NCER

National Centre for Econometric Research

NCER Working Paper Series

Modelling Spikes in Electricity Prices

R. Becker, S. Hurn and V. Pavlov

Working Paper #10

February 2007

Modelling spikes in electricity prices *

Ralf Becker

Economics, School of Social Sciences, University of Manchester

Stan Hurn and Vlad Pavlov

School of Economics and Finance, Queensland University of Technology

Abstract

During periods of market stress, electricity prices can rise dramatically. Electricity retailers cannot pass these extreme prices on to customers because of retail price regulation. Improved prediction of these price spikes, therefore, is important for risk management. This paper builds a time-varying-probability Markov-switching model of Queensland electricity prices, aimed particularly at forecasting price spikes. Variables capturing demand and weather patterns are used to drive the transition probabilites. Unlike traditional Markov-switching models, that assume normality of the prices in each state, the model presented here uses a generalized beta distribution to allow for the skewness in the distribution of electricity prices during high-price episodes.

Keywords

electricity prices, regime switching, time-varying probabilities, beta distribution

JEL Classification C22, C53, Q49

Corresponding author

Vlad Pavlov

School of Economics and Finance

Queensland University of Technology

Brisbane, 4001, Australia

email `v.pavlov@qut.edu.au`

Telephone number: 07-3138 5293

Fax number: 07-3138 4150

I. Introduction

Time-series data on electricity prices typically display two prominent features, both of which stem from the lack of practical ways to store electricity. *First*, predictable demand fluctuations over daily, weekly and yearly frequencies create a clear seasonal pattern. *Second*, during periods of market stress electricity prices can rise dramatically. These stressed market situations are associated with either unusually high demand or, more often, unexpected shortfalls in supply that can be caused, for example, by generator failures. Price surges or "spikes"are of particular importance to electricity retailers who, because of retail price regulation, cannot pass them onto final customers and end up bearing the price risk. While state governments may enter into compensation deals with retailers, acknowledging that this risk allocation is partly induced by the market design, it is of great interest to retailers to understand why and when these price surges occur and, once they have occurred, how long they will last.

This paper investigates the usefulness of a regime-switching model for modelling and forecasting price spikes. The main finding of the paper is that the ability to predict high-price episodes is enhanced by using a Markov-switching model with time-varying transition probabilities that are determined by demand patterns and meteorological variables. An interesting feature of the econometric model is the use of the generalized beta distribution to capture the skewness of electricity prices in the high price state.

The structure of the paper is as follows. Section 2 describes the main features of the price data and some stylized facts about the price process. A regime-switching model with time-varying-probabilities (TVP) of transition is introduced in Section 3. Section 4 discusses the nature of factors driving the time-varying transition probabilities. In Sections 5 and 6 the baseline and extended models of electricity prices are presented and their forecasting ability assessed. Section 7 is a brief conclusion.

II. The data

The national wholesale electricity market (NEM) has been operating in Australia since 1998. It comprises six state grids, five of which are physically linked by interconnectors which transfer electrical energy between states. Electricity prices change every five minutes, reflecting orders that come into the system and supply schedules submitted by generators, but trades are settled on half-hourly averages of five minute prices, the so called pool price. This paper investigates the features of average daily Queensland half-hourly or pool prices for the period from December 8, 1998 to June 30, 2005. The daily price is computed as a simple arithmetic average of pool prices between 6 *am* and 9 *pm* and will be referred to in the paper simply as the price.[1] The use of daily-average data is motivated by the objective of constructing a simple model relating the occurrence and duration of high price/high volatility to a small number of exogenous factors. Self-evidently, a clear understanding of the intra-day propagation of demand and supply shocks would require a detailed modelling of the market structure, a task at odds with the desire to keep the model as simple as possible. The use of average daily prices has two major benefits: namely, the modelling of the seasonal component of the series is considerably simplified; and the influence of important supply and demand fluctuations is more easily captured in terms of the available data on meteorological conditions.

The approach to modelling the price level (rather than price differences) presupposes the stationarity of the price process. This assumption is justifiable on the grounds that for the short timeframe over which the model is estimated and used for forecasting, the effect of neglected non-stationarity will be dominated by the seasonal pattern and day-to-day volatility of price changes.

It is instructive to consider the main empirical features of intra-day prices that give rise to daily averages. Figure 1 plots realizations of half-hourly prices over two typical weekly periods. Both panels of the figure show clearly the intra-day and weekly seasonal patterns – pronounced early morning and late afternoon demand-driven peaks – with the exception of Sundays, where the morning peak is absent. During the period between 9 *pm* and 6 *am*, following the tapering off of electricity demand overnight, half-hourly prices drop to about \$15 per Mega Watt hour (MWh) and show very little volatility around the seasonal pattern. Since prices during these off-peak times contain very little extra information they are excluded when constructing time-series of average daily prices used later in the econometric analysis[2]. The top panel of Figure 1 also includes a small singular price spike with prices jumping from around \$35 to \$81.17 per MWh at 6:30 pm on Tuesday. The bottom panel shows a more interesting price episode with prices jumping sharply on Thursday reaching a maximum of over \$150 per MWh, falling back to normal levels overnight, but quickly rising above $60 per MWh again on Friday. It is the prediction of these price spikes that is the primary interest of this paper.

[Figure 1 about here.]

Daily prices over the sample period are plotted in Figure 2. The sample starts on December 8, 1998 and ends on June 30, 2005. Notice that the full sample period has been divided into two sub-samples, namely the estimation sample (from December 8, 1998 to January 29, 2003) and the hold-out sample (from January 30, 2003 to June 30, 2005). The estimation sample is first used to select an appropriate model for electricity price forecasting and to conduct an initial forecast evaluation. Recognizing, however, that use of the same sample for model selection, estimation and forecasting can potentially create an overly optimistic assessment of the model, an extended sample of data, the hold-out sample, is retained and used to assess the robustness of the model forecasts in an out-of-sample environment.

[Figure 2 about here.]

It is also apparent from Figure 2 that price spikes dominate the volatility of the electricity price series and their characteristics are quite striking. While the average price over the period was about $55 per MWh, on a number of occasions prices rose considerably above this level (to the maximum of $2132 per MWh). An important feature, which provides the initial motivation for this paper, is the apparent clustering of periods of high price volatility. Another important observation concerns the change in the behavior of the series following the introduction of the interconnector between the Queensland and the NSW grids on February 18, 2001. It is reasonable to expect, *a priori*, that the additional supply available through the interconnector would moderate idiosyncratic shocks in both states with the extent of this moderating effect being determined by the capacity of the interconnector. Consistent with this argument, the price volatility during periods with no extreme price movements dropped markedly after the interconnector was brought into operation. However, the interconnector appears to have done very little to reduce the amplitude of extreme shocks.

