Munich Personal RePEc Archive

Modeling electricity spot prices: Regime switching models with price-capped spike distributions

Janczura, Joanna and Weron, Rafal

14 June 2010

Online at https://mpra.ub.uni-muenchen.de/23296/
MPRA Paper No. 23296, posted 14 Jun 2010 21:29 UTC

Modeling electricity spot prices: Regime switching models with price-capped spike distributions

Joanna Janczura
Hugo Steinhaus Center,
Institute of Mathematics and Computer Science,
Wrocław University of Technology
50-370 Wrocław, Poland
Email: joanna.janczura@pwr.wroc.pl

Rafał Weron
Institute of Organization and Management,
Wrocław University of Technology
50-370 Wrocław, Poland
Email: rafal.weron@pwr.wroc.pl

***Abstract*—We calibrate Markov regime-switching (MRS) models to spot (log-)prices from two major power markets. We show that while the price-capped (or truncated) spike distributions do not give any advantage over the standard specification in case of moderately spiky markets (such as NEPOOL), they improve the fit and yield significantly different results in case of extremely spiky markets (such as the Australian NSW market).**

***Index Terms*—Electricity spot price, Markov regime-switching model, Price spike, Price cap, Truncated distribution.**

I. Introduction

The aim of this paper is to test whether the common practice of ignoring market-imposed price caps in electricity spot price models is a harmless approximation or a dangerous procedure leading to under- or overestimation of price spike severity. To this end, we use the popular in the energy economics literature Markov regime-switching (MRS) models and evaluate the fit of models with standard, as well as truncated (or price-capped) spike regime distributions. Motivated by recent findings [11] we focus on MRS models with heteroskedastic base regime dynamics and shifted spike regime distributions. The rationale for the former comes from the observation that price volatility generally increases with price level, since positive price shocks increase volatility more than negative shocks (this is the so-called 'inverse leverage effect' [15]). Shifted spike distributions, on the other hand, are required for the calibration procedure to separate spikes from the 'normal' price behavior.

The analysis of spike size distributions in typical MRS models shows that, in some cases the fitted distributions are so heavy-tailed that the variance does not exist [10]. Yet, market prices are generally capped, yielding finite moments. Although model generated prices should comply with the market specifications, in the studies performed so far this issue was not taken into account. Therefore in this paper we introduce truncated spike distributions, which ensure that observations do not exceed a specified level and, hence, are well suited for modeling capped power market prices.

The paper is structured as follows. In Section II we introduce MRS models for electricity log-prices. Next, in Section III we present the datasets and explain the deseasonalization procedure. In Section IV we compare calibration results for the analyzed models. Finally, in Section V we conclude.

II. Markov regime-switching models

The idea underlying the Markov regime-switching (MRS) scheme is to model the electricity price (or any other observed stochastic process) by separate phases or regimes with different dynamics. The switching mechanism between the states is Markovian and is assumed to be governed by a latent random variable. The processes driving the individual regimes do not have to be Markovian, but in energy economics applications are often assumed to be independent from each other.

In this study, we let the average daily spot electricity price follow a 2-regime MRS model, which displays either normal (base regime $R_t = 1$) or high (spike regime $R_t = 2$) prices each day. The transition matrix $\mathbf{P}$ contains the probabilities p_{ij} of switching from regime i at time t to regime j at time $t+1$, for $i, j = \{1, 2\}$:

$$\mathbf{P} = (p_{ij}) = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} = \begin{pmatrix} p_{11} & 1-p_{11} \\ 1-p_{22} & p_{22} \end{pmatrix}. \quad (1)$$

The current state R_t at time t depends on the past only through the most recent value R_{t-1} and the probability of being in state j at time $t+m$ starting from state i at time t is given by $P(R_{t+m} = j \mid R_t = i) = (\mathbf{P}')^m \cdot e_i$, where $\mathbf{P}'$ is the transpose of $\mathbf{P}$ and e_i is the ith column of the identity matrix.

To our best knowledge, the MRS models were first applied to electricity prices in [6]. A two state specification was proposed, in which in both regimes the log-prices were governed by autoregressive processes of order one, i.e. AR(1), with the same error term. Huisman and de Jong [9] proposed a model for deseasonalized log-prices with two independent regimes – a stable, mean-reverting AR(1) regime and a spike regime modeled by a normal distributed random variable whose mean and variance were higher than those of the base regime process. This simple yet versatile model was further extended by admitting lognormal, Pareto [2] and exponential [1] spike regime distributions, as well as, autoregressive Poisson driven spike regime dynamics [4] or shifted spike distributions and heteroskedastic CIR-type base regime dynamics [10].

