest that the evolutionary nature of power price formation requires more robust methods. It is an open question, therefore, whether quantile methods can be developed to provide this capability.

Prediction of VaR out-of-sample for energy commodities can be found in Aloui (2008), Chan et al (2006), Füss et al. (2010), Giot et al (2003), and Hung et al. (2008). The majority of these studies use various GARCH models with different specifications of the innovation processes, the main conclusions being that GARCH models need to have fat tailed and possibly skewed distributions to work well. Quantile regression approaches are beginning to be applied in this context (Nowotarski et al, 2014, and Jonsson et al, 2014). Furthermore, it appears to follow both from the evident GARCH properties in power prices suggesting a persistence of shocks, and the in-sample regime-switching models between electricity fundamentals suggesting the need to accommodate varying effects of exogenous factors, that there is a *prima facie* case for examining the joint specification of both fundamental factors and volatility as explanatory factors. As Xiao et al (2009) observe, this presents a complicated nonlinear quantile formulation which is elusive to conventional estimation. One approach taken by Chen et al (2012) uses a smooth transition regime switching GARCH model, with innovations modelled as skewed Laplace distributions to facilitate the quantile estimations. Using Bayesian estimation, this succeeds in estimating quantile regression under conditions of heteroscedasticity and regime switching, but applications with more than one exogenous factor remain challenging to estimate. We adopt an alternative two stage approach of first estimating the GARCH process with a factor model for the price levels, and then augmenting a multifactor model for quantiles with this GARCH process (Xiao et al, 2009, demonstrated the convergence of this process in the univariate case). The benefit of this two stage approach that it preserves a more transparent representation of volatility in its conventional manner, but still isolates how, alongside other factors, its impact varies by quantile.

This approach is tested against a range of best-practice alternatives under the most stringent condition of the evening peak (6:30–7:00pm), which is the most volatile, skewed and spiky trading period in the British power market. The British market is a transparent one with useful exogenous day-ahead information available to market participants, and, being well-established, there is a long time series provide an extensive out-of-sample analysis of day ahead forecasts. With this methodological proposal, it is important to benchmark this quantile regression approach against the most appropriate GARCH models, typical in practice, and also the potential use of CAViaR models in this context. We provide convincing results that the quantile factor model with volatility provides better validated and more accurate forecasts than conventional alternatives.

In the next section we provide a short review of the price formation properties in wholesale electricity, motivate the use of the exogenous variables and indicate why they may be nonlinear. Then, we describe the data, followed by the methods. In-sample and out-of-sample results are then reported and a concluding section summarises the key research contributions.

## **2. ELECTRICITY PRICE FORMATION FUNDAMENTALS**

Electricity is a flow, rather than a stock commodity: it is produced and consumed continuously and instantaneously. Traded physical products are therefore defined and sold in the form of metered contracts for the constant delivery of a specified amount of power over a specified period of time, eg one megawatt for one hour (MWh). Most “spot” markets deal in such hourly products, although some, eg Britain and Australia, have finer granularity at half-hourly intervals. These hourly (or half-hourly) spot prices emerge either from an auction process whereby generators and retailers make offers and bids (which may be held once on the previous day to set all hourly prices for the subsequent day) or continuous trading on an exchange platform from a day ahead until a particular time before actual delivery (eg an hour in Britain). Spot price formation itself, because consumers are price inelastic in the short term and cannot store electricity once generated, is mainly a function

of the demand, technology and competition amongst generators. For a particular level of demand at a particular time, there will be a stack of generating technologies available, and the market-clearing price, is usually taken as the offer price of the most expensive plant needed over that trading period. Thus, if the market were competitive, and generators offered at short-run marginal costs, market price volatility would be envisaged as the projection of demand volatility on to the supply function offered by the generators (Stoft, 2002). Given the various plant technologies available for dispatch, differentiated in terms of costs and operational constraints, this short-run supply cost function is intrinsically steeply increasing, discontinuous and convex. Hence there is an induced skewness from demand volatility into price volatility.

In the presence of these characteristics, and since relatively few customers are able to respond to real-time price signals, spot prices are sensitive to the stochastic variables, such as demand shocks and plant outages. Thus, expectations of spot prices involve at least considerations of the underlying fuel (for short term marginal costs) and the reserve margin (for scarcity pricing above marginal cost). Moreover, as almost all electricity markets are oligopolies, at times of scarcity, when the reserve margin (available supply minus demand) is low, those generators with market power may offer and create market prices substantially above marginal costs (e.g. Wolfram (1999); Wolak et al (1997)). This creates a further behavioural and possibly nonlinear element to extreme price formation. The nonlinear implications of this price formation process for the exogenous factors in a quantile regression model suggest the following functional form propositions:

1. **Demand effect is positive and increases nonlinearly with higher quantiles.** To the extent that the supply function is convex in available capacity, we would expect to see a positive price coefficient to demand with higher effects at higher prices. Thus, one should find the coefficients increase nonlinearly with higher quantiles.
2. **Reserve margin effect is negative and decreases nonlinearly with higher quantiles.** Scarcity induces a propensity for generators to offer at higher prices. Oligopolies characterize power markets and their market power increases with a declining reserve margin. Extreme scarcity can create a residual monopolist. Furthermore, reserve margin shocks may be due to unusual outages in generating capabilities and, once it is known to the market that a generator is short, this may induce a selling squeeze by the other generators and hence higher prices.
3. **Fuel (including carbon allowance prices) coefficients will be positive but may have nonlinear, non monotonic functional relation across quantiles.** Even though power prices in a predominantly thermal system will consist, at the market clearing price, of fuel prices including carbon costs (where levied, eg Bunn et al, 2007), plus the generator's profit margin, to the extent that changes in the fuel and carbon prices may cause a change in the merit order, a nonlinear response to fuel and carbon price changes could be expected, but this could also be non monotonic. For example if gas prices go down, gas generation will become baseload, increasing its low quantile coefficients, with gas no longer affecting the high quantile peakload prices. But if gas prices are higher than coal, and if scarcity in the carbon allowance market means that carbon prices are determined by fuel-switching from coal to gas, this may neutralize the impact of gas-coal price discrepancies. On the other hand, low carbon prices, that do not cause fuel switching, will have a decreasingly positive effect as price rise, if coal is cheaper than gas, and an increasingly positive effect if gas is cheaper than coal (since the carbon intensity of coal is higher). It is clear therefore that fuel prices will have an idiosyncratic pattern of influence on power prices that will not be linear, and

indeed is likely to change over time according to movements in the coal, gas and carbon markets.

4. **Adaptive behavior (manifest as a lagged price) effect will be positive and nonlinear across quantiles.** Adaptive behavior will be manifest in terms of reinforcing previously successful offers. High prices will tend to be followed with high prices. Furthermore, if there is an element of repeated gaming in power markets, as often suggested (Rothkopf, 1999), signaling between market agents will encourage this and motivate a positive coefficient for lagged prices. This will become stronger at higher quantiles, as the market becomes less competitive and gaming more possible and plausible.