III. The regime switching model

The basic premise upon which the model is based is that the wholesale electricity market displays two distinctly different states. The first state, which may be called the *benign state*, is one in which extreme events are unlikely to occur, while the second state, referred to as the *stressed state*, is characterized by the presence of either extreme supply or demand shocks, or both, likely resulting in a significantly higher average daily price. Importantly, a stressed state is not immediately identified by the occurrence of extreme prices, rather it should be thought of as a state associated with a greater price risk than a benign state. A two-state regime-switching model (Hamilton, 1990) is the natural starting point for this modelling exercise.[3]

A traditional Markov regime switching model can be written as

$$y_t \sim \begin{cases} f_1\left(x_t; \theta_1\right) & \text{if } s_t = 1 \\ \\ f_2\left(x_t; \theta_2\right) & \text{if } s_t = 2 \end{cases} \tag{1}$$

$$\mathbf{P} = \begin{pmatrix} p_{11} & p_{21} \\ \\ p_{12} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1 - p_{22} \\ \\ 1 - p_{11} & p_{22} \end{pmatrix}. \tag{2}$$

Note that the distribution of the observable variable, y_t, is either $f_1\left(x_t; \theta_1\right)$ or $f_2\left(x_t; \theta_2\right)$ depending on the value taken by the unobserved state variable, s_t. Furthermore, the distributions $f_1\left(x_t; \theta_1\right)$ and $f_2\left(x_t; \theta_2\right)$ are typically assumed to be normal with means conditional on a set of exogenous and/or lagged dependent variables x_t and parameters $\theta_i, i = 1, 2$. Finally, the transitions between states are modelled as a Markov process with the transition matrix $\mathbf{P}$. The elements p_{ij}, $i, j = 1, 2$, indicate the probability of moving from state i to state j, $P\left(s_t = j | s_{t-1} = i\right)$. The higher the probabilities p_{11} and p_{22}, the higher is the probability of seeing prolonged periods in state 1 or 2 respectively.

In this application, the traditional Markov regime switching model is modified in two ways. Cursory examination of Figure 2 suggests a positively-skewed unconditional distribution for electricity prices. To allow for the possibility that this feature is replicated in the two states, the beta distribution, rather than the normal distribution, is used for $f_1\left(x_t; \theta_1\right)$ and $f_2\left(x_t; \theta_2\right)$. The beta distribution is defined on $[0, 1]$ and is described by two parameters, α, $\beta \in \mathcal{R}^+$, which determine the shape and location of the distribution. Several examples of the distribution, indicating its versatility, are displayed in Figure 3. The upper price boundary (regulatory cap) for electicity pool prices is $10,000 per MWh and scaling by this figure is thus a readily available way of fitting prices into the unit segment.

[Figure 3 about here.]

The second modification to the traditional Markov switching model concerns the transition probabilities. Self evidently the assumption of constant transition probabilities made by Hamilton (1990) is not appropriate for electricity prices for at least two reasons. *First*, there is a straightforward weekend effect on transition probabilities. Extreme price outcomes are relatively less frequent on weekends. For example, out of 121 instances of prices in excess of $100 per MWh only 19 fall on Saturday or Sunday. In contrast, 40 out of 121 are observed on Mondays. Any reasonable Markov model should at least allow for the probabilities of a price spike either starting on, or spilling over into, a weekend or a public holiday to be smaller than the corresponding probabilities for weekdays. *Second*, it is reasonable to expect that the duration of a spike should depend on the type and features of the environmental and economic conditions prevailing at the time. For example, a prolonged period of extremely high temperatures or loads may put generators under stress, thus exacerbating both the frequency and severity of failures. It may also be that generators change their economic behavior and bidding strategies in response to meteorogical conditions.

The regime-switching model based on the beta distribution and with time-varying transition probabilities may be formulated as follows:

$$y_t \sim \begin{cases} beta\left(x_t; \alpha_1, \beta_1\right) & \text{if } s_t = 1 \\ beta\left(x_t; \alpha_2, \beta_2\right) & \text{if } s_t = 2 \end{cases} \tag{3}$$

where s_t is the unobserved state variable and the transition probability matrix is given by

$$\mathbf{P} = \begin{pmatrix} p_{11}(\mathbf{z}_t, \gamma_1) & 1 - p_{22}(\mathbf{z}_t, \gamma_2) \\ 1 - p_{11}(\mathbf{z}_t, \gamma_1) & p_{22}(\mathbf{z}_t, \gamma_2) \end{pmatrix}. \tag{4}$$

The $k \times 1$ vector $\mathbf{z}_t$ contains explanatory variables (including a constant term) influencing the transition probabilities and γ_1 and γ_2 are parameter vectors. Following Filardo (1994) the relationship between transition probabilities and $\mathbf{z}_t$ is specified using a Logit transformation

$$p_{11}(\mathbf{z}_t, \gamma_1) = \frac{\exp(\mathbf{z}_t' \gamma_1)}{1 + \exp(\mathbf{z}_t' \gamma_1)} \tag{5}$$

$$p_{22}(\mathbf{z}_t, \gamma_2) = \frac{\exp(\mathbf{z}_t' \gamma_2)}{1 + \exp(\mathbf{z}_t' \gamma_2)} \tag{6}$$

Define $x_t = (y_{t-1}, ...; z_t, z_{t-1}, ...)$ as all observations on the predetermined and exogenous variables up to time t respectively[4]. The conditional likelihood for the regime-switching model can then be written as

$$l(\theta) = \sum_{t=1}^{T} \log f(y_t | x_t) = \sum_{t=1}^{T} \sum_{i=1}^{2} P(s_t = i | x_t)\ beta(y_t; \alpha_i, \beta_i), \tag{7}$$

where the beta distribution is defined by

$$beta(y; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} y^{\alpha - 1} (1 - y)^{\beta - 1} \tag{8}$$

$$B(\alpha, \beta) = \int_0^1 y^{\alpha - 1} (1 - y)^{\beta - 1} \, dy, \tag{9}$$

with the moments of the distribution being functions of the two parameters α and β (see, for example, Johnson and Kotz, 1970). The likelihood is computed by iterating between the optimal inference on the current state and the one-step ahead forecast using the following two equations

$$P(s_t = j|y_t, x_t) = \frac{P(s_t = j|s_{t-1}, x_t)\ beta\,(y_t; \alpha_j, \beta_j)}{\sum_{i=1}^{2} P(s_t = i|s_{t-1}, x_t)\ beta\,(y_t; \alpha_i, \beta_i)} \quad (10)$$

$$P(s_{t+1}|x_t) = \mathbf{P} \times P(s_t = j|y_t, x_t) \quad (11)$$

The forecast iterations are started by setting the filtered probabilities $P(s_t|y_{t-1}, x_{t-1})$ to the ergodic probabilities of the transition matrix evaluated at the mean value of the exogenous variables.

IV. Modelling transitional probabilities

As noted in the previous section, the regime switching model proposed for this application requires the selection of a $k \times 1$ vector $\mathbf{z}_t$ containing the explanatory variables that drive the transition probabilities. All the variables that were considered are tabulated in Table 1 and a few comments on these variables are in order.

[Table 1 about here.]

It is reasonable to expect that the probability of either moving or remaining in the stressed state drops during off-peak (weekends and public holidays) periods. This observation is incorporated into the model specification using a dummy for peak periods, $Peak_t$.

Visual inspection of the data (Figure 2) reveals that the introduction of the interconnector between Queensland and NSW electricity grids on February 18, 2001 had a marked effect on the behavior of prices. The most apparent change is the reduction of the price volatility in the benign state.