Some of the more recent second generation models (as classified in [11]) used fundamental information (system constraints, weather variables) to better model regime-switching. Mount et al. [16] proposed a 2-regime model with two AR(1) regimes for log-prices and transition probabilities dependent on the reserve margin. They concluded that the estimated switching probability from the base (low) to the spike (high) regime predicts price spikes well if the reserve margin is measured accurately. In a complementary study Huisman [8] used temperature as a proxy and showed that the probability of spike occurrence increases when temperature deviates substantially from mean temperature levels. However, in general, temperature does not provide as much information as the reserve margin.

Finally, in a recent review paper Janczura and Weron [11] tested a range of MRS models and concluded that the best structure was that of an independent spike 3-regime model with time-varying transition probabilities, heteroscedastic diffusion-type base regime dynamics and shifted spike and price drop regime distributions. Not only did it allow for a seasonal spike intensity throughout the year and consecutive spikes or price drops, which is consistent with market observations, but also exhibited the 'inverse leverage effect' reported in the literature for spot electricity prices.

The above mentioned models have a common feature. Namely, they ignore the fact that in organized markets, like power exchanges or power pools, electricity prices are generally capped. To address this issue we introduce truncated spike distributions. Motivated by [11] we use shifted spike regime distributions which assign zero probability to prices below a certain quantile (here: the third quartile $m = F^{-1}(0.75)$, where F is the cumulative distribution function) of the dataset. We consider the lognormal (LogN) distribution:

$$\log(X_t - m) \sim \mathrm{N}(\alpha_2, \sigma_2^2), \quad X_t > m, \tag{2}$$

and the truncated lognormal (TLogN) distribution:

$$f(x) = \frac{C \exp\left(-\frac{(\log(x-m)-\alpha_2)^2}{2\sigma_2^2}\right)}{(x-m)\sigma_2\sqrt{2\pi}}, \quad x > m, \tag{3}$$

where $C = \Phi((\log(L) - \alpha_2)/\sigma_2)$ is a normalizing constant, Φ is the standard Gaussian cumulative distribution function (cdf) and L is the truncation level.

For the base regime dynamics we use a mean-reverting heteroskedastic process of the form:

$$dX_t = (\alpha_1 - \beta X_t)dt + \sigma_1 X_t^{\gamma} dW_t. \tag{4}$$

Note, that in this model the volatility is dependent on the current price level X_t, i.e for a positive γ the higher the price level the larger are the price changes. Consequently, compared to the commonly used AR(1) dynamics, in this model the less extreme price changes will be generally classified as 'normal' and not spiky.

Calibration of MRS models is not straightforward since the regime is only latent and hence not directly observable. Hamilton [7] introduced an application of the Expectation-Maximization (EM) algorithm [5] where the whole set of parameters θ is estimated by an iterative two-step procedure. The algorithm was later refined by Kim [14]. In the first step the conditional probabilities $P(R_t = j|X_1, ..., X_T; \theta)$ for the process being in regime j at time t, so-called 'smoothed inferences', are calculated based on starting values $\hat{\theta}^{(0)}$ for the parameter vector θ of the underlying stochastic processes. Then, in the second step, new and more exact maximum likelihood (ML) estimates $\hat{\theta}$ for all model parameters are calculated. Compared to standard ML estimation, where for a given probability density function f the log-likelihood function $\sum_{t=1}^{n} \log f(X_t, \theta)$ is maximized, here each component of this sum has to be weighted with the corresponding smoothed inference, since each observation X_t belongs to the jth regime exactly with probability $P(R_t = j|X_1, ..., X_T; \theta)$.

The parameters of the 'shifted lognormal' regime are obtained as the ML estimates of the standard lognormal distribution fitted to log-prices shifted by m and weighted by the smoothed inferences. In the truncated lognormal case the ML estimation requires numerical maximization of the likelihood function. Finally, the base regime parameters are estimated via ML with each price being weighted by the smoothed inferences. Following [12] we replace the latent values from the base regime with their expectations. In every iteration the EM algorithm generates new estimates $\hat{\theta}^{(n+1)}$ as well as new estimates for the smoothed inferences. Each iteration cycle increases the log-likelihood function and the limit of this sequence of estimates reaches a (local) maximum of the log-likelihood function.

