

A Time Series Analysis of  
Price Formation in Power Markets

by

Ibrahim Khan  
B.A., University of Victoria, 2014

A Thesis Submitted in Partial Fulfillment  
of the Requirements for the Degree of

MASTER OF ARTS

in the Department of Economics

© Ibrahim Khan, 2017  
University of Victoria

All rights reserved. This thesis may not be reproduced in whole or in part, by  
photocopy or other means, without the permission of the author.

A Time Series Analysis of  
Price Formation in Power Markets

by

Ibrahim Khan  
B.A., University of Victoria, 2014

**Supervisory Committee**

Dr. Judith A. Clarke, Department of Economics

**Supervisor**

Dr. G. Cornelis van Kooten, Department of Economics

**Departmental Member**

## Abstract

This study examines price formation in one of the largest wholesale electricity markets in the world: the Pennsylvania Jersey Maryland Interconnection, which serves 13 states and the District of Columbia with over 60 million consumers. The contribution of this thesis is to apply a variety of time series models offered in the literature to a large data set describing a single market, allowing for a comparison of their performance as well as demonstrating their validity. A central question that drives market deregulation is if it has created efficiency gains. To formalize this notion of efficiency, we implement tests for stationarity to measure the degree of randomness over time, finding that short run volatility can result in the outcomes for these tests that are inconclusive. We explore this volatility structure using Asymmetrical Power Autoregressive Conditional Heteroskedastic (APARCH) framework which captures the asymmetric nature of price shocks, finding that this behavior is unique to electricity returns, and that APARCH offers a better modelling alternative than simpler representations. Additionally, we account for long memory given the seasonal drivers of electricity prices which are persistent using Autoregressive Fractionally Integrated Moving Averages (ARFIMA). Temperature related market drivers are further modelled using Fourier based seasonality functions which enable us to capture cycles over multiple frequencies. Lastly, we provide an application of Markov Regime Switching models to account for the possibility of multiple states. Although appealing from a theoretical perspective, we find that the increased complexity of the model does not necessarily translate to better performance over simpler non-switching alternatives. These findings highlight the importance of establishing the features of the time series before selecting an appropriate model, and motivating it with economic rationale.

## **Table of Contents**

- 1. Introduction**
- 2. Market Structure**
- 3. The Economic Theory**
- 4. Data**
- 5. Descriptive Statistics**
- 6. Asymmetric Volatility**
- 7. Diagnostic Tests**
- 8. Fractional Integration**
- 9. Seasonality**
- 10. Markov Regime Switching**
- 11. Conclusion**
- 12. References**

## Tables

**Table 1:** Power Production Function

**Table 2:** Descriptive Statistic of Financial and Electricity Returns

**Table 3:** Descriptive Statistic of Financial and Electricity Prices

**Table 4:** GARCH Results for Financial Returns

**Table 5:** GARCH Results for Electricity Returns

**Table 6:** APARCH Results for Financial Returns

**Table 7:** APARCH Results for Electricity Returns

**Table 8:** ADF Unit-Root Test Results

**Table 9:** PP Unit-Root Test Results

**Table 10:** KPSS Stationary Test Results

**Table 11:** ADF Unit-Root Test Results of Differenced Series

**Table 12:** PP Unit-Root Test Results of Differenced Series

**Table 13:** KPSS Stationary Test Results of Differenced Series

**Table 14:** Lo's Modified R/S Long Range Dependency Test Results

**Table 15:** GPH Estimate of the Order of Fractional Integration

**Table 16:** Lo's Modified R/S Test Results of Differenced Series

**Table 17:** GPH Estimate of the Order of Integration of Differenced Series

**Table 18:** Seasonality Regression with Dummy Variable

**Table 19:** Seasonality Regression with Sinusoids

**Table 20:** RT Off Peak MR

**Table 21:** RT Peak MRS

**Table 22:** DA Off Peak MR

**Table 23:** RT Peak MRS

**Table 24:** Non-Switching AR(1)

## Figures

**Figure 1:** Average Hourly Loads

**Figure 2:** Distributions of Financial and Electricity Returns

**Figure 3:** Hypothetical PJM Generation Stack

**Figure 4:** News Impact Curves of Financial Returns

**Figure 5:** News Impact Curves of Electricity Returns

**Figure 6:** Autocorrelation Function Plots of First Differences of Price Series

**Figure 7:** Long Run ACF Plots of the Average Daily Prices in RT and DA

**Figure 8:** Autocorrelation Function Plots of Representative Price Series

**Figure 9:** Electricity Time Series Plots

**Figure 10:** Peak Hour Fourier Seasonality

**Figure 11:** Off Peak Hour Fourier Seasonality

**Figure 12:** Load Fourier Seasonality

**Figure 13:** RT Off Peak Switching Regimes Probabilities

**Figure 14:** RT Peak Switching Regimes Probabilities

**Figure 15:** DA Off Peak Switching Regimes Probabilities

**Figure 16:** DA Peak Switching Regimes Probabilities

## **1. Introduction**

The reliable production and transmission of electricity is a necessary condition for the industrial advancement of any society in the modern era. Electricity generation is capital intensive, and like many high fixed cost industries, its market has traditionally been organized as a natural monopoly with the public sector taking a lead role in the maintenance of its critical infrastructure. This has changed in recent decades as advanced economies have deregulated the industry with varying degrees of liberalization. For instance, in the United States, the supply chain is composed of three types of utilities: public or private generators owned by shareholders or municipalities, co-ops owned by consumers, and federally owned agencies. The nation's power supply is divided into three alternating current grids: the Eastern, Western, and Texas Interconnections, these are sub-divided into regional markets or Independent System Operators (ISOs). This study examines price formation in one of the largest wholesale electricity markets in the world: the Pennsylvania Jersey Maryland Interconnection, hereafter referred to as the PJM, which serves 13 states and the District of Columbia with over 60 million consumers. Changes to these market's structure and trading rules are still on-going as regulators continually try to address the economic welfare implications of market participant strategies.