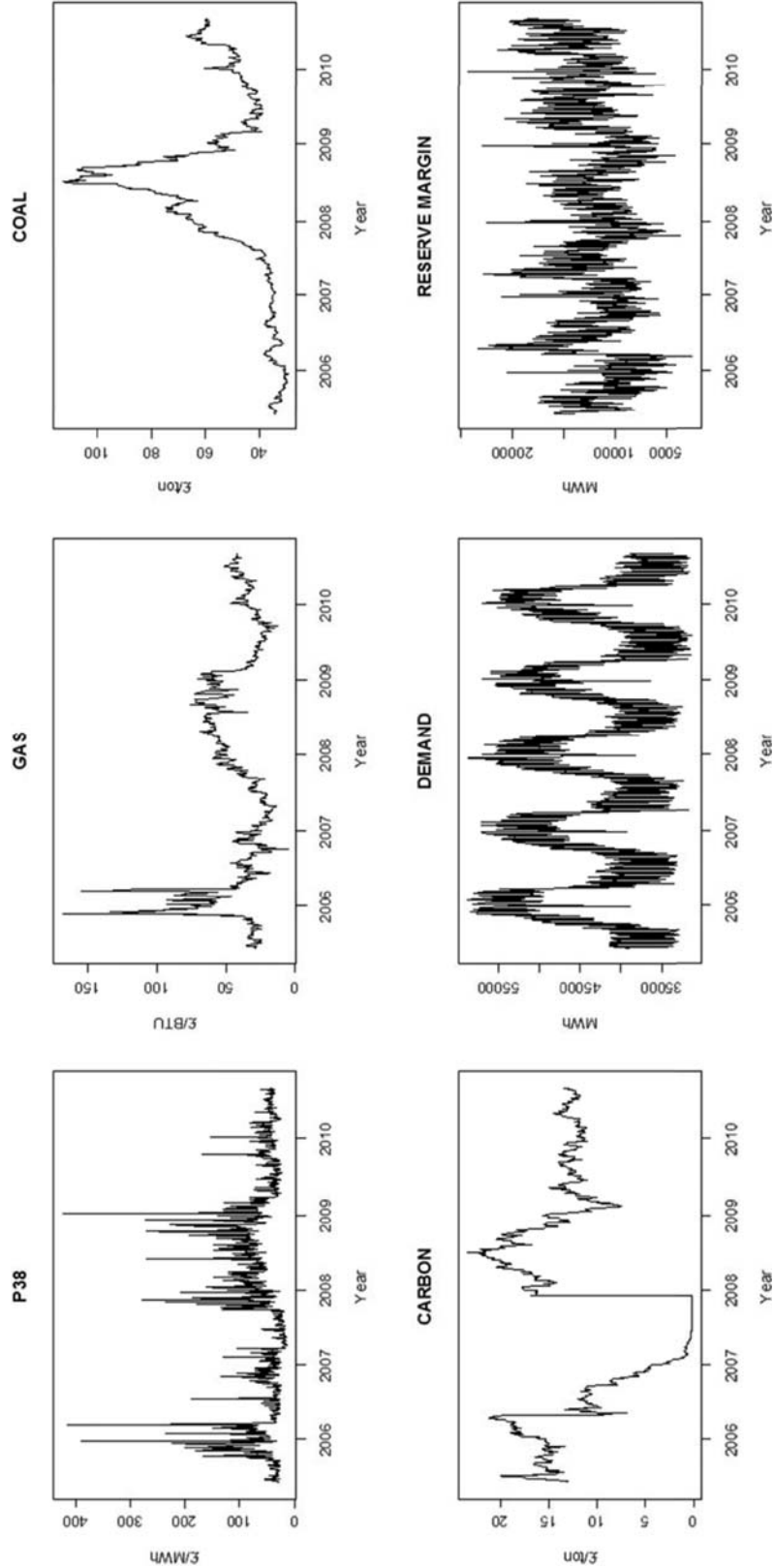
### 3. DATA

The British electricity market, liberalised in the early 1990s, is perhaps one of the most mature wholesale power markets in the world. In March 2001 the New Electricity Trading Arrangement (NETA) were implemented and introduced bilateral and voluntary forward trading in England and Wales to replace the compulsory, day-ahead uniform auction Pool that had existed since 1990. In April 2005 the British Electricity Trading and Transmission Arrangement (BETTA) extended this to include Scotland, whilst the EU carbon emissions market started at the beginning of that year. More specifically, the design of the reformed British market is based upon fully liberalised trading and plant self-scheduling, and hence, in this voluntary bilateral environment, most energy is traded with forward contracts. Close to physical delivery, agents fine tune their positions, from blocks (peak and baseload) to half-hourly resolutions, in the Power Exchanges that have emerged. These operate continuously up to 1 hour prior to each half-hourly physical delivery period, a point defined as Gate Closure, and are effectively the spot markets. After Gate Closure, the System Operator administers a market for system balancing, and invites offers and bids for load increases or decreases in real time.

In the British market, the main reference for spot trading has been the UKPX (now APX) power exchange. The spot prices are volume-weighted averages of all trades ahead of each trading period. Each day consist of 48 trading and load periods. The UKPX started its operation in March 2001. Compared to OTC contracting, UKPX is an anonymous exchange market place for trading and clearing. There are no locational prices in the British market and congestion, although dealt with in the real time system balanced by the system operator, does not contaminate the UKPX energy prices.

Our data spans 8<sup>th</sup> June 2005 to 4<sup>th</sup> September 2010. We selected period 38, corresponding to 18:30 to 19:00 in the winter (and an hour earlier in the summer), as the most challenging time series of prices for comparative analysis. It is the evening peak when household activities and the onset of darkness create the highest prices and, if supply shortages happen, so do spikes. The standard deviation of prices in this sample varies from around 11 during the night time hours, around 20 during the day to over 30 within the periods 33 to 39. Demand and demand uncertainty are also greatest during P38, with average demand in P38 in the sample being 44.9GW, and a standard deviation of 7.2GW, compared to averages across all periods of 38.7GW and 5.5GW respectively. With the onset of darkness (sunset) fluctuating at the load centre of GB between 15:53pm in winter and 21:22pm in summer, P38 is the median period. P38 has also been used in other published studies on British data (Chen et al, 2010; Bordignon et al, 2013) as being typical of the evening peak period in Britain, and therefore its choice here aids comparisons. Figure 1

**Figure 1:** Price of UKPX Period 38 (18:30–19:00) in £/MWh, UK Day Ahead Forward Gas Price (£/BTU) from the National Balancing Point, Daily Steam Coal Europe-ARA Index (translated into £/ton), EEX-EU Carbon Emission Price Daily Spot Price (translated into £/ton), the UK National Demand Forecast for Period 38 from the System Operator (MWh), the UK National Forecast of Reserve Margin for Period 38 from the System Operator (MWh)



The data spans from 8th June 2005 to 4th September 2010 (1915 observations altogether).



shows the development of the UKPX period 38 prices together with the day ahead spot prices of gas, coal, and carbon, as well as the day ahead demand and reserve margin forecasts as conveyed to the market by the system operator.

The dependent variable used in the analysis is:

**UKPX Period 38 Price.** UKPX is the day-ahead and on-the-day power exchange, allowing high frequency (half hourly) trading, up to an hour before real time. Period 38 (18:30 to 19:00 in winter; an hour earlier in summer) represents peak demand. Prices are quoted in £/MWh and represent the volume weighted prices for this period as cleared on the exchange in the preceding 24 hours.

The factors used in the analysis are all known to the market before the power exchange closes for the trading period concerned, and can therefore be considered exogenous market information for the power price formation:

**Lagged UKPX Period 38 prices.** These are the UKPX Period 38 prices lagged by one day.

**NBP Gas Price.** We use daily UK natural gas one-day forward price, from the main National Balancing Point (NBP) hub. The price is quoted in £/BTU (British Thermal Unit).

**Coal price.** We use the daily steam coal Europe-ARA (Amsterdam, Rotterdam, and Antwerp) index, taking into account the \$/£ rate.

**Carbon emission price.** We use the EEX-EU daily carbon emission allowance one year forward price taking into account the €/£ rate.

**Demand forecast.** This forecast is made available by the System Operator for each half-hourly trading period. In our study we use the period 38 demand forecast for the next day. Since it is released to the market at 18:17, one operational day ahead of closing prices, it reflects information available to participants and avoids the endogeneity issue of using actual demand.

**Reserve margin forecast (indicated margin).** The System Operator also makes forecasts of the available reserve margin for each half-hourly trading period. This is defined as the difference between the sum of the maximum available output capacities, as initially nominated by each generator prior to each trading period, and the demand forecast described above. In our study we use the period 38 reserve margin forecast, released at 18:16 for the next day.

Figure 1 (upper left exhibit) displays the period 38 price profile reflecting the “super peak” period with distinct demand, costs and operating constraints. The figure reveals typical spot electricity price features of spikes, mean reversion, seasonality, high (and time varying volatility). In Table 1 we present descriptive statistics that confirm these characteristics with a high price standard deviation, substantial skewness and kurtosis, rejection of normality according to the Jarque-Bera test, rejection of unit roots according to the ADF test and clear signs of positive serial correlation at different lags according to  $p$  and  $Q$  tests. We also show the empirical quantiles at 1%, 5%, 10%, 90%, 95%, and 99% levels which reveal high price risk in particular for consumers or traders having a short position in the period 38.

#### 4. PRICE DISTRIBUTION MODELLING AND FORECASTING

In this Section we briefly describe the theoretical framework.

**Table 1: UKPX Period 38 Prices**

Statistics	Mean	Med	Min	Max	Std	Skew	Kurt	JB	ADF	$\rho_1$	$\rho_{10}$	Q(10)
$P_t$	58.79	46.93	13.22	421.72	37.54	2.92	18.92	23040	-6.93	0.71	0.52	3268
$\ln P_t$	3.93	3.85	2.85	6.04	0.52	0.49	3.19	79	-4.55	0.84	0.72	5586
			Quantiles	1 %	5 %	10 %	90 %	95 %	99 %			
			$P_t$	18.18	24.12	28.89	98.38	130.07	194.06			
			$\ln P_t$	2.90	3.18	3.36	4.59	4.87	5.27			

The table shows the mean, median, min, max, standard deviation, skewness, excess kurtosis, Jarque-Bera, Augmented Dickey Fuller with constant and control lags according to the SIC criteria, autocorrelation at lag 1 and 10 and Ljung-Box statistics with 10 lags. We also show the empirical 1%,5%,10%,90%,95%, and 99% quantiles. Critical values at 1% level for JB is 9.21, for ADF-test -3.43, and for LB(10) 23.21.

#### 4.1 Quantile Regression Models

Quantile regression methods develop explicit models for specific quantiles of the distribution of a dependent variable, using exogenous variables with different coefficients at each quantile. Quantile regression was introduced by Koenker et al (1978) and is fully described in Koenker (2005) and Hao et al (2007). Applications in financial risk management can be found in Alexander (2008).