The interconnector also appears to have reduced the frequency of high price episodes. For example, in the pre interconnector sub-sample 10% of observations register prices in excess of \$100 per MWh (18% - over \$65 per MWh). Following the introduction of the interconnector the incidence of over \$100 per MWh prices dropped to 6% (10% - over \$65 per MWh). On the other hand, the interconnector did very little to reduce the amplitude of price peaks. To account for the effect of the interconnector a dummy variable, QNI_t, is used to allow the behaviour of the transition probabilities to differ before and after February 18, 2001. This means, in effect, that there are three regimes in the model, one stressed state and two benign states, before and after interconnection respectively. A fundamental simplifying assumption is made to the effect that the interconnector does not interact with other exogenous variables. This assumption is necessary in order to limit the model-specification search to manageable proportions.

All the other variables attempt to capture stress accumulation in the system by means of unexpected demand expansions and extreme weather conditions. Weather extremes are of course responsible for major fluctuations of demand around its trend values, but are also included separately because of their technological effects on the generation and transmission process (e.g. transmission losses depend on temperature).

The variable that most directly represents demand is the load.[5] Load displays very regular intra-day, weekly and yearly seasonal patterns and electricity generators will take these regularities into account when bidding. The deterministic seasonal component of the load variable is modelled using a mixture of dummy variables and trigonometric functions. Intra-day data is used to construct two measures of demand variation, namely, average daily load and maximum daily load.

The meteorological data contains half hourly observations on temperature and dewpoint for Archerfield, a Brisbane weather station.[6] The difference between temperature and dewpoint indicates the degree of humidity with smaller differences corresponding to higher humidity. We use the raw data to compute the daily average and maximum temperature and the daily average and maximum difference between temperature and dewpoint. For inclusion in $\mathbf{z}_t$ we consider the values of the temperature and humidity series, as well as their deviations from seasonal cycles fitted by means of trigonometric functions.

[Figure 4 about here.]

Figure 4 illustrates a trigonometric function fitted to maximum daily temperatures for a representative period in the data set. The annual pattern in the maximum temperature is clearly visible.[7] It should be noted that both very high and very low temperatures have considerable effect on demand. To capture this effect, the specification of transition probabilities includes squared deviations of maximum and average daily temperatures from their seasonal trends. Temperature and humidity can also have an indirect effect on the supply side of the electricity market as a prolonged run of extreme temperatures or high humidity can raise the probability of a generator failure. It is possible to capture the effect of accumulated stress caused by a prolonged periods of extreme weather by including weekly moving averages of temperature variables.

An important point concerns the timing of both the load and the meteorological data. All these variables can be expected to have a contemporaneous effect on the probability of demand or supply surges and therefore in forecasting price extremes, one-day ahead forecasts of these variables are required. Rather than trying to formulate forecasts, in this simple first attempt to predict price spikes, two simple rules are used. For the temperature variables, actual values of the variables at time t will be assumed known at time $t-1$. While this is clearly unrealistic it is a better assumption, for example, than using lagged temperature as a forecast for time t. Temperature forecasts from meteorology agencies relying on a wide array of data and sophisticated physical modelling are readily available to all market participants and are likely to be superior to any purely statistical forecasts. For the load variable, a very simple forecast model, namely the observation of seasonally adjusted load at time $t-1$, will be used as the forecast for the load at time t.

The last explanatory variables are the lagged price and its moving average. It is likely that any conditions at time $t-1$ determining the price may still prevail the next day. The only clear seasonal pattern in the actual average prices is between peak and off-peak days. Hence the variable proposed here is the deviation of price from the peak or off-peak average. The inclusion of the lagged dependent variable will capture the impact of persistent but neglected factors.

A final word needs to be said on the modelling strategy. Ideally a formal general-to-specific strategy would be employed, but this is not feasible because of computational cost and also because of the high correlation between some of the potential variables used in the model. As a result a more crude trial and error method is employed.

V. *A baseline model*

The results from a simple baseline model that uses only deterministic variables to capture changes in the transition probabilities, $\mathbf{z}_{t-1} = (Peak_t, QNI_t)$, are reported in Table 2 and the resultant matrix of transition probabilities pre- and post-interconnector are in Table 3. Note that in in this and subsequent tables, regime 1 is identified as the stressed state; regimes 2a and 2b refer to benign states before and after the introduction of the interconnector respectively.

[Table 2 about here.]

[Table 3 about here.]

Perhaps the most notable feature of the results is the significant variation in the parameters α_i and β_i of the beta distributions for each of the three regimes. A simple test of the plausibility of these beta parameters is to check if the model produces produces reasonable estimates of the distributions in each state and of the stationary distribution of the price process. The price distributions in the three states estimated by the model are displayed in the top panel of Figure 5. The bottom panel of Figure 5 shows an estimate of the stationary distribution of prices obtained from a kernel estimator applied to the actual data, and compares this with a similar nonparametric kernel estimator applied to data generated from the model by simulation.[8] The model appears to capture the main features of the stationary distribution very well, particularly the marked positive skewness.

[Figure 5 about here.]

The baseline model accounts only for the lower probability of high prices in off-peak periods and the apparent effect of the introduction of the interconnector on price volatility in the benign state. The parameters on the $peak_t$ dummy have the predicted signs, so that the model picks up the basic pattern of a declining probability of high average prices on off-peak days. The same pattern should be true of probabilities pre- and post-interconnector, but it is interesting to note that the QNI_t dummy is not particularly significant in the transition probability equations of this baseline model. It is a common practice with Markov-switching models to classify observations into states on the basis of filtered probabilities. The correlation between the filtered and conditional transition probabilities indicates the degree to which this classification is predictable. Despite the rudimentary nature of the baseline model, the correlation between the conditional transition probability $P\left(s_t = 1|y_{t-1}, x_{t-1}\right)$ and filtered probability of being in state i, $P\left(s_t = i|y_t, x_t\right)$, is $\rho_{pp} = 0.707$. One might further correlate $P\left(s_t = 1|y_{t-1}, x_{t-1}\right)$ with the average pool price, y_t, itself, and this correlation is $\rho_{py} = 0.346$. This is a respectable result given that this is a minimalist model and the fact that the filtered probabilities will not exhibit the same extreme behaviour as the pool price.