This liberalization of the industry has resulted in electricity becoming a widely traded commodity, with market participants tasked with making optimal production decisions in which they take financial risk as they manage large portfolios of physical assets. These market participants must be able to accurately capture the price and supply dynamics of the part of market they operate in, in order to maximize return from the capital-intensive resources they deploy. This requires that risks can be quantified, and accurately measured, which is particularly challenging in electricity markets as its price display features that are distinct from other financial markets. Electricity is a commodity that can be transported instantaneously, but cannot be stored economically. This requires that generators which produce energy, and load serving entities that buy it in the wholesale market, be able to forecast demand precisely and schedule delivery accordingly. This can

be problematic from a market organization perspective, as the ability to increase output is limited by physical constraints, whereas demand can shift rapidly. This inherent risk is mitigated by a market structure that organizes and centrally clears demand in addition to financial contracts such as derivatives. This, however, requires robust modelling of the price process of electricity, a difficult exercise given the statistical challenges related to volatility, especially when trying to disentangle its deterministic and stochastic components.

The contribution of this thesis is to apply a variety of time series models offered in the literature to a large data set describing a single market, allowing for a comparison of their performance as well as demonstrating their validity. A central question that drives market deregulation is if it has created efficiency gains. When considered in the financial context, efficiency implies that prices are random if participants have already acted on any predictability within them. To formalize this notion of efficiency, we implement tests for stationarity to measure the degree of randomness over time, finding that short run volatility can result in the outcomes for these tests that are inconclusive. We also highlight how this volatility is measurably distinct from other financial markets using models that can capture its asymmetrical nature. Since stationarity is a useful condition for many autoregressive time series models and is difficult to establish given the observed volatility, we also apply fractional integration methods, demonstrating that electricity prices can exhibit long memory. These representations imply that prices are correlated over long periods of time, suggesting that they could be influenced by drivers that are somewhat constant over time. To further motivate this, we also examine the seasonal characteristics of both prices and volumes using trigonometric representations which offer a more nuance and granular alternative for capturing seasonality than conventional dummy variables. Lastly, we implement regime switching models which posit that the market behaves in a distinct way depending on whether it is in equilibrium or not.

This thesis is organized as follows: section two provides a summary of the market structure that enables participants to trade a non-storable commodity. We

describe the bidding mechanism that participants use, and the choices this offers them for buy and selling into the market. Section three describes the economic framework that governs the optimal decision of producers. In addition, we test if the market marginal cost curves is convex as would be implied by the type of resources used to produce energy, and implement three different specifications for representing price as a function of quantity, to which we apply to the observed data. In section four we give an overview of the data utilized in this thesis. Section five presents descriptive statistics for this data, demonstrating that electricity prices are statistically distinct from financial assets. In section six we further examine if electricity price returns behave distinctly from financial returns by testing for asymmetrical volatility, demonstrating that the convexity of supply can result in higher volatility following upward, rather than downward price shocks. In section seven we implement three tests that explore for stationarity. Two of the tests assume a unit-root null hypothesis, and one test that assumes stationarity as the null hypothesis. We apply these tests to 450 different sub-samples finding inconclusive outcomes.

In section eight we provide an alternative representation of the series that incorporates the notion that series exhibit behavior that is between stationary and non-stationary. We apply the modified rescaled range test for long memory, in addition to generating estimates for the order of integration. These indicate that fractional integration methods offer more appropriate representations, given the observed long memory in the series. In section nine we use Fourier based seasonality functions that define the series as a sum of sinusoidal waves. This method adequately captures the cyclical nature of volume, but less so for prices. It is, however, still an improvement over dummy variable methods, as it does not assume that the means change in discrete shifts. In section ten, we motivate that the distinct volatility observed during peak demand indicates that the market can exist in multiple states, or regimes. We implement a two-state Markov regime switching representation of an autoregressive time series model that allows for all parameters switch between regimes. Despite the economic rationale behind the existence of

multiple regimes, we find that they do not outperform simpler non-switching alternatives. We conclude in the last section.

## 2. Market Structure

The demand and price uncertainties faced by electricity producers and retailers play a central role in how these markets are structured. Electricity is characteristically distinct from other tradable commodities because the underlying asset involves the flow of energy that is non-storable. This also means that the energy inputted into the electrical grid, which can be thought of as a central market or pool, must at any point in time equal the energy taken out of the system through consumer use; this implies a maintenance of constant equilibrium between supply and demand, a feature which fundamentally drives its price formation. Consequently, these markets often use a spot-forward settlement system where the forward delivery is only 24 hours ahead of the spot market, with these two markets being referred to as the real-time (RT) and day-ahead (DA) markets.