Let  $q \in [0,1]$  be the quantile, e.g. 1%, 5%, ..., 99%. Let  $Y_t$  be the dependent variable (e.g. log of el. price) and  $\mathbf{X}_t$  a  $d$ -dimensional vector of explanatory variables (e.g. log of the gas price, log of demand etc.), including a constant. The conditional quantile function is given by

$$Q_q(Y_t | \mathbf{X}_t) = \mathbf{X}_t \boldsymbol{\beta}_q \quad (4.1)$$

in which, as implemented in several software packages (eg Eviews, R, Stata), the parameter  $\boldsymbol{\beta}_q$  is derived according to the following optimization:

$$\arg \min_{\boldsymbol{\beta}_q} \sum_{t=1}^T (q - \mathbf{1}_{Y_t \leq \mathbf{X}_t \boldsymbol{\beta}_q}) (Y_t - \mathbf{X}_t \boldsymbol{\beta}_q) \quad (4.2)$$

where

$$\mathbf{1}_{Y_t \leq \mathbf{X}_t \boldsymbol{\beta}_q} = \begin{cases} 1 & \text{if } Y_t \leq \mathbf{X}_t \boldsymbol{\beta}_q \\ 0 & \text{otherwise} \end{cases}.$$

Details on estimating standard errors for coefficients, inference and goodness of fit can be found in Koenker et al (1999). To specify a model that incorporates the time series aspects, let  $\dot{\mathbf{X}}_t$  be a subset of information variables that become available at time  $t$  and  $\tilde{\mathbf{X}}_{t-1}$  be variables  $\{\dot{\mathbf{X}}_j\}_{j=0}^{t-1}$ , so that  $\tilde{\mathbf{X}}_t = (\dot{\mathbf{X}}_t, \tilde{\mathbf{X}}_{t-1})$ . Chernozhukov et al (2001) formulate the following general quantile regression model

$$Q_q(Y_t) = \dot{\mathbf{X}}_t \boldsymbol{\beta}_q + \gamma_q f_1(\tilde{\mathbf{X}}_t, q) \quad (4.3)$$

$$f_1(\tilde{\mathbf{X}}_t, q) = \theta_q f_1(\tilde{\mathbf{X}}_{t-1}, q) + f_2(\dot{\mathbf{X}}_t, \varphi_q).$$



Here  $f_1(\cdot)$  and  $f_2(\cdot)$  represents functions of the information set. Models of this form are useful as parsimonious regressions that represent VaR/quantiles (Chernozhukov et al (2001)). One example involves the model

$$f_1(\tilde{X}_t, q) = \sigma(Y_t - \mu_t | \tilde{X}_t),$$

where  $\sigma^2(Y_t - \mu_t | \tilde{X}_t)$  is the conditional variance of the de-meaned  $Y_t$ . It can take the form of an ARCH model, see Koenker et al (1996). As suggested by Chernozhukov et al (2001), a simpler strategy is to first estimate  $\sigma^2(Y_t - \mu_t | \tilde{X}_t)$  via a GARCH model and use it as a regressor in the linear model. A similar approach is also done in Xiao et al (2009). They first use quantile regression to estimate the volatility, and then include lagged volatility and the lagged dependent variable as explanatory variables in the linear quantile regression model. They establish the convergent efficiency of this approach.

The CAViaR class of models also falls into this framework. Let  $\varepsilon_t$  be the de-meaned process,  $\varepsilon_t = Y_t - \mu_t$ . Engle et al (2004) suggest four different CAViaR specifications, all of which are first-order autoregressive. The Indirect GARCH(1,1) CAViaR model,

$$Q_q(\varepsilon_t) = (1 - 2I(q < 0.5))(\alpha_1 + \alpha_2 Q_q(\varepsilon_{t-1})^2 + \alpha_3 \varepsilon_{t-1}^2)^{0.5}. \quad (4.4)$$

The Symmetric Absolute Value CAViaR model,

$$Q_q(\varepsilon_t) = \alpha_1 + \alpha_2 Q_q(\varepsilon_{t-1}) + \alpha_3 |\varepsilon_{t-1}|. \quad (4.5)$$

The Asymmetric Slope CAViaR model,

$$Q_q(\varepsilon_t) = \alpha_1 + \alpha_2 Q_q(\varepsilon_{t-1}) + \alpha_3 |\varepsilon_{t-1}| I(q \geq 0.5) + \alpha_4 |\varepsilon_{t-1}| I(q < 0.5). \quad (4.6)$$

The Adaptive CAViaR model,

$$Q_q(\varepsilon_t) = Q_q(\varepsilon_{t-1}) + \alpha_1 [q - [1 + \exp(K(\varepsilon_{t-1} - Q_q(\varepsilon_{t-1})))]^{-1}]. \quad (4.7)$$

Here  $Q_q(\varepsilon_t)$  is the  $q$ -percentile of the de-meaned price distribution at time  $t$ ,  $I$  is a indicator function and  $K$  is a smoothing parameter which may be chosen or estimated. Quantile regression models are robust to distributional misspecifications, as no explicit distributional assumptions need to be made. In fact, CAViaR models can be used for situations with constant volatilities but changing error distributions, or situations in which both error densities and volatilities are changing (Engle et al (2004)).

## 4.2 Fully Parametric Location-scale Models

Fully parametric models are among the most commonly used market risk measures, and are typically based on the assumption that the price distribution can be described by a parametric density function together with a model for the conditional variance. Popular specifications of the conditional density are Gaussian or skew Student-t, and a GARCH model for the conditional variance. As a methodological benchmark for the quantile regression models, we use location-scale

models where the volatility dynamics follows a GARCH(1,1) structure. That is, we assume that the distribution of the dependent variable can be expressed in the form

$$Y_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t z_t$$

$$z_t \stackrel{iid}{\sim} f_Z(\cdot)$$
(4.8)

Here  $\mu_t$  is the mean of  $Y_t$ ,  $\sigma_t = \sqrt{\alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2}$ , with  $f_Z$  is a zero-location, unit-scale probability density that can have additional shape parameters. The one-step forecast of the  $q$ -percentile of  $Y_t$  based on information up to time  $t$  is given by

$$Q_q(Y_{t+1}) = \hat{\mu}_{t+1} + \hat{\sigma}_{t+1} Q_q(z)$$
(4.8)

where  $Q_q(z)$  is the  $q$ -percentile implied by  $f_Z$ . Approaches differ with respect to the specification of the conditional mean,  $\mu_t$ , and conditional density  $f_Z$ . Note that the Indirect GARCH(1,1) CAViaR model (4.4) corresponds to a model on the form given by (4.8), but with zero location parameter,  $\mu_t \equiv 0$ , and an iid error distribution.

### 4.3 Forecasting

To validate the predictive performance of the models, we consider the unconditional test of Kupiec (1995), the conditional coverage test of Christoffersen (1998) and two backtests based on regression. The Kupiec (1995) test is a likelihood ratio test designed to reveal whether the model provides the correct unconditional coverage. More precisely, let  $\{H_t\}_{t=1}^T$  be an indicator sequence where  $H_t$  takes the value 1 if the observed price,  $Y_t$ , is below the predicted quantile,  $Q_t$ , at time  $t$

$$H_t = \begin{cases} 1 & \text{if } Y_t < Q_t \\ 0 & \text{if } Y_t \geq Q_t \end{cases}$$
(4.9)

Under the null hypothesis of correct unconditional coverage the test statistic is

$$LR_{uc} = -2 \log \left[ \frac{(1 - \pi_{exp})^{n_0} \pi_{exp}^{n_1}}{(1 - \pi_{obs})^{n_0} \pi_{obs}^{n_1}} \right] \stackrel{asy}{\sim} \chi_1^2$$

where  $n_1$  and  $n_0$  is the number of violations and non-violations respectively,  $\pi_{exp}$  is the expected proportion of exceedances and  $\pi_{obs} = n_1/(n_0 + n_1)$  is the observed proportion of exceedances. In the Kupiec (1995) test only the total number of ones in the indicator sequence  $\{H_t\}_{t=1}^T$  counts, and the test does not take into account whether several quantile exceedances occur in rapid succession, or whether they tend to be isolated. Christoffersen (1998) provides a joint test for correct coverage and for detecting whether a quantile violating today has influence on the probability of a violating tomorrow. The test statistic is defined as follows:

$$LR_{cc} = -2 \log \left[ \frac{\pi_{exp}^{n_1} (1 - \pi_{exp})^{n_0}}{(1 - \pi_{01})^{n_{00}} \pi_{01}^{n_{01}} (1 - \pi_{11})^{n_{10}} \pi_{11}^{n_{11}}} \right] \stackrel{asy}{\sim} \chi_2^2$$

where  $n_{ij}$  represents the number of times an observations with value  $i$  is followed by an observation with value  $j$ .  $\pi_{01} = n_{01}/(n_{00} + n_{01})$  and  $\pi_{11} = n_{11}/(n_{11} + n_{10})$ . The  $LR_{cc}$  test is only sensitive to one violation immediately followed by another, ignoring all other patterns of clustering.