It has been stressed that the objective of the time-series model in this paper is to identify conditions that make price spikes likely and then use this knowledge to predict these abnormal price events. In other words, the value of the model in practical situations depends on its ability to identify high risk periods and give prior warning of future price spikes. A direct test of the forecasting ability of the model, therefore, is to check how often a model prediction of an elevated likelihood of entering the stressed state is subsequently confirmed with the occurrence of a price spike. For the purposes of this exercise, price spikes will be defined primarily in terms of two levels of threshold prices, namely, \$65 and \$100. Since these price thresholds are essentially chosen only by inspection of the data, a density-based criterion is also used. According to this criterion, if $beta\left(y_t;\alpha_1,\beta_1\right) > beta\left(y_t;\alpha_2,\beta_2\right)$ the period t is identified as state 1, otherwise period t is identified as state 2. Relative to this criterion the absolute thresholds are more conservative in associating observations with the stressed state or regime 1, since inspection of the top panel of Figure 5 reveals that this cross-over point between the beta distributions is located just below \$50 per MWh. The performance of the model may then be summarized using the following prediction matrices:

$$A = \begin{pmatrix} a\left(high, stressed\right) & a\left(high, benign\right) \\ a\left(low, stressed\right) & a\left(low, benign\right) \end{pmatrix}. \tag{12}$$

The top left value $a\left(high, stressed\right)$ gives the number of days when the model prediction of a stressed state, $P\left(s_t = 1|y_{t-1}, x_{t-1}\right) > 0.5$, was followed by a price observation at time t deemed to be a price spike. The value $a\,(high, benign)$ refers to the number of instances when the price exceeded the threshold but $P\,(s_t = 1|y_{t-1}, x_{t-1}) < 0.5$. The definitions of $a\,(low, stressed)$ and $a\,(low, benign)$ are similar. The off-diagonal elements $a\,(low, stressed)$ and $a\,(high, benign)$ indicate the number of instances when the model failed to give a warning of an impending high price period and when it gave a false signal. Ideally the off-diagonal elements should be zero. As with the choice of price thresholds, the cut-off point of 0.5 in this classification is an essentially arbitrary choice. However it does lend itself to a reasonably natural interpretation, namely using 0.5 implies that an occurrence of the less likely transition is effectively defined as a failure of the forecast.[9]

The graph of average daily prices (Figure 2) clearly demonstrates the tendency of high price days to cluster. Not surprisingly the estimated probability of staying in the stressed state is high in both peak and off-peak periods (Table 3). It should therefore be relatively easy to predict high price outcomes once the process has entered the stressed state. A bigger challenge is to predict transitions from the benign to the stressed state. Consequently, separate matrices of prediction frequencies will be reported for periods in which the system is deemed to be in the stressed and benign state respectively. In order to provide some evidence of the efficacy of the model an alternative against which its performance can be measured must be provided. Here the predication classification of the baseline model is compared with predictions from a “naive” model. This naive model predicts that the process will stay in the benign state with probability one;

that the process will remain in the stressed state with probability one on peak-days and revert to the benign state on off-peak days. Table 4 displays the results of this comparative exercise.

[Table 4 about here.]

The two panels in the Table 4 report prediction matrices separately for transitions from periods classified as stressed and benign according to the filtered probabilities, as outlined previously. Each panel contains 6 prediction matrices. The first two rows (Baseline) in the table report forecasts from the baseline model while the bottom panel gives the outcomes of predictions based on the naive model. To aide interpretation of Table 4, a detailed interpretation of one prediction matrix for the baseline model using the $65 absolute cut-off as the ex-post classification criterion (figures are in bold), is provided. The number in the top left corner of the matrix indicates that on 102 occasions the model predicted that the market would remain in the stressed state and the subsequent price observation exceeded $65, resulting in a correct prediction. In 106 cases (p1/r2) the model predicted that the system would remain in the stressed state but the actual price realization did not exceed the threshold. On five days (p2/r1) the model predicted a transition from the stressed to the benign state but the next days price exceeded the threshold and on 23 days (p2/r2) this prediction turned out to be correct.

Comparing the performance of the models, it is clear that the baseline model offers little, if any, improvements over the naive model in either the stressed or benign states. In particular the model is unable to predict any switches from the benign to the stressed state. This is hardly surprising as the variables used to model transition probabilities in the baseline are those implicitly entering the naive rule.

VI. An extended model

As noted earlier, a comprehensive general-to-specific modelling strategy, involving all of the possible specifications for the time-varying transition probabilities, is not feasible due to computational considerations. As a consequence, a number of different models for the transition probabilities were estimated on a trial and error basis using data from December 1998 to January 2003. This section reports only the most successful model and examines its forecasting performance against a simple benchmark. In addition, data for the period from January 2003 to June 2005 is used for a further forecasting exercise designed to check the robustness of the model.

(i) In sample model selection and forecasting

The TVP specification includes the following exogenous variables $Peak_t$, QNI_t, NSP_{t-1}, U_{t-1}, Tma_t, $T^{+}trigsq_t$, Td_t^{-}, $Ltrigma_{t-1}$ and Pex_{t-1}. It is important to re-iterate that the specification uses lagged values of the load and contemporaneous values of meteorological variables, reflecting the assumption that accurate external weather forecasts are generally available. The parameter estimates are in Table 5.

[Table 5 about here.]

Somewhat surprisingly, the coefficient on the interconnector dummy variable is positive in both equations indicating that both the benign and stressed states are more persistent after interconnector. For the benign state, one possible explanation for this is that the interconnector has had a moderating influence on small shocks. More difficult to explain is the fact that the model also predicts longer lasting periods of market stress after the introduction of the interconnector. This may be due to the unusually prolonged period of market turbulence in May-June 2002. Another possible explanation may be related to correlation between demand and supply shocks in the NSW and Queensland grids.

As expected, the coefficient on the dummy variable $\gamma_{1,Peak}$ is positive and significant, indicating that the probability of staying in the stressed state rises on peak days and drops for off-peak days. The coefficient on the peak dummy $\gamma_{2,Peak}$ in the equation for p_{22} is negative, indicating that the probability of exiting the benign state drops on off-peak days. Large price spikes clearly depress the probability of staying in the benign regime, reflecting the observed clustering of high-price days. Another useful indicator of increasing persistence of stress is the number of pool prices over a trading day exceeding the \$65 threshold. The coefficient on NSP_{t-1} is positive suggesting that the number of high price episodes is linked to higher probabilities of staying in a stressed state. For example, two instances of pool prices of \$100 on a trading day provide a better indication of a persistent period of market stress than one \$200 price spike, which may only be indicative of a stressed period with a very short (less than 24 hours) duration.

The coefficient on NSP_{t-1} in the equation for p_{22} is insignificant, which is consistent with this observation. The interpretation of the effect of U_{t-1}, the duration of the period with no prices over \$100 per MWh is similar.

Turning to the meteorological variables, higher values of the 7 day moving average of the temperature seasonal ($\gamma_{1,Tma}$) are associated with lower probabilities of staying in both states, so both benign and stressed states would be expected to last longer in winter. The effect of the temperature deviations from seasonal patterns on transition probabilities is negative in both states. Importantly, temperature shocks considerably reduce the probability of staying in the benign state and the effect is statistically significant. The negative effect on the persistence of stressed conditions is more difficult to interpret. This effect is large but statistically insignificant. A possible explanation is that temperature variations have higher frequency than other technological shocks to the market, so the duration of spikes caused purely by unexpected extreme temperatures is relatively short. The coefficient on the $T^{+}trigsq_t$ variable is negative, indicating that deviations from the seasonal pattern of maximum daily temperatures, in either direction, tend to increase the probability of moving into the stressed state.

The log-likelihood for the preferred model is -4853.38 while for the baseline model in reported in the previous section the log-likelihood is -4779.37. This implies the likelihood ratio for the extended against the baseline model of 148.02 and the hypothesis that all of the extra variables included in the preferred model are statistically insignificant may be rejected. The correlations ρ_{pp} and ρ_{py} also increase to $\rho_{pp} = 0.827$ and $\rho_{py} = 0.468$ respectively.