This structure forms the two choices in the optimal forward position model of Bessminder and Lemmon (2002) that we use in this study as the economic theory governing price formation. The multi-settlement system used in the PJM gives participants the option to submit bids and offers for delivery in DA, as well as the ability to buy and sell megawatt hours (MWh) for delivery in the RT. This settlement system can be thought of as a central clearing mechanism that allows for the ISO to coordinate production decisions in an organized and structured process to ensure the reliability and efficiency of the market as a whole.

The energy sent over each hour of the day is treated as an independent asset, so implying that there are 24 different assets in each of the RT and DA markets. Bids in the DA market (for electricity to be delivered at some hour  $i$  on day  $t = 2^1$ )

---

<sup>1</sup> The electricity market data takes  $t$  as the time series variable representing the day of the observation.

are placed with forecasts on demand for that hour made at  $t = 1$  with information set  $\Omega_t$ . Once  $t = 2$  arrives, the difference between the predicted demand given information set  $\Omega_t$  and actual demand realized in information set  $\Omega_{t+1}$  is reconciled in the RT market at  $t = 2$ . Market participants can bid continuously in the RT, where prices are booked, or settled, every 5 minutes and the corresponding quantity is delivered hourly (as each hour is a separate asset). Price formation in the DA market is largely driven by the PJM's centralized matching algorithm. Unlike equity markets where the National Best Bid and Offer (NBBO) regulation demands minimization of bid/ask spreads, in electricity markets the objective of central clearing is to equalize quantity supplied ( $q^S$ ) and quantity demanded ( $q^D$ ). Therefore, the matching algorithm minimizes the spread between  $q^S$  and  $q^D$ , allowing the price to fluctuate instead. This rule is a distinct feature in electricity markets globally and has the aim of limiting situations where power is ultimately withheld, or not supplied, due to unprofitable generation in the short run.

In the PJM's DA, each market participant is required to submit ten price/quantity combinations to the ISO by noon the day before its delivery. Bidding in the DA is then closed from 12pm to 4pm while the PJM runs a matching algorithm in the form of an optimization problem where bids and offers for quantity are matched subject to minimizing Locational Marginal Prices (LMPs)<sup>2</sup>, the fuel needed for generation, and the spread between  $q^S$  and  $q^D$ . The prices for each hour of the day are then posted at 4pm where re-bidding is allowed for two hours, and at 6pm the final DA prices are posted. We now present some economic theory that motivates our analysis.

### 3. The Economic Theory

---

<sup>2</sup> LMPs occur when the market is segmented due to congestion or transmission constraints resulting in multiple prices.

Applying market equilibrium based models to explain fundamental price behavior in power markets originates from the seminal work of Bessminder and Lemmon (2002), who theorize that price formation is better explained by the decisions of physical market participants (who actually give and take delivery) than by the actions of speculators. Explaining price formation using the objective functions of both sides of the market is an example of working within the systematic approach. This framework assumes that the prevailing strategy of producers is to minimize volatility in their cash flows, in addition to maximizing profits. This means that a physical market participant's risk minimization (or optimal) strategy is ultimately a function of the higher statistical moments about the means of spot and forward prices.

In this setup, the link, or equilibrium, relationship between the RT and DA price processes requires two assumptions: 1) that demand is an inelastic function of exogenous factors such as temperature, and 2) that total cost (TC) of supply for each producer is a convex function of quantity and fixed costs (FC). Such an economic framework motivates a non-linear time series model. Specifically, if prices are a non-linear function of quantity, then even small, perhaps even predictable, movements in demand can result in significant price changes. To examine this, following Bessminder and Lemmon (2002), we estimate a production function to measure the convexity of the market production cost curve in the PJM. Their methodology assumes a profit function  $\pi$  for each producer  $i$  such that:

$$\pi_i = \underbrace{P_{RT}Q_{RT,i} + P_{DA}Q_{DA,i}}_{Revenue} - \underbrace{\left(FC + \frac{a}{c}(Q_{RT,i} + Q_{DA,i})^c\right)}_{Costs}.$$

Here  $P_{RT}$  and  $P_{DA}$  are market RT and DA prices as each firm is a price taker,  $FC$  reflects fixed costs,  $Q_{RT,i}$  and  $Q_{DA,i}$  are the quantities sold in the RT and DA, and  $Q_{Tot,i} = Q_{RT,i} + Q_{DA,i}$ . The parameters  $a$  and  $c$  are constants, ensuring that TC are not a linear function of  $Q_{Tot,i}$ . The profit maximizing quantity sold in the spot market, as a function of spot prices, is then:

$$Q_{RT,i} = \left(\frac{P_{RT}}{a}\right)^{\frac{1}{c-1}}$$

Bessminder and Lemmon (2002) assume that total forward contracts are in zero net supply, and that total retail demand equals total supply. They provide the following form of the above relationship between spot prices and quantity:

$$P_{RT} = a \left( \frac{Q_{Tot}}{N} \right)^{c-1}$$

where  $N$  is the total number of producers, each with identical production technology. In their implementation, however, the average load observed over a day is used for  $\frac{Q_{Tot}}{N}$  to be able to estimate the production cost function. Further, the right side of this equation is linearized using natural logs, and prices observed in the Peak RT market for each day  $t$ :

$$\ln(\bar{P}_{RT\ Peak,t}) = a + (c - 1)\ln(\bar{Q}_{Tot,t}) + \varepsilon_t.$$

This equation includes a disturbance term  $\varepsilon_t$  to account for the fact that transmission constraints can result in price spikes during times of low system load. The price used in the regression is  $\bar{P}_{Peak\ RT}^t$ , the average price for peak hours on day  $t$  (5PM to 7PM), and the quantity  $\bar{Q}_{Tot}^t$  is the average load for that day<sup>3</sup>. Our model is consistent with that adopted by Bessminder and Lemmon (2002), who applied this model to the California Power Exchange (CALPX) and PJM<sup>4</sup>.