The Kupiec (1995) and Christoffersen (1998) tests only use information on past quantile violations, and therefore might not have sufficient power to detect misspecified risk models. To increase the power we may also want to consider whether violations can be predicted by including other data in the information set such as previous outcomes or the quantile estimate for the period itself. The advantage of increasing the information set is not only to increase power, but also to help us understand the areas in which the risk model is misspecified (Christoffersen (2010)). We consider the following two regression based backtests, similar to Engle et al (2004);

$$H_t = \beta_0 + \beta_1 H_{t-1} + \beta_2 H_{t-2} + \beta_3 H_{t-3} + \beta_4 H_{t-4} + \varepsilon_t \quad (4.10)$$

$$H_t = \beta_0 + \beta_1 H_{t-1} + \beta_2 H_{t-2} + \beta_3 H_{t-3} + \beta_4 H_{t-4} + \beta_5 Q_t + \varepsilon_t. \quad (4.11)$$

Here the indicator variable  $H_t$  is defined in Equation (4.9), and  $Q_t$  denotes the quantile estimate itself. The four lags were the longest periods of significance in the data. In Equation (4.10) we test the hypotheses  $H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = 0$  against the alternative that at least one of  $\beta_1, \beta_2, \beta_3, \beta_4$  is significant different from zero using a standard  $F$  test. This provides insights into clustering. In Equation (4.11) we test  $H_0: \beta_5 = 0$  whether quantile exceedances, under circumstances where there may be clustering, are also linked to the scale of the quantile forecasts using a simple  $t$  test.

## 5. EMPIRICAL ANALYSIS

First we conduct an empirical analysis of the price sensitivities with respect to the factors described in Section 3. Then we evaluate the ability of different methods to forecast the tails of the price distribution.

### 5.1 In Sample Modelling Using Quantile Regression

We perform in-sample analysis using all data from 9<sup>th</sup> June 2005 to 4<sup>th</sup> September 2010 which consist of 1915 observations for the period 38 UK electricity prices. The coefficients of lagged prices, lagged gas/coal/carbon prices, forecast of demand and reserve margin are investigated at the quantiles 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95%, and 99%. In order to variance-stabilise the data and interpret all parameters as elasticities, we log-transform both the dependent and independent variables, as in equation (5.1).

$$Q_q(\ln P38_t) = \beta_0^q + \beta_1^q \ln P38_{t-1} + \beta_2^q \ln Gas_{t-1} + \beta_3^q \ln Coal_{t-1} + \beta_4^q \ln Carbon_{t-1} + \beta_5^q \ln Demand_t + \beta_6^q \ln Reserve_t + \beta_7^q \sigma(\ln P38_t - \mu_t) \quad (5.1)$$

Here  $\sigma^2(\ln P38_t - \mu_t)$  is the conditional variance of the de-meaned log prices, and is estimated using *Model 4* described in Section 5.2.1 below. The index  $q$  refers to the specific quantile. All calculations are performed in EViews (The QREG procedure). We recognise that this is a reduced form representation, but structural estimation from a quantile formulation is beyond the scope of this research and our approach is comparable to many single period models used elsewhere in research on electricity price formation.

**Table 2: Quantile Regression Results**

Quantile	lag P38	Gas	Coal	Carbon	Demand	Reserve Margin	Volatility	R <sup>2</sup> -adjusted
1 %	0.22***	0.31***	0.31***	0.07***	0.08	−0.37***	−0.22***	49.0
5 %	0.28***	0.27***	0.33***	0.06***	0.23***	−0.28***	−0.11***	53.6
10 %	0.31***	0.27***	0.31***	0.05***	0.24***	−0.27***	−0.02	55.2
25 %	0.38***	0.23***	0.30***	0.04***	0.30***	−0.27***	−0.02	57.5
50 %	0.47***	0.19***	0.27***	0.03***	0.25***	−0.35***	0.06**	58.9
75 %	0.55***	0.20***	0.22***	0.02***	0.26***	−0.46***	0.02	58.2
90 %	0.59***	0.20***	0.16***	0.01	0.29***	−0.54***	0.04	58.9
95 %	0.50***	0.20***	0.22***	0.02***	0.34**	−0.63***	0.21***	59.8
99 %	0.35**	0.26*	0.30	0.02	0.42	−0.86***	0.16	58.3

\*, \*\* and \*\*\* indicate significance at the 10%, 5% or 1% level, respectively.

In table 2 we show the parameter values at different quantiles for model (5.1). The explanatory power measured by Koenker et al (1999) pseudo R-squared is in the range of 49% to 59%.

**Lagged prices.** The significance and sign of the lagged electricity price is consistent with mean-reversion (positive serial correlation for prices and negative serial correlation for returns). Except for the 95% and 99% quantiles we find increasing serial correlation for higher prices.

**Gas prices.** The gas price coefficients are generally positive in line with the previous discussion of supply function fundamentals. Using logtransform prices, the coefficients are in the range of 0.19–0.31 and all significant. There is no clear pattern in the coefficient values across the different quantiles. This is consistent with our previous proposition about the potentially nonmonotonic effects of fuel prices.

**Coal prices.** The coal price coefficients are generally positive and also here in line with the previous discussion of supply function fundamentals. The sensitivities values are generally higher than gas. Low and moderate quantiles are more sensitive than high quantiles, which may reflect the fact that low prices are more likely to be driven by fuel fundamentals than high prices, where scarcity may be more determinate.

**Carbon emission prices.** The carbon emission price coefficients are generally positive. Their magnitude is rather low (in the 0.01 to 0.07) range, decreasing with price levels.

**Volatility.** The coefficient of volatility changes sign from negative (low prices) to positive (high prices). During times of low prices, an increase in volatility tends to drive prices even lower than we can explain using fundamental factors alone. When prices are high, an increase in volatility tends to drive prices even higher. This suggests that both low and high electricity prices overshoot the fundamentals when the price uncertainty is high, which is a remarkable but plausible observation.

**Demand forecast.** The generally positive and increasing sensitivities with higher prices reflects the intuitive price and demand relationships from the increasing supply function discussed earlier. The coefficients are all significant apart from the 1% and 99% quantile for logtransform prices.

**Reserve margin forecast.** The negative signs for margin intuitively reflect the fact that the lower reserve margin, the higher the price becomes. The effect is also increasing with higher prices as one would expect according to the supply function and scarcity considerations. All parameters are significant.

It should be noted that other exogenous variables (such as several seasonal dummies and trading volume) were tested initially but were not found significant. Seasonal effects are to a large extent captured in forecasted demand.

## 5.2 Forecasting Tail Probabilities

VaR modelling requires accuracy in the forecasting of the tails of the price density rather than in the main body of the price distribution.

### 5.2.1 Fully Parametric Location Scale-models

From the fully parametric location-scale models we use two different conditional densities; Gaussian and skew Student-t. In addition, we consider two different specifications of the conditional mean; (1)  $\mu_t$  is a linear function of the lagged prices, and (2)  $\mu_t$  is a linear function of the lagged price,  $\ln P_{t-1}$ , and the factors described in Section 3.

*Models 1–2:*

$$\ln P38_t = \beta_0 + \beta_1 \ln P38_{t-1} + \beta_2 \ln P38_{t-2} + \beta_3 \ln P38_{t-3} + \beta_4 \ln P38_{t-4} + \beta_5 \ln P38_{t-5} + \beta_6 \ln P38_{t-6} + \beta_7 \ln P38_{t-7} + \sigma_t z_t$$

Here the volatility follows a GARCH(1,1) process,  $\sigma_t = \sqrt{\alpha_0 + \alpha_1 \sigma_{t-1}^2 + \alpha_2 \varepsilon_{t-1}^2}$ , and the conditional density  $f_Z(\cdot)$  is set to Gaussian (Model 1) and skew Student-t (Model 2), respectively.

*Models 3–4:*

$$\ln P38_t = \beta_0 + \beta_1 \ln P38_{t-1} + \beta_2 \ln \text{Gas}_{t-1} + \beta_3 \ln \text{Coal}_{t-1} + \beta_4 \ln \text{Carbon}_{t-1} + \beta_5 \ln \text{Demand}_t + \beta_6 \ln \text{Reserve}_t + \sigma_t z_t$$

As above,  $\sigma_t = \sqrt{\alpha_0 + \alpha_1 \sigma_t^2 + \alpha_2 \varepsilon_{t-1}^2}$ , and  $f_Z(\cdot)$  is Gaussian (Model 3) and skew Student-t (Model 4), respectively. The estimation of models 1–4 is done in two steps. First the  $\beta$ -parameters are estimated using an ordinary least square (OLS) regression where log prices is regressed against explanatory variables. Then a GARCH(1,1) model is fitted to the residuals from the regression.<sup>1</sup>

### 5.2.2 Quantile Regression Models

From the quantile regression class of models, being semi parametric without a specified error term, the quantile distributions come from quantile dependent regression coefficients, of which, we use the following functional forms:

*Model 5:* A linear model with only lagged prices as explanatory variables

$$Q_q(\ln P38_t) = \beta_q^0 + \beta_q^1 \ln P38_{t-1} + \beta_q^2 \ln P38_{t-2} + \beta_q^3 \ln P38_{t-3} + \beta_q^4 \ln P38_{t-4} + \beta_q^5 \ln P38_{t-5} + \beta_q^6 \ln P38_{t-6} + \beta_q^7 \ln P38_{t-7}$$

1. Whilst GLS is generally preferred for estimating the standard errors of the regression coefficients, the first stage OLS provides unbiased conditional expectations of the price and that was sufficient for the purposes of this study. We did not need to undertake a process of in-sample regressor elimination according to significance levels and the comparative modelling evaluations were mainly based upon out-of-sample backtesting.