More important, however, is the improved ability of the model to predict changes in regime. Table 6 reports that same prediction exercise as carried out in the previous section and reported in Table 4. Note that in interpreting these results it should be recognized that the extended model is now used to classify the observations into stressed and benign states. As this classification into different regimes is made on the basis of filtered probabilities, and is therefore model dependent, the outcomes for the naive model will be different to those reported in Table 4 even though the rules used to generate naive predictions are exactly the same as stated previously.

[Table 6 about here.]

Comparing the predicted switches from a stressed state, the extended model cuts the instances of false predictions of switching to a benign regime from 117 to 96 when using $65 as a threshold and from 149 to 130 when $100 is used. The false prediction of moving from stressed to a benign state has also improved from 18 for the naive model to 6 using the $65 threshold and from 12 to 2 for the $100 threshold. The extended model, therefore, shows considerable promise in forecasting switches from a benign to a stressed state or, given the high correlation between prices and filtered state probabilities, from a period of low prices to a period of high prices. The model would have predicted 18 out of 42 episodes of greater than $65 average daily prices and 11 of 42 periods with the average price exceeding $100.

To put these numbers in perspective, it may be noted that, by the very nature of the TVP regime switching model, both state classification and prediction are imprecise. TVP can only provide signals that high price outcomes are relatively more likely than low price outcomes. Therefore, by construction the model would not be able to predict all of the switches. Also, a proportion of price spikes can be expected to be associated with fundamentally unpredictable technological events. Taking these considerations into account, it could be argued with some justification that this approach to predicting price spikes shows a great deal of potential.

(ii) Out of sample results

Relying merely on in-sample forecast results leaves the preferred model vulnerable to criticism based on data-mining bias. As a consequence an extensive out-of-sample prediction exercise has also been undertaken. Data for the period from January 2003 to June 2005 were not used in the model-selection phase of the modelling exercise. Starting from January 2003, the model was iteratively re-estimated on a rolling window of data comprising 1514 observations, the length of the original estimation sample. The parameter estimates obtained for each window of data were then used to generate weekly out-of-sample forecasts. The window was then rolled one week forward and the exercise repeated. It is important to note that at each iteration the structure of the model remains constant and no additional model selection is undertaken. This process provides a check that the actual preferred model selected is not merely an artefact of the particular sample used for the original estimation. It should also be noted that, by conventional standards, the hold-out sample is a large one and almost as large as the original sample itself. This out-of-sample forecast evaluation, summarized in Table

7, therefore provides a stringent test of the preferred model.

[Table 7 about here.]

There are a number of important conclusions to be drawn from the results reported in Table 7. In the first instance, the improvement in the ability to predict switches between the regimes of the preferred specification over the naive model is of a similar order to that observed in-sample (reported in Table 6). For example, the incorrect predictions of the naive model, for switches from the stressed to the benign state, are reduced from 35 to 26 using the density based criterion, from 42 to 33 using the $65 threshold and from 28 to 21 using $100 cut-off. Quantitatively this is very similar to the improvement recorded in-sample. The model does not demonstrably improve the false predictions of switches into the benign regime but, given that the naive model generates so few false predictions, this result is not surprising.

Furthermore, as noted in the previous sub-section, an important aspect of the model is its ability to explain a significant percentage of switches from the benign regime into the stressed regime. This ability is also found in the out-of-sample results and is of the same order of magnitude as those reported in-sample. This is despite the fact that no additional model selection has been attempted on the extended data. For example, the full model predicts 17% of abnormal price events according to the density criterion (compared to 20% in the estimation sample) and 30% of incidences of greater than $65 price events (compared to 24% for the original sample). The performance in term of predicting excedences of $100 is weaker than the in-sample results

(12% versus 32%). This last observation may be due to changes in the extreme behaviour of the series in the hold-out sample.

Overall, it may be argued that the results show that the model performance in the hold-out sample is consistent with that reported in sample. Furthermore, the sustained improved performance of the extended model over the naive alternative cannot solely be due to data mining given the size of the hold-out sample. The conclusion that this approach to predicting price spikes in electricity markets is a useful one is therefore reinforced.

VII. Conclusion

This paper has demonstrated that a time-varying-probability regime-switching model can help to predict spikes in Queensland electricity pool prices. The model clearly identifies two separate states; a low risk or benign regime, with a peaked state distribution concentrated in the low to medium price range, and a high risk or stressed regime with a very disperse distribution of the state variable. The model indicates that meteorological and demand factors are important in explaining the persistence of each regime. Importantly, it appears that a reasonable proportion of switches from the benign to the stressed regime are predictable from the information contained in load variables and a relatively small subset of meteorological variables. The performance of the preferred model in a substantial hold-out sample demonstrates the potential of this modelling approach for forecasting abnormal events in electricity prices.

References

Engle R.F. and Russell (1998) "Autoregressive conditional duration: A new model for irregularly spaced transaction data" *Econometrica,* **66**, 1127-1162.

Filardo A.J. (1994) "Business-cycle phases and their transitional dynamics" *Journal of Business & Economic Statistics*, **12**, 299-308.

Hamilton J.D. (1990) "Analysis of time series subject to changes in regime" *Journal of Econometrics*, **45**, 39-70.

Hamilton J.D. and Jorda O. (2002) "A Model of the Federal Funds Target," *Journal of Political Economy*, **110**,1135-1167.

Johnson N.L. and Kotz S. (1970) *Continuous Univariate Distributions - 2*, John Wiley & Sons: New York.

Notes

*The financial support of Ergon Energy in terms of a Collaborative Research Grant with Industry is gratefully acknowledged. The research was also partly funded by ARC Discovery Project DP0666266. We wish to thank Luke Evans of Ergon Energy, Harry Bloch and two anonymous referees for very useful comments on earlier drafts of this paper. The remaining errors are those of the authors.

[1]The data are available from the NEMMCO website - http://www.nemmco.com.au.

[2]The exclusion of the overnight data has no material effect on the main conclusions of the paper.

[3]The other possible approach that springs immediately to mind is to model price surges directly using a duration model, such as that proposed by Hamilton and Jorda (2002) following earlier work by Engle and Russell (1998). This, however, would mean that the entire model construction would be critically dependent on the definition of the price spike.

[4]Note that the random variable y_t and its observed value are not differentiated notationally.

[5]Since electricity cannot be stored, the amount of electricity drawn from the grid and the amount generated should balance at any particular given time. This figure is usually called the load and, in the absence of blackouts or other abnormal events, the load is equivalent to the demand for electricity.

[6]Brisbane temperature data was chosen because the Brisbane area load

is by far the largest in the state.

[7]There were short periods where meteorological data were unavailable. These observations were dropped during estimation and care was taken to ensure the proper treatment of dynamics around these missing data.

[8]For each period an observation is drawn from $B(\alpha_1, \beta_1)$ with probability $P(s_t = 1|h_t)$ or from $B(\alpha_2, \beta_2)$ with the complementary probability. This process is then repeated 10 times and the resulting observations are passed to the kernel estimator.

[9]Note that limited experimentation with other probability cut-off thresholds did not materially alter the results of the paper.