III. Data preprocessing

In this paper we concentrate on the New South Wales power market (NSW; Australia) and the New England Power Pool (NEPOOL; U.S.). We use mean daily spot log-prices from the period January 1, 2006 – December 26, 2009. The sample consists of 1456 daily observations (208 weeks). The bid caps are equal to 10000 AUD in the NSW market and 1000 USD in the NEPOOL market.

It is well known that electricity prices exhibit seasonality on the annual, weekly and daily level [3], [18]. Hence, we follow the 'industry standard' and represent the spot price P_t by a sum of two independent parts: a predictable (seasonal) component f_t and a stochastic component X_t, i.e. $P_t = f_t + X_t$. Further, we let f_t be composed of a weekly periodic part s_t and a long-term seasonal trend T_t, which represents both the changing climate/consumption conditions throughout the year and the long-term non-periodic structural changes. As in [11] the deseasonalization is conducted in three steps.

First, T_t is estimated from daily spot prices P_t using a wavelet filtering-smoothing technique (for details see [17], [18]) with the 8th level (or S_8) approximation, which roughly corresponds to annual ($2^8 = 256$ days) smoothing. The price series without the long-term seasonal trend is obtained by subtracting the S_8 approximation from P_t. Next, the weekly periodicity s_t is removed by subtracting the average week calculated as the median of prices corresponding to each day of the week. The median is used instead of the commonly used mean as it is more robust to outliers (extreme prices); this is especially important for the extremely spiky Australian prices. Finally, the deseasonalized prices, i.e. $P_t - T_t - s_t$, are shifted

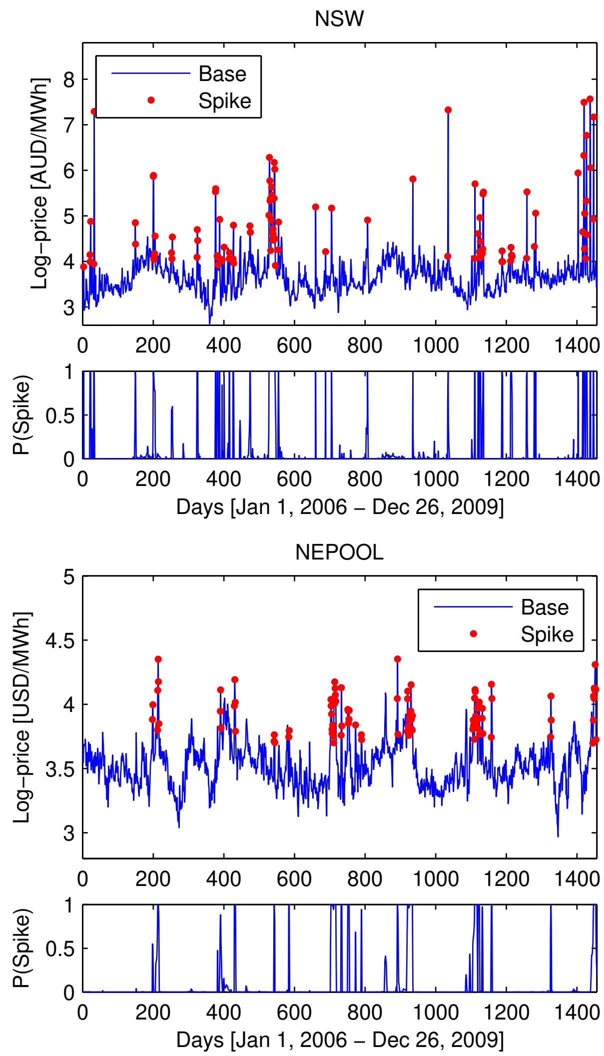

Fig. 1. Calibration results for the MRS models with truncated lognormal spikes and mean-reverting heteroskedastic base regime dynamics fitted to NSW (*top*) and NEPOOL (*bottom*) log-prices. The corresponding lower panels display the probability $P(R_t = 2|x_1, x_2, ..., x_T)$ of being in the spike regime. The prices classified as spikes, i.e. with $P(R_t = 2|x_1, x_2, ..., x_T) > 0.5$, are additionally denoted by dots.

so that the minimum of the new process is the same as the minimum of P_t (the latter alignment is required if log-prices are to be analyzed). The resulting deseasonalized time series X_t can be seen in Figure 1.