*Model 6:* A linear model with lagged price and fundamental explanatory variables

$$Q_q(\ln P38_t) = \beta_0^q + \beta_1^q \ln P38_{t-1} + \beta_2^q \ln Gas_{t-1} + \beta_3^q \ln Coal_{t-1} + \beta_4^q \ln Carbon_{t-1} + \beta_5^q \ln Demand_t + \beta_6^q \ln Reserve_t$$

*Model 7:* A linear model with lagged price, fundamental variables and volatility as explanatory variables

$$Q_q(\ln P38_t) = \beta_0^q + \beta_1^q \ln P38_{t-1} + \beta_2^q \ln Gas_{t-1} + \beta_3^q \ln Coal_{t-1} + \beta_4^q \ln Carbon_{t-1} + \beta_5^q \ln Demand_t + \beta_6^q \ln Reserve_t + \beta_7^q \sigma(\ln P38_t - \mu_t)$$

Here  $\sigma^2(\ln P38_t - \mu_t)$  is the conditional variance of the de-measured log prices, and is estimated using *Model 4* described above. That is, the estimation is performed in two steps; first a GARCH models is used to estimate the volatility, next, treating volatility as an observed variable, the linear quantile regression is estimated.

*Models 8–11:*

$$Q_q(\ln P38_t) = \beta_0 + \beta_1 \ln P38_{t-1} + \beta_2 \ln Gas_{t-1} + \beta_3 \ln Coal_{t-1} + \beta_4 \ln Carbon_{t-1} + \beta_5 \ln Demand_t + \beta_6 \ln Reserve_t + Q_q(\varepsilon_t)$$

Here the residual error terms,  $\varepsilon_t$ , define a CAViaR process specified around same multifactor conditional expectation as we used in models 3 and 4. Observe, that, as with models 1–4, the beta coefficients are not specific to the quantiles. More precisely, in *Model 8* an indirect GARCH(1,1) CAViaR model is chosen for  $\varepsilon_t$ , in *Model 9* a symmetric absolute value CAViaR model, in *Model 10* a asymmetric slope CAViaR and finally *Model 11* the adaptive CAViaR model is chosen for the error term (for the adaptive model, we follow Engle et al (2004) and set  $K = 10$ ).

*Models 8–11* are estimated in two steps. Since conventional CAViaR specifications do not include exogenous variables, we proceed by first estimating  $\beta$  using OLS regression, and then a CAViaR model is fitted to the residuals from the regression. The first stage conditional expectation is therefore the same as in models 1–4, and this also helps the modelling comparison to be based purely upon the volatility aspect. Estimation of the CAViaR models is complicated by the fact that the quantiles are latent and are dependent on the unknown parameters. We use the Matlab code of Manganelli (2002) to estimate the CAViaR models. An alternative estimating strategy is Markov Chain Monte Carlo, as in Gerlach et al. (2011).

### 5.2.3 Empirical Results

It is important to validate and compare the models in an out-of-sample forecasting context, as it is well-known that elaborate, well-specified ex-post models may not forecast better than simpler, more robust models, because of overfitting and/or evolutionary changes in the price formation process. In contrast, the usefulness of well specified ex-post models is more often argued for their value in ex-post market performance analysis, e.g., market monitoring, and ex ante, for facilitating multiple scenario simulations, e.g., for risk management. The ultimate aim is of course to have a well specified model both in-sample and out-of-sample. We use two approaches for in-sample and out-of-sample:



**Expanding window (EW) in sample.** Here we estimate the models using the first 730 observations. We then forecast quantiles (1%, 5%, 10%, 90%, 95%, and 99%) of observation 731. Thereafter we estimate the models with the first 731 observations. We then forecast quantiles of observation 732 and so on. At the end, we estimate models with the first 1914 observations and forecast quantiles of the last observation 1915. That will leave us with  $1915 - 730 = 1185$  observations to verify tail forecasting performance.

**Rolling window (RW) in sample.** We start out estimating the models using the first 730 observations. We then forecast quantiles (1%, 5%, 10%, 90%, 95%, and 99%) of observation 731. Thereafter estimate the models using observation 2 to 731. We then forecast quantiles of observation 732 and so on. At the end, we estimate the models using observation 1185 to 1914 and forecast quantiles of the last observation 1915. Again, this will leave us with  $1915 - 730 = 1185$  observations to verify tail forecasting performance.

Tables 3–5 report the percentage of times the observed price is below the estimated quantile, the  $p$ -values for the unconditional coverage test by Kupiec (1995), the  $p$ -values for the conditional coverage test by Christoffersen (1998), and the  $p$ -values from the two different regression based tests defined by (4.10) and (4.11). Note that the 5% significance tests have the null hypotheses of correct %, so ideally the actual exceedences should *not* be significantly different. In other words we would not like to see bold ( $p < 5\%$ ) significantly different exceedences in the results. We also record the average values of the quantile estimates. A good risk model method should not only pass the calibration tests described above, it should also provide narrow prediction intervals, as the width of the intervals is linked to the precision of using the method in practice. Gneiting et al (2007) refer to “sharpness” as well as “calibration” as desirable for the assessment of density forecasts.

The overall results from the tests suggest that the parametric location-scale models based on Gaussian distribution are seriously flawed, failing about half of the tests. Using a skew Student-t distribution leads to clear improvements, and the models using this distribution estimated under a rolling window has good performance with no significantly difference exceedences under the UC and CC coverage tests at  $p = 5\%$ . The linear quantile regression models including only lagged prices (Model 5) do not provide satisfactory forecasts of the quantiles, showing performance on par with the Gaussian based parametric location-scale model. Introducing fundamental factors in the linear quantile regression model (Model 6) significantly improves the results, underpinning the importance of these risk factors in predicting the tail probabilities of the electricity spot price. Adding volatility as an explanatory variable (Model 7) improves the tail predictions further, indicating that volatility is not adequately encapsulated through the factors in the fundamental model. This is the only model to have no significantly different exceedences under the expanding window forecasting. The CAViaR models show relatively good results, with the Adaptive CAViaR (Model 11) performing well on the coverage tests with a rolling window, but not so well with an expanding .

Going into more details for each of the tests, we find, using a 5% significance level and an expanding window, that only the linear quantile regression model with fundamental factors and volatility as explanatory variables (Model 7) provide the correct percentage of exceedences. Fitting the models to a rolling window generally improves the unconditional coverage, indicating that the data generating process may change over time. Using a rolling window, Model 4, 6, 7, 8 and 11 all provide correct unconditional coverage. Examining the clustering of exceedences of the quantiles, we find that the location scale GARCH models all perform very well on the regression based test defined by (4.10). The joint test of unconditional coverage and independence of Christoffersen (1998) shows less encouraging results, but this is generally caused by incorrect percentage of violations. The same conclusion can be drawn for the CAViaR models except the Adaptive model,