Three different specifications are estimated: one equation that includes a time trend  $t$  because of our longer span of data; the second specification includes a time trend with dummy variables indicating the month of the year to capture possible seasonal effects; and a third model that does not include any time trend or monthly dummy variables. Seasonality is incorporated because of the seasonal variation in production decisions that occur when producers schedule maintenance during low

---

<sup>3</sup> Note: this regression is taken ex-post, meaning that realized spot prices associated with the observed load with that day-hour combination are used.

<sup>4</sup> The equation used in Bessminder and Lemmon (2002) takes the form:

$\ln(P_t) = a + (c - 1)\ln(Q_t + \mathbf{d}_{it}) + \varepsilon_t$  where  $\mathbf{d}_{it}$  is a vector of dummy variables indicating the month  $i$ .

volume times of the year. The parameter estimates for  $(c - 1)$ , shown in Table 1, suggest that  $c$  falls between 2.4 and 3 depending on the specification of seasonality. This is significant because if  $c > 2$ , then the production cost function is convex with respect to quantity. This also means that the distribution of prices will have a positive skew, even with a symmetric distribution of  $Q_{Tot}$ , given that production cost function increases significantly when there is insufficient supply to meet higher quantities of demand. These results are in line with Bessminder and Lemmon (2002), who estimate  $c$  to be 4.8 in the PJM, and 5.81 in the CALPX (both estimates also greater than 2). These results indicate that production costs are convex with respect to output, and go up at an increasing rate as quantity increases, and that this convexity has diminished in the PJM since 2002. This finding will be explored further by analyzing asymmetric volatility but first we examine the descriptive statistics that support this framework.

It is important to qualify these findings with the fact that Besminder and Lemmon's (2002) research, on which ours is based, assumes a simplified TC function that is uniform for all producers, which is not the case in actuality. Given the diverse fuel mix present in electricity markets, especially well-developed ones as the PJM, both fixed and variable costs vary across producers. In reality, screening curves, which reflect the revenue requirement for a specific fuel source to be profitable for a given quantity, in addition to duration curves, which reflect the distribution of demand for each hour, are two of many inputs to the large-scale optimization problem that drives the production decision. The matching algorithm cited in Section 2 is one of these optimization problems managed at the ISO level. In the application above, focus is given to the spot market, which assumes forward contracts are in zero net supply, which means that real-time production decisions are ex-post to the clearing of the day-ahead market. The least cost production decision to supply electricity in the real-time market therefore requires that the fuel source can be deployed immediately, or that excess spinning reserves exists for fuel sources already being used to meet expected demand. The notion that price

formation is driven by a convex production cost curves, which can result in asymmetrical volatility is further explored in Section 6.

#### 4. Data

The electricity market data in our study is sourced directly from the PJM's web based data mining tool<sup>5</sup>. We consider aggregate (market-wide) data for three main variables: prices in the spot market (RT); the one-day forward market (DA); and quantity for each hour of the day, which combines the volume sold in both markets. Volume also serves as the demand variable due to the general convention in electricity markets that requires all demand bids to be supplied by producers. Prices for the RT and DA are measured in \$/MWh. Each hour of the day is viewed as an independent asset, in other words there are 24 separate daily time series for each of the RT, DA, and Load variables. This panel treatment of the data, where the “cross section” variable is the hour of the day  $i$ , and the “time series” variable is the day  $t$ , is motivated by the market's structure, where each hour is traded independently. Our approach is consistent with the literature that accounts for intraday heterogeneity; e.g., Barlow (2002) and Karakatsani and Bunn (2008). Alternative methods for structuring data can involve averaging across all hours of the day (e.g., Kanamura and Ohashi, 2008), using the maximum observed price (e.g., Rambharat et al., 2005), or an hourly series but with autoregressive lags that are a multiple of 24 (e.g., Weron and Misiorek, 2008). It is important to note that these series represent market-wide information, that is, the data excludes LMPs. Our data, which spans from July 2007 to February 2016 with 75,984 observations (or 3166 observations per hour) for each variable, allows us to observe the market over multiple years and business cycles making seasonality analysis more robust, as its highly cyclical behavior typically requires long spans of data to be visible.

---

<sup>5</sup> <https://dataminer.pjm.com/dataminerui/pages/public/energypricing.jsf>