IV. Empirical results

The log-prices X_t and the conditional probabilities of being in the spike regime $P(R_t = 2|x_1, x_2, ..., x_T)$ for the analyzed datasets are displayed in Figure 1. The prices classified as spikes, i.e. with $P(R_t = 2|x_1, x_2, ..., x_T) > 0.5$, are additionally denoted by dots. The estimation results are summarized in Table I. Additionally in this table we provide probabilities of staying in each regime p_{ii}, unconditional probabilities $P(R = i)$ of being in regime i, moments and values of the log-likelihood function (LogL). Moreover, in order to evaluate

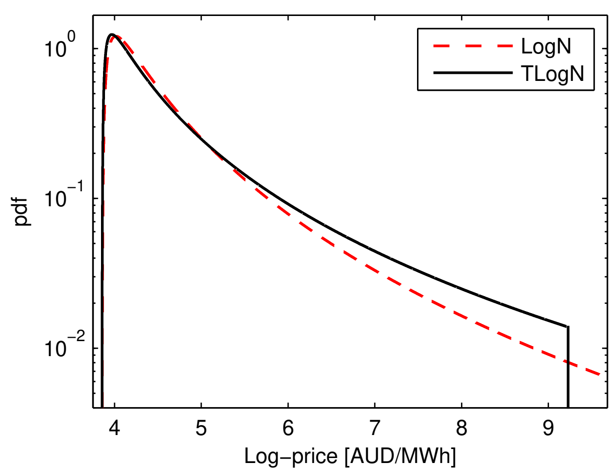

Fig. 2. Comparison of lognormal (red dashed line) and truncated lognormal (black solid line) spike regime probability distribution functions (pdf) fitted to NSW log-prices. The logarithmic scale is used on the y axis to better illustrate the differences in the tails.

the goodness-of-fit, we report the K-S test p-values. For details on the testing procedure see [13].

The results of the K-S tests indicate that acceptable fits are obtained for all considered models, since all p-values are larger than 0.01. Recall, that p-values larger than 0.01 indicate that we cannot reject the hypothesis about the chosen price model at the 1% significance level.

Considering spike occurrences we see a similar picture for both datasets. In each case there are about 8% of log-prices classified as coming from the spike regime. As expected, the probability of remaining in the base regime is very high: from 0.9670 for the NSW log-prices up to 0.9844 for the NEPOOL log-prices. The probability of remaining in the spike regime is lower, but still relatively high.

Regarding the base regime parameters we observe that positive γ was obtained in all cases. This is consistent with the 'inverse leverage effect' reported for electricity prices, reflecting the observation that positive electricity price shocks increase volatility more than negative shocks [11], [15]. The speed of mean reversion, represented by the parameter β, is similar for both datasets and equals 0.23 and 0.24 for the NSW and NEPOOL log-prices, respectively.

Comparing results obtained for the models with lognormal and truncated lognormal spike distributions calibrated to the NSW log-prices, we observe that the base regime parameters and all probabilities are pretty much the same for both models. This suggests that the classification to the base and spike regimes was the same in both cases. The spike regime parameters, however, differ significantly between the truncated and the non-truncated specification. The truncated distribution yields slightly higher α_2 and evidently higher σ_2. This leads to a higher mean and variance in the truncated specification. Especially apparent is the difference of variances: 2.48 in the lognormal model and 8.30 in the truncated lognormal one.

TABLE I

CALIBRATION RESULTS, MOMENTS, K-S TEST p-VALUES AND LOG-LIKELIHOOD (LOGL) FOR THE NSW AND NEPOOL LOG-PRICES

		Parameters				Probabilities		Moments		K-S test p-values		
Regime	Distribution	β	α_i	σ_i^2	γ	p_{ii}	$P(R_t = i)$	$E(X_{t,i})$	$Var(X_{t,i})$	Regime	Model	LogL
						NSW						
Base		0.24	0.52	0.0003	1.8	0.9675	0.9235	3.61		0.01	0.04	499
Spike	LogN		-0.60	1.2263		0.6074	0.0765	4.87	2.48	0.17		
Base		0.24	0.52	0.0003	1.8	0.9670	0.9231	3.61		0.01	0.04	502
Spike	TLogN		-0.52	1.6409		0.6044	0.0769	5.23	8.30	0.27		
						NEPOOL						
Base		0.23	0.47	0.0004	1.2	0.9844	0.9249	3.51		0.22	0.23	1565
Spike	LogN		-1.39	0.3443		0.8079	0.0751	3.93	0.04	0.18		
Base		0.23	0.47	0.0004	1.2	0.9844	0.9249	3.51		0.22	0.23	1565
Spike	TLogN		-1.39	0.3446		0.8079	0.0751	3.93	0.04	0.18		

The estimated spike probability density function (pdf) for the truncated, as well as, the non-truncated model is given in Figure 2. Looking at the goodness-of-fit measures we see a similar picture. While the p-values obtained for the base regime are the same, the truncated specification yields a better fit of the spike distribution (0.27 for the truncated specification versus 0.17 for the non-truncated one). Moreover, the whole model log-likelihood is higher for the specification with truncated spikes. This clearly shows, that the introduction of a truncated distribution might be beneficial when considering spike distributions. It is not only consistent with the market specifications but also provides a better statistical fit.