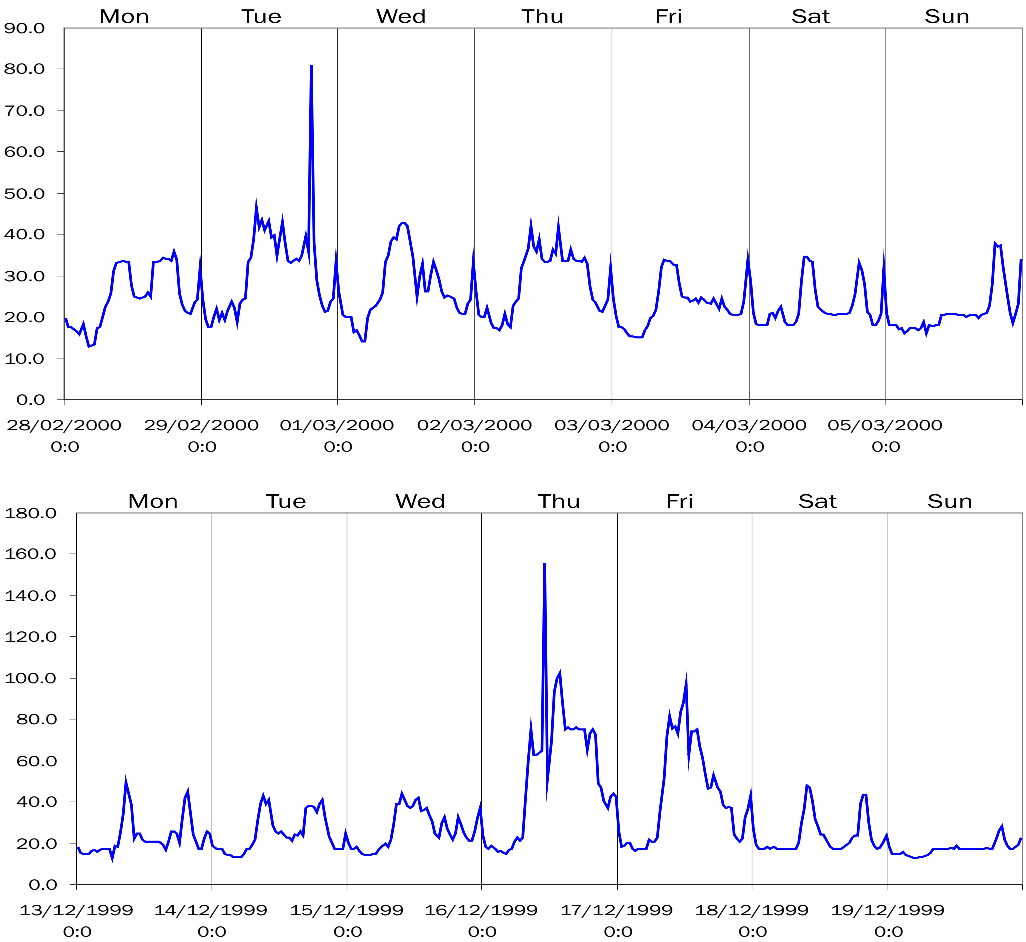

FIGURE 1

Queensland pool prices

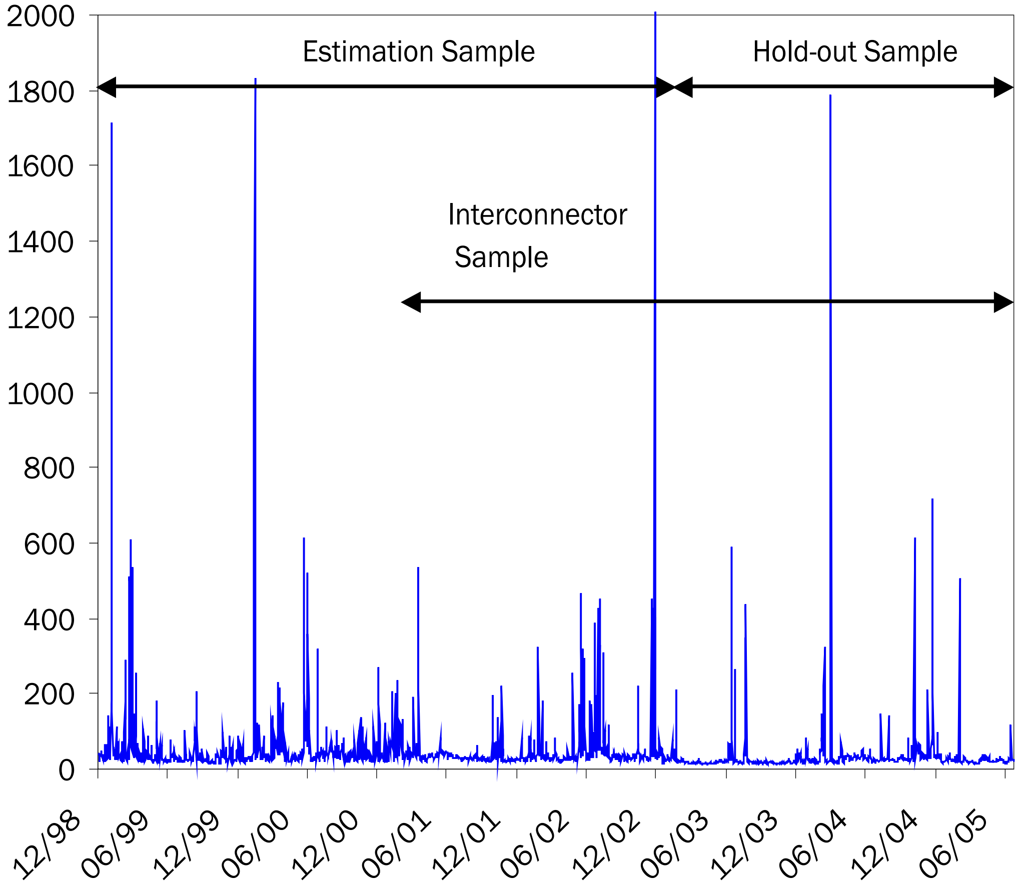

FIGURE 2
Average daily prices from 8-12-1998 to 30-06-2005

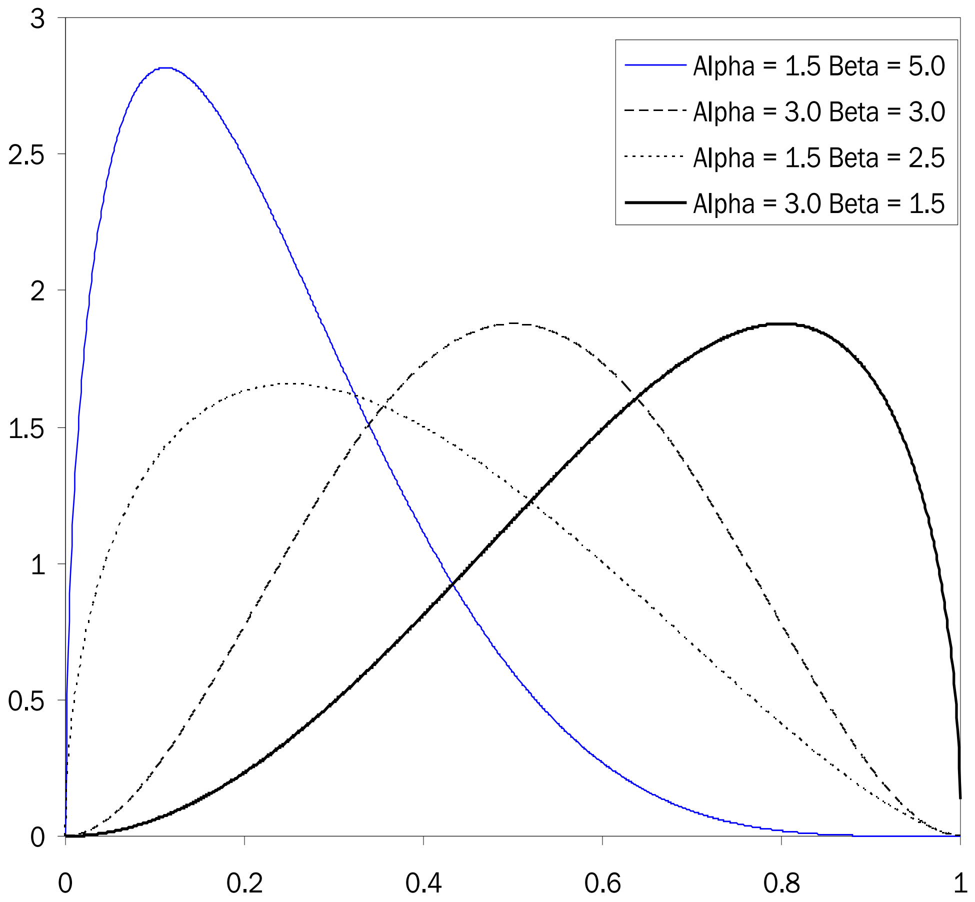