We use the S&P500 Index, USD/EUR (\$/€) exchange rate, spot natural gas prices at the Henry Hub (\$/MBtu), and West Texas Intermediate crude oil (\$/barrel) as heuristics in this study. We use these asset classes because they are either storable or can be borrowed, and represent liquid markets, thereby allowing us to examine how these characteristics make electricity distinct in its price formation. We provide descriptive statistics for both log returns and prices in this section. It is important to note that due to the non-storability property of electricity, it does not technically accrue a return over a time horizon per se as it is not being held like other financial assets. Although electricity generation assets, and financial assets can be bought, held, and sold to accrue a return, this is not the case with electricity itself. We use log returns as they allow for a common basis of comparison between the different asset classes since indexes, exchange rates, and prices represent different units. In addition, the volatility framework applied in this study has conventionally used returns. The log returns series are generated from the raw series using first differences of the natural logarithm of each observation, in other words, for a generic series  $P_t$ , then  $r_t = (1 - L)\ln(P_t)$ , where  $L$  is the usual lag operator. As there are 48 different price series for the PJM (24 each for the RT and DA markets), representative hours of the day are considered. Since demand fluctuates throughout the day, each of these price series and their associated hour have distinct supply and demand characteristics that cycle throughout the day. To accommodate for this, we use the four combinations of Peak, Off-Peak, RT and DA as categories to capture supply and demand heterogeneity. On this criteria, the four representative series used are the 3am RT, 6pm RT, 3am DA, and 6pm DA, chosen as these hours represent the minima and maxima of load (see Figure 1).

## 5. Descriptive Statistics

Descriptive statistics for returns and prices are shown in Table 2 and 3 respectively. A cursory examination of the first through fourth moments shows the extent to which power markets display significantly distinct behavior from other assets. All four electricity price series show positive skewness as suggested by our convexity

analysis. The electricity price and return series display the highest sample variances,  $\sigma^2$ , which is also shown as annualized volatility defined by  $\frac{\sigma}{\sqrt{T/365}}$ , and  $\frac{\sigma}{\sqrt{T/250}}$  for electricity and financial assets respectively with  $T$  equal to the number of observations.<sup>6</sup> A major contributor to higher volatilities in power markets is the presence of outliers, or price spikes, which likely also explains the higher observed sample kurtosis of electricity returns. The RT prices display the largest volatility amongst the series, with this being significantly higher than that for DA prices. We expect this feature given that these markets are used to balance supply and demand. Transmission constraints, inaccurate demand forecasts, and the inability to use inventories to smooth the mismatches that occur between  $q^S$  and  $q^D$  require purchases in the RT market, resulting in non-linear and erratic behavior in that market. Day ahead markets display less variation because of the bid-offer matching algorithm that implicitly involves averaging prices submitted to the market.

Electricity returns also display significantly more skewness than financial returns, again likely due to the inability to arbitrage, and the convexity of the marginal cost curve. The Jarque and Bera (1987), or JB test is applied to determine whether the series are normally distributed; given our findings so far, we do not anticipate this to be the case. Using sample standardized skewness  $S$ , and kurtosis  $K$ , the test statistic takes the form<sup>7</sup>:

$$\frac{T}{6} \left( S^2 + \frac{1}{4} (K - 3)^2 \right)$$

with an asymptotically  $\chi_2^2$  null distribution with 2 degrees of freedom. All assets strongly reject the null hypothesis of normality, with estimated JB statistics significantly greater than the 1% critical value of 9.21.

---

<sup>6</sup> As financial assets are traded on weekdays, and power markets operate 24/7, a simple volatility scaling function is implemented to enable comparison of estimates from data taken over the same span, but with different number of observations  $T$ . It is important to note that given the non-normality of prices, the square root scaling of volatility can generate downwardly biased estimators of annualized volatility (see, e.g., Dacorogna et al. 2001).

<sup>7</sup> Here  $T$  is the length of the series  $y_t$ , with  $S = \frac{1}{T} \sum_{t=1}^T \left[ \frac{y_t - \bar{y}}{\sigma} \right]^3$  and  $K = \left[ \frac{1}{T} \sum_{t=1}^T \left[ \frac{y_t - \bar{y}}{\sigma} \right]^4 \right]$ .

A portmanteau Ljung and Box (1979), or LB, test is also undertaken to determine the degree of autocorrelation persistent in these series. Let  $\rho_k$  be the autocorrelation between a series, say  $p_t$ , and its  $k$ 'th lag,  $p_{t-k}$ . The LB  $Q_k$ -statistic examines the null hypothesis  $H_0: \rho_1 = \dots = \rho_k = 0$  against the alternative hypothesis that at least one of the autocorrelations is non-zero. The test statistic is:

$$Q_k = T(T + 2) \sum_{j=1}^k \frac{\hat{\rho}_j}{T - j}$$

where  $\hat{\rho}_j$  is a consistent estimator of  $\rho_j$ . This  $Q_k$ -statistic is asymptotically  $\chi_k^2$  distributed with  $k$  degrees of freedom. Lag choices of  $k = 30, 60,$  and  $90$  are used in our tests. The results for the LB test with  $k = 30, 60,$  and  $90$  lags (with 1% critical values of 50.90, 88.38, and 122.95 respectively) for both return and price series are presented in Tables 2 and 3, respectively. These indicate that, with the exception of EUR/USD returns, all return and price series reject the null for all lag lengths at their 1% critical values. These results suggest that the autocorrelation structures are significant, and need to be accounted for in any modeling.

From our examination of these descriptive statistics and preliminary explorations, we conclude that electricity prices and returns display significantly greater dispersion than those of financial assets. Additionally, they are generally more right skewed as well, driven by price spikes that produce extreme values. This is also evident in their generally higher sample kurtosis statistics. Additionally, the series display significant autocorrelation structures, except for exchange rates, perhaps not surprisingly as currencies are extremely liquid and large markets that are unlikely to retain long memory. Lastly, the presence of skewness and relatively high variances in electricity returns seems to call for a more granular examination of their volatility processes, as these appear to be unique features. We investigate these characteristics in the following section.