The calibration results for the less spiky NEPOOL log-prices lead, however, to significantly different conclusions. The various statistics are identical for the truncated and non-truncated specifications. Only σ_2's differ ... but by less than 0.1%. Clearly, even the highest log-prices are still far from the market cap of $\log(1000 \text{ USD}) = 6.91$. As a consequence, the estimated probability of exceeding the market cap is close to zero, implying pretty much the same spike distribution in both cases.

V. CONCLUSIONS

Our empirical study provides evidence that the introduction of a truncated (or price-capped) spike distribution not only is consistent with market regulations, but also can be useful for modeling extremely spiky electricity spot prices. The calibration results for the Australian NSW power market show that there are significant differences between the estimated spike distributions in the truncated and non-truncated cases and that the statistical fit is better for the former. This indicates that the truncation should not be neglected when modeling the NSW market or alike. On the other hand, results obtained for the NEPOOL market show that in case of less spiky electricity prices the truncated and non-truncated specifications lead to similar model estimates and goodness-of-fit.

ACKNOWLEDGMENTS

J.J. acknowledges partial financial support from the European Union within the European Social Fund. The work of R.W. was partially supported by ARC grant no. DP1096326.

REFERENCES

[1] M. Bierbrauer, C. Menn, S.T. Rachev, S. Trück (2007) Spot and derivative pricing in the EEX power market. *Journal of Banking & Finance* 31: 3462-3485.

[2] M. Bierbrauer, S. Trück, R. Weron (2004) Modeling electricity prices with regime switching models. *LNCS* 3039: 859-867.

[3] D.W. Bunn, ed. (2004) Modelling Prices in Competitive Electricity Markets. Wiley, Chichester.

[4] C. de Jong (2006) The nature of power spikes: A regime-switch approach. *Studies in Nonlinear Dynamics & Econometrics* 10(3), Article 3.

[5] A. Dempster, N. Laird, D.B. Rubin (1977) Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society* 39: 1-38.

[6] R. Ethier, T. Mount (1998) *Estimating the volatility of spot prices in restructured electricity markets and the implications for option values*. PSerc Working Paper 98-31.

[7] J. Hamilton (1990) Analysis of time series subject to changes in regime. *Journal of Econometrics* 45: 39-70.

[8] R. Huisman (2008) The influence of temperature on spike probability in day-ahead power prices. *Energy Economics* 30, 2697-2704.

[9] R. Huisman, C. de Jong (2003) Option pricing for power prices with spikes. *Energy Power Risk Management* 7.11: 12-16.

[10] J. Janczura, R. Weron (2009) *Regime switching models for electricity spot prices: Introducing heteroskedastic base regime dynamics and shifted spike distributions.* IEEE Conference Proceedings (EEM'09), doi:10.1109/EEM.2009.5207175.

[11] J. Janczura, R. Weron (2010) An empirical comparison of alternate regime-switching models for electricity spot prices. *Energy Economics*, doi:10.1016/j.eneco.2010.05.008. Working paper version available at MPRA: http://mpra.ub.uni-muenchen.de/22876.

[12] J. Janczura, R. Weron (2010) *Efficient estimation of Markov regime-switching models*. Working paper.

[13] J. Janczura, R. Weron (2010) *Goodness-of-fit testing for regime-switching models.* Available at: http://mpra.ub.uni-muenchen.de/22871.

[14] C.-J. Kim (1994) Dynamic linear models with Markov-switching. *Journal of Econometrics* 60: 1-22.

[15] C.R. Knittel, M.R. Roberts (2005) An empirical examination of restructured electricity prices. *Energy Economics* 27(5): 791-817.

[16] T.D. Mount, Y. Ning, X. Cai (2006) Predicting price spikes in electricity markets using a regime-switching model with time-varying parameters. *Energy Economics* 28: 62-80.

[17] S. Trück S, R. Weron, R. Wolff (2007) *Outlier treatment and robust approaches for modeling electricity spot prices*. Proceedings of the 56th Session of the ISI. Available at http://mpra.ub.uni-muenchen.de/4711/.

[18] R. Weron (2006) *Modeling and forecasting electricity loads and prices: A statistical approach*. Wiley, Chichester.