FIGURE 3
Symmetric and positively skewed beta distributions

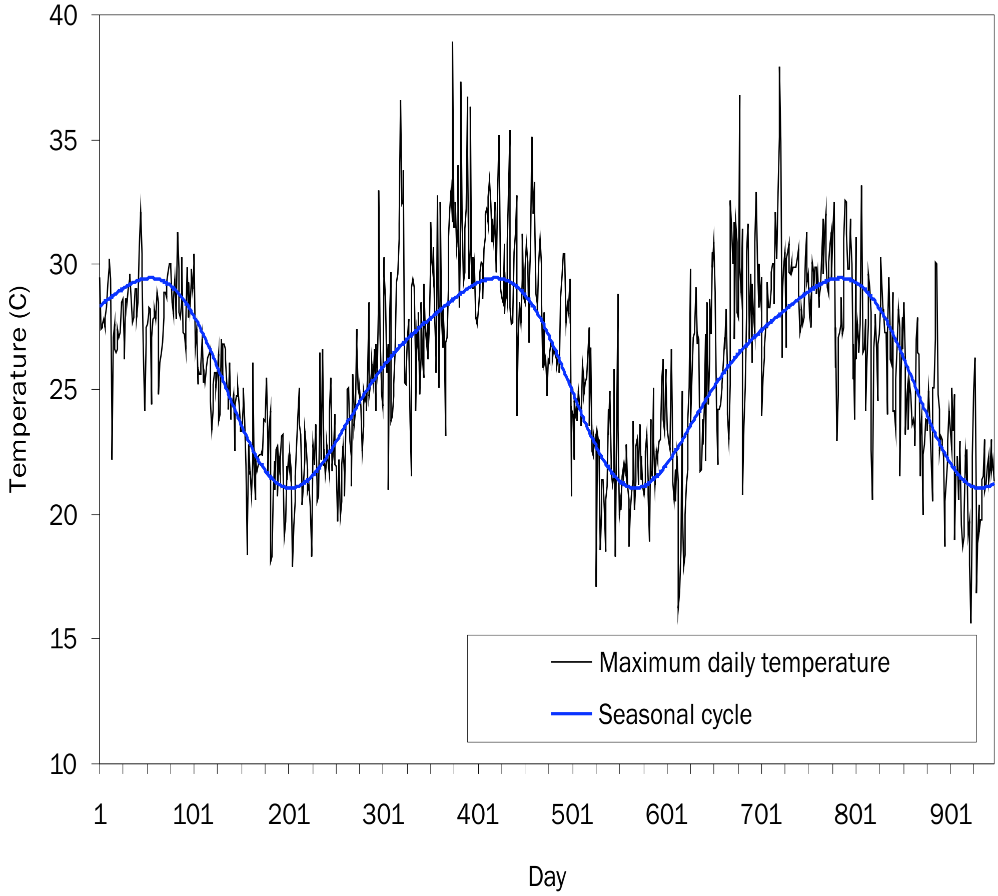

FIGURE 4
Deterministic approximation to the annual cycle in daily maximum temperatures