**Table 3: Forecasting Results Out of Sample for GARCH Models with Expanding Window (left) and Rolling Window (right)**

Model	Quantile	Expanding Window					Rolling Window						
		Violations	LR <sub>UC</sub>	LR <sub>CC</sub>	REG <sub>Hit</sub>	REG <sub>Var</sub>	Quantile	Violations	LR <sub>UC</sub>	LR <sub>CC</sub>	REG <sub>Hit</sub>	REG <sub>Var</sub>	Quantile
Gaussian, lag prices (Model 1)	1	0.00 %	NA	<b>0.000</b>	NA	0.056	3.422	0.00 %	NA	<b>0.000</b>	NA	0.054	3.432
	5	1.35 %	<b>0.000</b>	<b>0.000</b>	0.921	0.776	3.594	1.94 %	<b>0.000</b>	<b>0.000</b>	0.672	0.559	3.603
	10	5.06 %	<b>0.000</b>	<b>0.000</b>	0.469	<b>0.003</b>	3.686	5.91 %	<b>0.000</b>	<b>0.000</b>	0.158	<b>0.000</b>	3.695
	90	89.11 %	0.315	0.242	0.745	0.143	4.335	89.62 %	0.665	0.695	0.889	0.093	4.338
	95	93.16 %	<b>0.006</b>	<b>0.020</b>	<b>0.039</b>	<b>0.040</b>	4.427	92.74 %	<b>0.001</b>	<b>0.003</b>	0.238	0.091	4.429
	99	97.47 %	<b>0.000</b>	<b>0.000</b>	0.644	0.561	4.599	97.38 %	<b>0.000</b>	<b>0.000</b>	0.787	0.522	4.601
Skew Student-t, lag prices (Model 2)	1	0.42 %	<b>0.024</b>	0.076	0.999	<b>0.048</b>	3.498	0.25 %	<b>0.002</b>	<b>0.009</b>	1.000	0.147	3.509
	5	4.05 %	0.122	0.214	0.814	<b>0.045</b>	3.665	4.39 %	0.260	0.503	0.782	0.257	3.669
	10	9.03 %	0.258	0.432	0.433	<b>0.001</b>	3.737	9.45 %	0.526	0.540	0.613	<b>0.001</b>	3.740
	90	88.35 %	0.065	<b>0.031</b>	0.543	0.400	4.325	88.61 %	0.117	0.239	0.873	0.371	4.334
	95	93.67 %	<b>0.043</b>	0.080	0.065	0.148	4.457	94.01 %	0.128	0.130	0.178	0.055	4.464
	99	98.90 %	0.741	0.811	0.968	0.996	4.783	98.99 %	0.965	0.875	0.977	0.841	4.776
Gaussian, fundamental (Model 3)	1	0.51 %	0.059	0.163	0.998	<b>0.023</b>	3.461	0.68 %	0.232	0.461	0.994	<b>0.004</b>	3.471
	5	2.19 %	<b>0.000</b>	<b>0.000</b>	<b>0.012</b>	<b>0.000</b>	3.620	2.45 %	<b>0.000</b>	<b>0.000</b>	0.071	<b>0.000</b>	3.630
	10	6.41 %	<b>0.000</b>	<b>0.000</b>	0.976	<b>0.000</b>	3.705	6.33 %	<b>0.000</b>	<b>0.000</b>	0.647	<b>0.000</b>	3.714
	90	90.04 %	0.961	0.397	0.297	<b>0.023</b>	4.304	90.55 %	0.526	0.635	0.471	<b>0.003</b>	4.312
	95	94.43 %	0.377	0.593	0.918	0.268	4.389	94.43 %	0.377	0.316	0.216	0.322	4.396
	99	97.72 %	<b>0.000</b>	<b>0.001</b>	0.771	0.926	4.548	97.55 %	<b>0.000</b>	<b>0.000</b>	0.359	0.335	4.555
Skew Student-t, fundamental (Model 4)	1	1.01 %	0.965	0.875	0.970	<b>0.000</b>	3.520	1.10 %	0.741	0.811	0.217	<b>0.001</b>	3.537
	5	3.71 %	<b>0.034</b>	<b>0.018</b>	0.210	<b>0.000</b>	3.669	4.98 %	0.973	0.784	0.884	<b>0.000</b>	3.683
	10	8.44 %	0.066	<b>0.015</b>	0.970	<b>0.000</b>	3.739	10.30 %	0.736	0.072	0.185	<b>0.000</b>	3.751
	90	89.87 %	0.885	0.334	0.380	<b>0.011</b>	4.303	90.04 %	0.961	0.586	0.464	<b>0.026</b>	4.308
	95	95.36 %	0.567	0.754	0.870	0.506	4.415	95.44 %	0.478	0.705	0.353	0.371	4.422
	99	98.82 %	0.542	0.694	0.950	0.898	4.667	99.07 %	0.802	0.866	0.980	0.811	4.680

Table shows the  $p$ -values of the backtests defined in Section 4.3.  $\text{REG}_{\text{Hit}}$  and  $\text{REG}_{\text{Var}}$  are the regression based tests given by Eq. (4.10) and (4.11), respectively. Bold type  $p$ -values are significant at the 5% level, which implies poor calibration. Quantile denotes the average value of the quantile estimates.

Table 4: Forecasting Results Out of Sample for the Linear Quantile Regression Models with Expanding Window (left) and Rolling Window (right)

Model	Quantile	Expanding Window					Rolling Window						
		Violations	LR <sub>UC</sub>	LR <sub>CC</sub>	REG <sub>Hitt</sub>	REG <sub>Var</sub>	Quantile	Violations	LR <sub>UC</sub>	LR <sub>CC</sub>	REG <sub>Hitt</sub>	REG <sub>Var</sub>	Quantile
Linear quantile regression, lag prices (Model 5)	1	0.68 %	0.232	0.461	<b>0.000</b>	<b>0.007</b>	3.393	1.01 %	0.965	0.875	<b>0.001</b>	<b>0.001</b>	3.563
	5	2.87 %	<b>0.000</b>	<b>0.001</b>	<b>0.030</b>	<b>0.006</b>	3.598	3.54 %	<b>0.016</b>	<b>0.003</b>	<b>0.001</b>	0.059	3.707
	10	6.24 %	<b>0.000</b>	<b>0.000</b>	<b>0.001</b>	0.402	3.697	8.02 %	<b>0.019</b>	<b>0.000</b>	<b>0.000</b>	0.122	3.768
	90	91.90 %	<b>0.025</b>	<b>0.022</b>	0.354	0.308	4.357	91.81 %	<b>0.032</b>	0.069	0.169	0.911	4.301
	95	96.20 %	<b>0.048</b>	0.132	0.992	0.423	4.520	95.95 %	0.122	0.221	0.748	0.163	4.403
	99	98.90 %	0.741	0.297	<b>0.002</b>	<b>0.023</b>	4.872	98.99 %	0.965	0.875	0.112	<b>0.008</b>	4.697
Linear quantile regression, fundamental (Model 6)	1	0.42 %	<b>0.024</b>	0.076	0.999	0.350	3.519	1.43 %	0.158	<b>0.026</b>	<b>0.001</b>	0.112	3.418
	5	3.88 %	0.067	0.058	0.081	0.118	3.694	5.40 %	0.532	<b>0.021</b>	<b>0.000</b>	0.096	3.615
	10	8.10 %	<b>0.025</b>	<b>0.022</b>	0.057	<b>0.041</b>	3.755	10.97 %	0.272	<b>0.003</b>	<b>0.000</b>	0.071	3.708
	90	88.69 %	0.141	0.136	0.284	0.098	4.285	89.70 %	0.736	0.834	0.368	0.320	4.362
	95	94.26 %	0.254	0.433	0.962	0.158	4.394	94.60 %	0.532	0.553	0.439	0.128	4.515
	99	98.73 %	0.377	0.273	0.119	0.431	4.647	98.65 %	0.250	0.409	0.921	<b>0.006</b>	4.854
Linear quantile regression, fundamental and volatility (Model 7)	1	0.76 %	0.385	0.635	<b>0.011</b>	0.276	3.559	1.35 %	0.250	0.232	0.169	0.151	3.581
	5	4.30 %	0.260	0.262	<b>0.043</b>	0.095	3.698	5.99 %	0.128	0.064	<b>0.001</b>	0.105	3.710
	10	8.52 %	0.083	0.099	0.201	<b>0.040</b>	3.757	10.55 %	0.532	0.388	<b>0.002</b>	<b>0.044</b>	3.767
	90	89.03 %	0.272	0.202	0.374	0.174	4.288	89.45 %	0.532	0.687	0.534	0.224	4.299
	95	94.77 %	0.716	0.654	0.435	0.206	4.398	94.94 %	0.921	0.520	0.830	0.271	4.417
	99	98.73 %	0.377	0.551	0.385	0.507	4.647	98.40 %	0.055	<b>0.017</b>	<b>0.028</b>	<b>0.002</b>	4.690

Table shows the  $p$ -values of the backtests defined in Section 4.3.  $\text{Reg}_{\text{fit}}$  and  $\text{Reg}_{\text{var}}$  are the regression based tests given by Eq. (4.10) and (4.11), respectively. Bold type  $p$ -values are significant at the 5% level, which implies poor calibration. Quantile denotes the average value of the quantile estimates.

Table 5: Forecasting Results Out of Sample for the CAViaR Models with Expanding Window (left) and Rolling Window (right)