## 6. Volatility Analysis

As shown in Figure 2, the magnitude of volatility of PJM returns far exceeds that of the other series presented in financial and commodity markets. We explore some of the patterns in the volatility process in this section. Specifically, we consider if the non-storability of electricity affects its market volatility in ways that are distinct from markets in which the asset can be held. To examine this question, we adopt the autoregressive conditional heteroscedasticity (ARCH) framework of Engle (1982), and generalized in Bollerslev (1986). This modelling structure, along with the many extensions proposed in the literature, is broad, allowing us to integrate stylized facts into its methodology to conduct hypothesis tests about forms of volatility. The ARCH family of models formalizes the observation in financial returns that periods of low (high) volatility are often followed by continued periods of low (high) volatility.

To apply this to our context, we begin with the notion that there exists an upper limit on how much energy can be readily produced and delivered with the fuel sources already supplying energy to the market. During times of peak demand, less efficient, and, therefore, more expensive fuel sources come online. The implication being that periods of high volatility due to peak demand are associated with positive returns resulting from price spikes. Figure 3 displays a hypothetical marginal cost curve, or generation stack, representing the costs of the fuel sources in the PJM that motivates this notion. By using a volatility model that relaxes the assumption that positive and negative returns have an equivalent marginal impact on volatility, we can test whether there is an upward kink in the marginal cost curve. We demonstrate this by first describing the general volatility framework and then establishing the presence of asymmetrical volatility using diagnostic tests within that framework. Lastly, we use these results to assist specify a volatility model that can incorporate these stylized facts: the asymmetric-power ARCH (APARCH).

Black (1976) first examined asymmetric volatility in the context of volatility forecasting in stocks, terming the negative correlation between returns

and volatility as the “leverage effect<sup>8</sup>.” An important aspect of which is the direction, i.e., whether negative or positive shocks are associated with increased volatility. In financial markets, increased volatility is generally associated with negative shocks rather than positive ones. The current literature on asymmetric volatility in electricity returns shows overwhelmingly that the sign of the leverage effect is negative, using both hourly and daily prices. For example, Higgs (2005) examined this effect in Australian markets using GARCH, IGARCH, and asymmetric power ARCH (APARCH) models, with both Gaussian and Student-t distributions for the error terms, given the high degree of apparent kurtosis in their return series. In addition to incorporating deterministic seasonality variables in their variance equations, they find that positive shocks are associated with greater volatility than negative shocks, also called the “inverse leverage effect.” Monetro, Garcia-Centeno, and Fernandez-Aviles (2011) apply threshold asymmetric autoregressive stochastic volatility (TA-ARSV), the asymmetric GARCH (AGARCH), and exponential GARCH (EGARCH) models to daily returns in the Spanish electricity spot market. Their results show, irrespective of adopted model, the presence of inverse leverage effects, with the AGARCH model generating parameter estimates in the variance equation summing to greater than unity, implying non-stationarity. This implies that variance is persistent, and that IGARCH models may be better suited to capture the dynamics of volatility given that the effect of shocks on forecasts of conditional variance may be permanent and not temporary. In addition, they find that the TA-ARSV out-performs the EGARCH model based on commonly examined information criteria.

In U.S. markets, Hadsell, Marath, and Sawky (2004) find that the PJM, Palo Verde, Cinergy, Entergy, and California-Oregon markets also exhibit inverse leverage effects using daily returns and a threshold ARCH (TARCH) model. Erdogdu (2016) apply the EGARCH and TGARCH to price from fourteen different

---

<sup>8</sup> Here “leverage” refers to the debt to equity ratio of publicly traded companies, which increases as prices decline, thereby increasing the risk profile of the security. This in turn would cause them to be more volatile.

European wholesale markets with price data for multiple hours of the day. A total of 96 country and time period combinations were used, with their results showing that 68 out of the 96 (~71%) estimated models exhibit inverse leverage effects. They also find that the magnitude of the shock has a greater effect on volatility than the direction of the shock.

In financial markets, price shocks can be defined as the sudden inclusion of news or information not previously incorporated into the security's price. Although prices shocks are unexpected, market participants do have a history of returns series to forecast their marginal impact on volatility. The ARCH family of models formalizes this idea using an autoregressive time series model of the return series that separates expected returns from unexpected returns. More formally, the expected return  $m_t$  to occur at time  $t$  is the conditional expected value of  $r_t$ , given the historical information set  $\Omega_{t-1}$ , in other words:  $m_t \equiv E(r_t|\Omega_{t-1})$ . The conditional variance of the expected return is then  $s_t^2 \equiv Var(r_t|\Omega_{t-1})$ . The price shock, or unexpected return, is the difference between the expected and observed returns:  $\varepsilon_t \equiv r_t - m_t$ . If applied to the RT market's return series, a positive  $\varepsilon_t$  implies that an insufficient amount of load was sold in the DA market the day prior for the corresponding hour in the RT market. Since  $q^S$  and  $q^D$  must be held in constant equilibrium, any shortage in supply would cause price spikes as the demand curve intersects the supply curve at the upwardly kinked portion of the marginal cost curve.