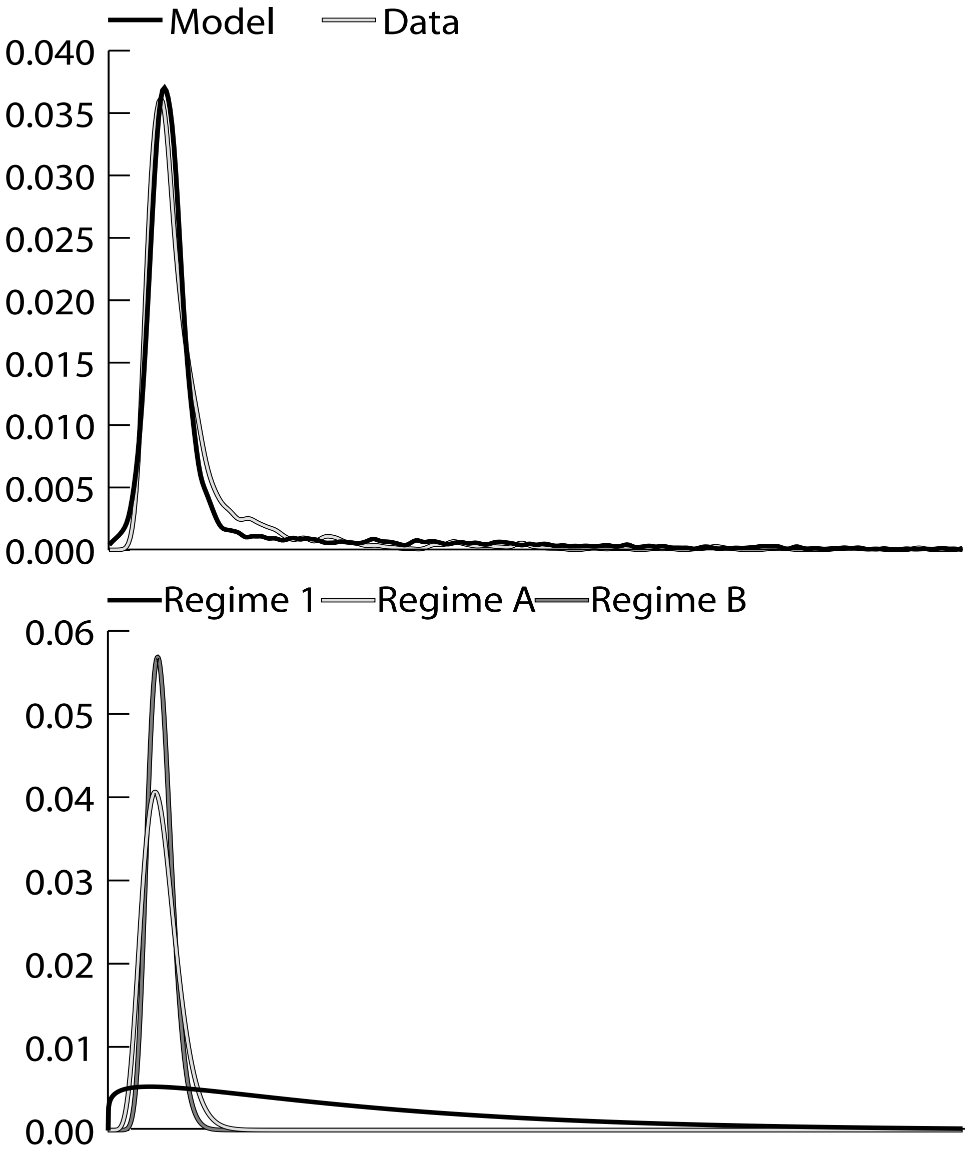

Figure 5

Densities for estimated Regime 1 and 2 (top panel). Empirical and baseline model densities for daily average prices

Name	**Description**
$Peak_t$	dummy; 1 on peak days (weekdays) -, 0 on off-peak days (weekends)
QNI_t	dummy for interconnector, 1 after 18/2/01
NSP_{t-1}	number of 30min prices > \$65/MWh
U_{t-1}	duration of period with no prices > \$100/MWh
Tma_t	7-day moving average of daily temperature
Tsq_t	squared deviation of average daily temperature from Tma_t
$Ttrig_t$	deviations of daily average temperature from its seasonal pattern
$Ttrigsq_t$	$Ttrig_t$ squared
T^+sq_t	squared deviation of maximum daily temperature from Tma_t
T^+trig_t	deviations of daily maximum temperature from the seasonal pattern
$T^+trigsq_t$	T^+trig_t squared
Td_t	average difference between temperature and dewpoint
Td_t^-	minimum difference between temperature and dewpoint
$Ltrig_{t-1}$	deseasonalised average daily load
L^+trig_{t-1}	deseasonalised maximum daily load
$Ltrigma_{t-1}$	7-day moving average of $Ltrig_{t-1}$
$L^+trigma_{t-1}$	7-day moving average of L^+trig_{t-1}
Pex_{t-1}	price deviation from peak or off-peak average
$Pexma_{t-1}$	7-day moving average of Pex_{t-1}

TABLE 1

Complete listing of all the variables considered in building a model for the time-varying transition probabilities.

	est	s.e.	t-stat
α_1	1.178	0.129	9.190
β_1	17.413	0.580	30.044
α_{2a}	7.442	0.658	11.309
β_{2a}	565.32	48.746	11.597
α_{2b}	17.755	1.325	13.404
β_{2b}	1423.5	103.23	13.789
$\gamma_{1,0}$	0.977	0.207	4.730
$\gamma_{1,Peak_t}$	0.686	0.204	3.362
γ_{1,QNI_t}	0.100	0.192	0.520
$\gamma_{2,0}$	2.792	0.174	16.054
$\gamma_{2,Peak_t}$	-0.592	0.246	-2.407
γ_{2,QNI_t}	0.175	0.144	1.214

TABLE 2
Parameter estimates for baseline model.

	Before QNI		After QNI	
Regime	Off-Peak	Peak	Off-Peak	Peak
1	0.727	0.841	0.746	0.854
2	0.942	0.9	0.951	0.915

TABLE 3
Probabilities of Staying in Regime.

	Stressed						**Benign**					
	Beta		$65		$100		Beta		$65		$100	
	p1	p2	**p1**	**p2**	p1	p2	p1	p2	p1	p2	p1	p2
Baseline												
r1	131	5	**102**	**5**	74	2	0	117	0	86	0	38
r2	77	23	**106**	**23**	134	26	0	997	0	1028	0	1076
Naive												
r1	120	16	92	15	66	10	0	117	0	86	0	38
r2	66	34	94	35	120	40	0	997	0	1028	0	1076

TABLE 4

Forecast evaluation for baseline model.

	est	$t-stat$		est	$t-stat$
α_1	1.183	9.863	α_{2a}	8.897	10.999
β_1	19.432	34.024	β_{2a}	711.36	11.154
			α_{2b}	17.98	14.006
			β_{2b}	1449.4	14.512
$\gamma_{1,0}$	-4.998	-2.153	$\gamma_{2,0}$	0.8910	2.695
$\gamma_{1,Peak}$	0.920	2.937	$\gamma_{2,Peak}$	-1.409	-5.928
$\gamma_{1,QNI}$	1.774	3.322	$\gamma_{2,QNI}$	0.694	3.310
$\gamma_{1,NSP}$	1.784	3.030	$\gamma_{2,NSP}$	0.127	0.212
$\gamma_{1,U}$	-12.55	-1.851	$\gamma_{2,U}$	-0.283	-2.246
$\gamma_{1,Tma}$	-0.685	-2.664	$\gamma_{2,Tma}$	-0.792	-3.647
$\gamma_{1,T^+trigsq}$	-0.398	-1.199	$\gamma_{2,T^+trigsq}$	-0.299	-2.068
$\gamma_{1,Td}$	-0.315	-1.386	$\gamma_{2,Td}$	0.058	0.323
$\gamma_{1,Ltrigma}$	-0.385	-1.520	$\gamma_{2,Ltrigma}$	0.021	0.112
$\gamma_{1,Pex}$	0.048	0.080	$\gamma_{2,Pex}$	-10.949	-5.456

TABLE 5
Parameter estimates for the preferred model.

	Stressed						**Benign**					
	Beta		$65		$100		Beta		$65		$100	
	p1	p2	p1	p2	p1	p2	p1	p2	p1	p2	p1	p2
Extended												
r1	147	14	112	6	78	2	23	90	18	57	11	23
r2	61	59	96	67	130	71	19	937	24	970	31	1004
Naive												
r1	140	21	100	18	68	12	0	113	0	75	0	34
r2	77	43	117	46	149	52	0	956	0	994	0	1035

TABLE 6
Forecast evaluation for the preferred model.

	Stressed						Benign					
	Beta		\$65		\$100		Beta		\$65		\$100	
	p1		p1		p1		p1		p1		p1	
Model												
r1	15	3	8	2	20	5	3	15	3	7	6	43
r2	26	17	33	18	21	15	8	794	8	802	5	766
Naive												
r1	15	3	8	2	22	3	0	18	0	10	0	49
r2	35	8	42	9	28	8	0	802	0	810	0	771

TABLE 7

Evaluation of the out-of-sample forecasting performance of the preferred model.