Model	Quantile	Expanding Window					Rolling Window				
		Violations	LR <sub>UC</sub>	LR <sub>CC</sub>	REG <sub>Hit</sub>	REG <sub>Var</sub>	Violations	LR <sub>UC</sub>	LR <sub>CC</sub>	REG <sub>Hit</sub>	REG <sub>Var</sub>
Indirect GARCH(1,1) CaViaR (Model 8)	1	1.18 %	0.542	0.694	0.951	<b>0.001</b>	3.522	0.095	<b>0.003</b>	0.900	<b>0.000</b>
	5	3.71 %	<b>0.034</b>	0.086	0.760	<b>0.000</b>	3.673	0.532	<b>0.031</b>	0.255	<b>0.000</b>
	10	8.44 %	0.066	0.145	0.982	<b>0.000</b>	3.740	0.219	<b>0.036</b>	0.120	0.308
	90	89.79 %	0.809	0.471	0.606	0.195	4.301	0.734	0.714	<b>0.027</b>	0.186
	95	94.77 %	0.716	0.878	0.500	0.556	4.410	0.662	0.341	<b>0.046</b>	0.777
	99	98.90 %	0.741	0.811	0.963	0.977	4.647	0.158	0.183	<b>0.003</b>	0.602
Symmetric absolute Value CAViaR, (Model 9)	1	1.35 %	0.250	0.409	0.462	<b>0.001</b>	3.543	<b>0.030</b>	0.060	<b>0.022</b>	<b>0.000</b>
	5	3.88 %	0.067	0.143	0.432	<b>0.000</b>	3.677	0.450	0.532	0.372	<b>0.000</b>
	10	7.85 %	<b>0.011</b>	<b>0.035</b>	0.988	<b>0.000</b>	3.741	<b>0.032</b>	<b>0.049</b>	<b>0.038</b>	<b>0.000</b>
	90	89.96 %	0.961	0.730	0.435	0.337	4.294	0.736	0.768	0.207	0.758
	95	94.26 %	0.254	0.433	0.766	0.923	4.405	0.254	0.283	0.138	0.768
	99	98.90 %	0.741	0.297	0.219	0.328	4.669	0.377	0.551	0.301	0.785
Asymmetric slope CAViaR, (Model 10)	1	0.93 %	0.802	0.223	0.076	<b>0.002</b>	3.533	<b>0.016</b>	<b>0.037</b>	<b>0.043</b>	<b>0.000</b>
	5	3.80 %	<b>0.048</b>	0.113	0.395	<b>0.000</b>	3.673	0.816	0.818	0.375	<b>0.000</b>
	10	8.19 %	<b>0.032</b>	0.086	0.540	<b>0.000</b>	3.749	0.809	0.470	0.248	<b>0.000</b>
	90	89.79 %	0.809	0.471	0.203	0.333	4.298	0.272	0.453	0.260	0.683
	95	95.44 %	0.478	0.406	0.110	0.225	4.413	0.973	0.789	0.455	0.547
	99	98.90 %	0.741	0.811	0.963	0.846	4.719	0.158	0.284	0.902	0.587
Adaptive CAViaR, (Model 11)	1	0.68 %	0.232	0.461	0.994	<b>0.000</b>	3.385	0.965	0.875	0.972	<b>0.000</b>
	5	3.12 %	<b>0.002</b>	<b>0.005</b>	0.221	<b>0.000</b>	3.651	0.160	0.285	0.722	<b>0.000</b>
	10	7.68 %	<b>0.006</b>	<b>0.014</b>	0.103	<b>0.000</b>	3.736	0.303	0.494	0.143	<b>0.000</b>
	90	90.30 %	0.734	0.162	<b>0.001</b>	<b>0.005</b>	4.293	0.884	0.069	<b>0.000</b>	<b>0.028</b>
	95	96.29 %	<b>0.034</b>	<b>0.026</b>	0.100	0.337	4.461	0.478	0.239	<b>0.008</b>	0.271
	99	99.16 %	0.579	0.781	<b>0.036</b>	0.629	4.764	0.741	0.811	<b>0.040</b>	0.949

Table shows the  $p$ -values of the backtests defined in Section 4.3.  $\text{Reg}_{\text{Hit}}$  and  $\text{Reg}_{\text{Var}}$  are the regression based tests given by Eq. (4.10) and (4.11), respectively. Bold type  $p$ -values are significant at the 5% level, which implies poor calibration. Quantile denotes the average value of the quantile estimates.

which shows more clustering than the other CAViaR models. This is not surprising, since the Adaptive model increases the quantile by the same amount regardless of whether the size of the residual term exceeded the quantile by a small or a large margin. The linear quantile regression models, especially the model including only lagged prices, show in general somewhat more clustering than the other methods. The reason may be the same as for the Adaptive model; the quantile does not depend directly on the last residual term. We also test whether the exceedances of the quantiles are independent of the conditional quantile estimator. The results suggest that both the location scale GARCH and CAViaR models suffer from exceedances being correlated with the forecast of the quantile itself, especially in the lower quantiles (1%, 5% and 10%). The results from the linear quantile regression models are more promising, particularly for the models including fundamental information.

Turning to the width of the predictions intervals, we find that the linear fundamental quantile regression models generally give the narrowest prediction intervals. That is, on average they provide relatively high estimates of the low quantiles (1%, 5% and 10%) and relatively low estimates of the high quantiles (90%, 95% and 99%). In particular, if we look at the three cases where there are no significant exceedances under both the UC and CC tests, ie model 4 (rolling), model 7 (expanding) and model 11 (rolling), the average widths for the 98% prediction intervals are 1.143, 1.088 and 1.240 respectively, and for the 90% intervals they are 0.739, 0.700 and 0.737 respectively, suggesting that of the three best calibrated models, model 7 (expanding) had the most precise prediction intervals on average. Since the precision and value of implementing the VaR method is linked to width of the intervals, our results suggest that the linear fundamental quantile regression models generally outperform the other models when it comes to the value of using the methods in practice.

The motivation for comparing an expanding and rolling estimation is to provide some evidence whether the data generating process changes over time. The electricity price in 2010 may have different dynamics compared to the dynamics in 2005. Using both an expanding and rolling window gives some indication if the methods we are testing are sensitive to the sample size chosen. The quantile forecasts from the parametric location-scale models show generally a slight improvement using a RW compared to using an EW. For the linear quantile regression the situation appears to be the opposite, with a tendency to more clustering of exceedances using a RW. All CAViaR models except the Adaptive model seem to perform slightly better using an EW. The reason may be due to data scarcity in estimating the extreme tails of the distribution. Chernozhukov et al (2001) observe that data scarcity problems are amplified by the presence of covariates, and that point estimates provided by regression quantiles can be biased in the tails.

## 6. CONCLUSIONS

We have characterised the nonlinear effects of exogenous factors on peak hour wholesale electricity price formation as well as in forecasting the price distribution. Using a dynamic quantile regression model with fundamental factors and conditional volatility as explanatory variables, we capture effects such as mean reversion, spikes, time varying volatility, and at the same time, estimate the rather complex relationships of this price to fundamentals. For this type of data, disentangling the intrinsic volatility from the induced effects of volatile fundamental factors is a substantial challenge, and the proposed approach was both transparent and comparatively effective. We demonstrated how lagged prices, prices of gas, coal and carbon, forecasts of demand and reserve margin in addition to price volatility influence the peak price distribution in quite intuitive ways. In general we find positive coefficients for the underlying fuel commodities (gas, coal and carbon prices) with,

as expected, no distinctly sustained pattern over the quantiles, as the relative impacts of coal and gas change over time and carbon prices adjust to some extent to their difference. The coefficient of demand is positive, as expected, with increased effects at higher quantiles. The coefficient of reserve margin is negative, also as expected, with increased impacts on higher quantiles. Conditional volatility of the shocks in expectations is mainly found to have an effect on extreme low and high quantiles, which is plausible. An increase in volatility drives low prices lower and high prices higher, suggesting that high price uncertainty combined low/high electricity prices may lead prices to overshoot the fundamentals.

We have also shown that the quantile regression models, taking into account the nonlinear effects of exogenous factors, outperform the CAViaR and GARCH models, as specified on the shocks to conditional expectations, regarding out-of-sample forecast of the price distribution quantiles. Furthermore, including conditional volatility as a factor captures heteroscedasticity in transparent way, with plausible market-induced effects, without the need for a fully parametric specification. Thus, using an expanding window the multifactor-with-volatility quantile regression models are well-calibrated and provide the most accurate forecasts. They are not outperformed and they are also easier to implement than some of the other state-of-the-art benchmark methods. The benefit of being able to model the distinctly different impacts of fundamentals, as well as the conditional volatility of the shocks to the expectations, is evident not only in its explanatory power, and potential for scenario construction in risk simulation, but also in its out of sample forecasting performance.

In summary, we have developed a practical multifactor quantile approach for predicting the electricity price distribution. This can be used for accurate day-ahead nonparametric density estimation (via multiple quantiles) of the spot price distributions, as an aid for producers, retailers, and speculators in determining their optimal strategies for short-term operations, risk management, hedging and trading.

## REFERENCES

- Alexander C. (2008). *Market Risk Analysis – Volume I, II, III, and IV*, John Wiley & Sons, Ltd.
- Alexander, G. and A. Baptista (2004), “A Comparison of VaR and CVaR Constraints on Portfolio Selection with the Mean-Variance Method,” *Management Science* 50(9): 1261–1273. <http://dx.doi.org/10.1287/mnsc.1040.0201>.
- Aloui, C. (2008). “Value-at-risk analysis for energy commodities: long-range dependencies at fat tails in return innovations,” *Journal of Energy Markets* 1: 31–63.
- Bordignon, S., D.W. Bunn, F. Lisi and F. Nan (2013). “Combining day ahead forecast for British electricity prices,” *Energy Economics* 35: 88–103. <http://dx.doi.org/10.1016/j.eneco.2011.12.001>.
- Bunn D.W. and C. Fezzi (2008). “A vector error correction model for the interaction among gas, electricity and carbon prices,” in F. Gulli (ed) *Markets for carbon and power pricing in Europe*, Edward Elgar. <http://dx.doi.org/10.4337/9781848445031.00012>.
- Burger M., B. Graeber and G. Schindlmayr (2008). *Managing energy risk – An integrated view on power and other energy markets*, Wiley.
- Chang K. F., P. Gray, and B.V. Campen (2008). “A new approach to characterizing and forecasting electricity price volatility,” *International Journal of Forecasting* 24: 728–743. <http://dx.doi.org/10.1016/j.ijforecast.2008.08.002>.
- Chan K. F., and P. Gray (2006). “Using extreme value theory to measure value-at-risk for daily electricity spot prices,” *International Journal of Forecasting* 22: 283–300. <http://dx.doi.org/10.1016/j.ijforecast.2005.10.002>.
- Chen, D., and D.W. Bunn (2010). “Analysis of the nonlinear response of electricity prices to fundamental and strategic factors,” *IEEE Transactions on Power Systems* 25:1–12. <http://dx.doi.org/10.1109/TPWRS.2009.2035129>.
- Chen, C.W.S., S. Lin and P.I.H. Yu (2012). “Smooth transition quantile capital asset pricing models with heteroscedasticity,” *Computational Economics* 40: 19–48. <http://dx.doi.org/10.1007/s10614-011-9266-y>.
- Chernozhukov, V. and L. Umantsev (2001). “Conditional value-at-risk: Aspects of modeling and estimation,” *Empirical Economics* 26: 271–292. <http://dx.doi.org/10.1007/s001810000062>.