To model some of these ideas, we first estimate a GARCH(1,1) specification that does not incorporate asymmetry into the responses. The mean and variance functions are:

$$\begin{aligned} r_t &= m + \varphi r_{t-1} + \theta e_{t-1} + e_t \\ \varepsilon_t &= s_t e_t ; e_t \sim iid, N(0,1) \\ s_t^2 &= \omega + \alpha \varepsilon_{t-1}^2 + \beta s_{t-1}^2 \end{aligned}$$

The  $\alpha$  and  $\beta$  parameters represent the ARCH and GARCH effects respectively, implying stationarity if  $\alpha + \beta < 1$ . The error term of the mean equation in this simpler specification is assumed to be i.i.d. and unit normal. Tables 4 and 5 display

parameter estimates for the GARCH(1,1) model fitted to financial and electricity returns respectively. Estimates for  $\alpha + \beta$  are close to 1 for all series, suggesting that shocks are highly persistent. This also indicates the GARCH process is likely to have a high degree of kurtosis if the assumption of stationarity is invalid<sup>9</sup>, we examine the concept of stationarity of prices in a later section.

In order to determine if positive and negative shocks need to be distinguished from each other in the variance equation, we apply the Engle and Ng (1993) diagnostic tests to detect the presence of leverage effects and their direction. These include the Sign Bias, Positive Size Bias, and Negative Size Bias tests. These tests regress the residual series of the GARCH model against dummy variables which code the direction of price shocks independently. If the coefficients of the dummy variables have significant explanatory power, then the original GARCH equation is likely misspecified by not distinguishing the asymmetric responses to shocks. The sign bias test examines if there is sufficient evidence to support incorporating the direction of the shock into the model. The negative and positive size bias tests examine the impact of negative and positive shocks respectively. This methodology codes dummy variables to indicate the direction of the shocks. Specifically, with a  $\hat{\cdot}$  denoting an estimator, let:

$$\begin{aligned} \hat{z}_t &= \hat{\varepsilon}_t / \hat{\sigma}_t: && \text{normalized residual} \\ S_{t-1}^- &= \begin{cases} 1; & \hat{\varepsilon}_t < 0 \\ 0; & \hat{\varepsilon}_t \geq 0 \end{cases}: && \text{dummy variable indicating a negative shock} \\ S_{t-1}^+ &= 1 - S_{t-1}^-: && \text{dummy variable indicating a positive shock} \end{aligned}$$

Then, consider the following auxiliary regressions:

$$\begin{aligned} \text{Sign Bias:} & \quad \hat{z}_t^2 = a_1 + b_1 S_{t-1}^- + u_{1t} \\ \text{Negative Size Bias:} & \quad \hat{z}_t^2 = a_2 + b_2 S_{t-1}^- \hat{\varepsilon}_{t-1} + u_{2t} \\ \text{Positive Size Bias:} & \quad \hat{z}_t^2 = a_3 + b_3 S_{t-1}^+ \hat{\varepsilon}_{t-1} + u_{3t} \\ \text{Joint Test:} & \quad \hat{z}_t^2 = a_4 + b_{4,1} S_{t-1}^- + b_{4,2} S_{t-1}^- \hat{\varepsilon}_{t-1} + b_{4,3} S_{t-1}^+ \hat{\varepsilon}_{t-1} + u_{4t} \end{aligned}$$

---

<sup>9</sup> This can be shown by defining kurtosis with the parameters of the GARCH(1,1) process:  
 $k = \frac{3(1+\alpha+\beta)(1-\alpha-\beta)}{1-\beta^2-2\alpha\beta-3\alpha^2}$ , which is an increasing function with respect to  $\alpha + \beta$  if they sum greater to one.

Then, for the first three regressions, we examine:

$$H_0: b_j = 0 \Rightarrow \text{symmetric volatility}$$

against

$$H_1: b_j \neq 0 \Rightarrow \text{asymmetric volatility}, j = 1, 2, 3$$

The test statistics for the sign bias, negative size bias, and positive size bias test are defined as the squares of the t-ratios of the  $b_j$  parameter in the above regressions, The bias tests statistics follow a  $\chi^2$  distribution, with one degree of freedom, and  $u_{jt}$  is assumed to be i.i.d.. A Lagrange Multiplier joint test, which examines the null hypothesis that  $b_{4,1} = b_{4,2} = b_{4,3} = 0$  is also undertaken, with support for the null hypothesis suggesting that the data are compatible with the originally specified volatility model. Our results for these tests, using a GARCH specification for the volatility model, are shown in Tables 4 and 5 for the financial and electricity time series respectively. These outcomes suggest a positive sign bias is present in the SP500 series, whereas a negative sign bias appears in both the DA Peak and DA Off Peak markets. All other series do not reject the null hypothesis of symmetry.

Given these outcomes, we also specify a more complex volatility model that can include certain stylized facts about the markets, specifically, we incorporate leverage effects into the model, distinguishing between negative and positive shocks. This is achieved using an indicator function, as adopted by Glosten, Jagannathan, and Runkle (1993) for instance, or by taking a transformation of the shocks and/or conditional volatility in the variance equation, as employed by, for example, Nelson (1991) and Engle et al. (1990). In our case, we choose to apply the APARCH model of Ding et al. (1993), which defines the conditional standard deviation as a function of lagged absolute residuals, and uses a power transformation to linearize the variance equation. The mean and variance equations for the APARCH (1,1) model are:

$$\begin{aligned} r_t &= \mu + \varphi r_{t-1} + \theta e_{t-1} + e_t \\ \varepsilon_t &= s_t e_t ; e_t \sim iid, t(0, \frac{v}{v-2}) \end{aligned}$$

$$s_t^\delta = \omega + \alpha(|\varepsilon_{t-1}| - \gamma\varepsilon_{t-1})^\delta + \beta s_{t-1}$$