- Christoffersen, P., (1998). "Evaluating interval forecast," *International Economic Review* 39: 817–840. <http://dx.doi.org/10.2307/2527341>.
- Christoffersen, P., (2010). *Backtesting. Encyclopedia of Quantitative Finance*, John Wiley & Sons, Ltd. <http://dx.doi.org/10.1002/9780470061602.eqf15018>.
- Engle R.F., and S. Manganelli (2004). "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles," *Journal of Business & Economic Statistics* 22: 367–381. <http://dx.doi.org/10.1198/073500104000000370>.
- Füss R., Z. Adams and D.G. Kaiser (2010). "The predictive power of Value-at-Risk Models in commodity futures markets," *Journal of Asset Management* 11: 261–285. <http://dx.doi.org/10.1057/jam.2009.21>.
- Giot, P., S. Laurent (2003). "Market risk in commodity markets: a VaR approach," *Energy Economics* 25: 435–457. [http://dx.doi.org/10.1016/S0140-9883\(03\)00052-5](http://dx.doi.org/10.1016/S0140-9883(03)00052-5).
- Gerlach R. H., C.W.S. Chen and N.Y.C. Chan (2011). "Bayesian time-varying quantile forecasting for Value-at-Risk in financial markets," *Journal of Business & Economic Statistics* 29: 481–492. <http://dx.doi.org/10.1198/jbes.2010.08203>.
- Glasserman, P., P. Heidelberger and P. Shahabuddin (2000). "Variance reduction techniques for estimating Value-at-Risk," *Management Science* 46 (10): 1349–1364. <http://dx.doi.org/10.1287/mnsc.46.10.1349.12274>.
- Guermat, C. and R.D.F. Harris (2001). "Robust conditional variance estimation and value-at-risk," *Journal of Risk* 4: 25–42.
- Guermat, C., and R.D.F. Harris (2002). "Forecasting value at risk allowing for time variation in the variance and kurtosis of portfolio returns," *International Journal of Forecasting* 18: 409–419. [http://dx.doi.org/10.1016/S0169-2070\(01\)00122-4](http://dx.doi.org/10.1016/S0169-2070(01)00122-4).
- Hao L., and D.Q. Naiman (2007), *Quantile Regression*, Sage Publication.
- Gneiting T, F. Balabdaoui, A.E. Raftery (2007). "Probabilistic forecasts, calibration and sharpness." *J. R. Statist. Soc. B* 69: 243–268. <http://dx.doi.org/10.1111/j.1467-9868.2007.00587.x>.
- Hong, L.J., and G. Liu (2009), "Simulating Sensitivities of Conditional Value at Risk," *Management Science* 55(2): 281–293. <http://dx.doi.org/10.1287/mnsc.1080.0901>.
- Hung J., M. Lee and H. Liu (2008), "Estimation of value-at-risk for energy commodities via fat-tailed GARCH models," *Energy Economics* 30: 1173–1191. <http://dx.doi.org/10.1016/j.eneco.2007.11.004>.
- Jeon, J., J.W. Taylor (2013), "Using Implied Volatility with CAViaR Models for Value at Risk Estimation," *Journal of Forecasting* 32 (1): 62–74. <http://dx.doi.org/10.1002/for.1251>.
- Jonsson, T., P. Pinson, H. Madsen and H. Nielson (2014). Predictive Densities for Day-Ahead Electricity Prices Using Time-Adaptive Quantile Regression. Technical University of Denmark.
- Energies* 7\*(9), 5523–5547; doi:10.3390/en7095523 <http://dx.doi.org/10.3390/en7095523>.
- Karakatsani N., and D.W. Bunn (2008), "Forecasting electricity prices: The impact of fundamentals and time-varying coefficients," *International Journal of Forecasting* 24: 764–785. <http://dx.doi.org/10.1016/j.ijforecast.2008.09.008>.
- Koenker R. and G. Bassett Jr. (1978), "Regression Quantiles," *Econometrica*, 46: 33–50. <http://dx.doi.org/10.2307/1913643>.
- Koenker R., and J.A.F. Machado (1999), "Goodness of fit and related inference processes for quantile regression," *Journal of American Statistical Association* 94: 1296–1310. <http://dx.doi.org/10.1080/01621459.1999.10473882>.
- Koenker R., (2005), *Quantile Regression*, Cambridge University Press. <http://dx.doi.org/10.1017/CBO9780511754098>.
- Kosater, P. and K. Mosler (2006), "Can Markov regime-switching models improve power-price forecasts? Evidence from German daily power prices," *Applied Energy* 83: 943–958. <http://dx.doi.org/10.1016/j.apenergy.2005.10.007>.
- Koenker, R. and Q. Zhao (1996), "Conditional quantile estimation and inference for ARCH models," *Econometric Theory* 12: 793–813. <http://dx.doi.org/10.1017/S0266466600007167>.
- Kuester, K., S. Mittnik and M.S. Paolella (2006), "Value-at-Risk Prediction: A Comparison of Alternative Strategies," *Journal of Financial Econometrics* 4: 53–89. <http://dx.doi.org/10.1093/jjfinec/nbj002>.
- Kupiec P., (1995), "Techniques for verifying the accuracy of risk measurement models," *Journal of Derivatives* 2: 183–184.
- Misiorek, A., S. Trueck and R. Weron (2006), "Point and Interval Forecasting of Spot Electricity Prices: Linear vs. Non-Linear Time Series Models." *Studies in Nonlinear Dynamics & Econometrics* 10(3): Article 2.
- Nowotarski, J., E. Raviv, S. Trueck and R. Weron (2014). "An empirical comparison of alternate schemes for combining electricity spot price forecasts, HSC Research Reports HSC/13/07," *Energy Economics*, forthcoming. <http://dx.doi.org/10.1016/j.eneco.2014.07.014>.
- Rothkopf, M.H. (1999), "Daily Repetition: A Neglected Factor in the Analysis of Electricity Auctions," *The Electricity Journal* 11: 60–70. [http://dx.doi.org/10.1016/S1040-6190\(99\)00010-X](http://dx.doi.org/10.1016/S1040-6190(99)00010-X).
- Stoft, S. (2002). *Power System Economics: Designing Markets for Electricity*, John Wiley & Sons, Ltd. <http://dx.doi.org/10.1109/9780470545584>.

- Taylor, J.W., and D.W. Bunn (1998), "Combining Forecast Quantiles using Quantile Regression: Investigating the derived weights, estimator bias and imposing constraints," *Journal of Applied Statistics* 25: 193–206. <http://dx.doi.org/10.1080/02664769823188>.
- Taylor, J.W., and D.W. Bunn (1999), "A Quantile Regression Approach for Generating Prediction Intervals," *Management Science* 45(2): 225–237. <http://dx.doi.org/10.1287/mnsc.45.2.225>.
- Taylor, J., (2005), "Generating Volatility Forecasts from Value at Risk Estimates," *Management Science* 51 (5): 712–725. <http://dx.doi.org/10.1287/mnsc.1040.0355>.
- Taylor J., (2008), "Using exponential weighted quantile regression to estimate value at risk and expected shortfall," *Journal of Financial Econometrics* 6: 382–406. <http://dx.doi.org/10.1093/jjfinec/nbn007>.
- Wolak, F., and R. Patrick (1997), "The Impact of Market Rules and Market Structure on the Price Determination Process in the England and Wales Electricity Market," *Working Paper*, University of California Energy Institute.
- Wolfram, C. (1999), "Measuring Duopoly Power in the British Electricity Spot Market," *American Economic Review* 79: 805–826. <http://dx.doi.org/10.1257/aer.89.4.805>.
- Xiao, Z. and R. Koenker (2009), "Conditional quantile estimation for generalized autoregressive conditional heteroscedasticity models," *Journal of the American Statistical Association* 104: 1696–1712. <http://dx.doi.org/10.1198/jasa.2009.tm09170>.