Many GARCH models can be nested using this functional form by setting parameter restrictions on  $\delta$  or  $\gamma$ . Due to the high degree of sample kurtosis in the return series, particularly in commodity assets, we assume a Student-t distribution for the random error term where  $e_t \sim t(0, \frac{v}{v-2})$  with  $v$  degrees of freedom. In our estimation of the APARCH model, both  $\delta$  and  $\gamma$  are free parameters to be estimated. The  $\delta$  exponent is a power transform that eases the assumption that the true model is linear in the parameters<sup>10</sup>, which may be a more appropriate specification given how severe the price shocks in the RT can be. These spikes indicate that the forecast of the conditional variance may be non-linear with respect to previous shocks, i.e. small price shocks are likely to be followed by larger ones. Additionally, Ding and Granger (1996) demonstrate that ARCH models that include a power transform outperform other specifications based on likelihood values when the Taylor effect is present. Our focus in this section is on the leverage effect, which is measured by  $\gamma$ , implying asymmetry if we cannot support that  $\gamma$  equals zero. Specifically:

- $\gamma > 0$  : leverage effect present; negative shocks have a greater effect on volatility
- $\gamma < 0$  : inverse leverage effect present; positive shocks have a greater effect on volatility

The full parameter set  $\{\mu, \varphi_t, \theta_t, s_t, v, \delta, \omega, \alpha, \beta\}$  is estimated using quasi maximum likelihood estimation<sup>11</sup>. This follows, for example, Giot and Laurent (2003), who use non-Gaussian APARCH models to estimate Value-at-Risk in commodity markets. The parameter estimates and t-ratios for the APARCH(1,1) specification are presented in Tables 6 and 7 for our financial and electricity returns respectively.

---

<sup>10</sup> This is included due to the stylized fact of our returns that  $\rho(|r_t|, |r_{t-k}|) > \rho(r_t^2, r_{t-k}^2) \forall k > 1$ . This property, termed the Taylor effect, underscores that financial return series are not i.i.d., see Taylor (1986) and Malmstem and Terasvirta (2010).

<sup>11</sup> Estimation was done using the Rugarch Comprehensive R Archive Network (CRAN) package in the R programming language, see Ghalanos (2017)

The values of the log-likelihood statistic, as well as the significances of the parameters, imply that the nine parameter APARCH(1,1) model generally fits the electricity returns better than the financial returns with the EUR/USD exchange rate having the worst fit and the RT Peak return series having the best fit. This result for the exchange rate can be motivated by the view that volatility in these markets is usually intraday when markets react to monetary policy developments, as opposed being spread over multiple days, making the incorporation of the power transform unnecessary as small shocks aren't likely to be followed by larger ones in the following days. These findings also demonstrate that the unique features present in electricity returns, specifically asymmetry and large price shocks, can be incorporated into the volatility modelling scheme.

The first three parameter estimates reported are for  $\mu, \varphi, \theta$ , which form the ARMA(1,1) conditional mean equation. All series report an estimate for  $\mu$  close to zero, which suggests that there are no returns at the one-day horizon. The own mean spillover from lagged returns is captured by the AR(1) term  $\varphi$ , which is estimated to be positive for the SP500 and EUR-USD series, and negative for the Natural Gas and WTI Crude series. The results also indicate that leverage effects are present in the SP500, EUR/USD, WTI Crude, and Off-Peak DA series. More interestingly, inverse leverage effects seem to be present for the RT and DA Peak hour markets<sup>12</sup>. This supports our supposition that the convexity of supply can cause a positive relationship between returns and volatility.

---

<sup>12</sup> This can be demonstrated visually using the news impact curve (NIC) of Engle and NG (1993). This method uses a partially non-parametric estimate of the volatility model by first creating break points in the  $\hat{\varepsilon}_t$  series where its sign changes. This allows parameters to be estimated with a piecewise linear spline, enabling allows us to visualize the degree to which the impact on volatility from negative shocks differs from positive shocks. The NICs for financial and electricity APARCH models are shown in Figures 4 and 5 respectively.

## 7. Diagnostic Tests

Unit-root analysis is used in this study to measure market efficiency in addition to establishing the necessary stochastic conditions needed for econometric modelling of price time series. Market efficiency is defined here by the unpredictability of spot and forward prices, and the possibility of discernable changes in their cointegrating relationship over time. This assumes production decisions of market participants are endogenous to the positions they take in the DA markets, and assuming they all utilize the same information set for demand expectations, prices in the RT should reflect the random nature of demand shocks. As DA prices reflect the marginal costs of the most efficient market participant, and given that they are heterogeneous in their maximum production capacity, the ability to exercise market power (an attribute of market efficiency) would lead to a predictability in DA prices. These concepts can be more formally examined by validating the no-arbitrage principle in derivative pricing models, an approach commonly taken in the literature on electricity price efficiency. This approach measures market efficiency (and the randomness of prices that should result) through the well-defined framework of testing for stationarity, in addition to calibrating time series models to price derivatives.

Examples of influential model validation studies for pricing electricity contingent claims include Lucia and Schwartz (2002), who apply one and two factor geometric Brownian motions, and Benth and Schmeck (2014), who use a non-Gaussian variant of the Schwartz and Smith (2000) reduced form model commonly used for longer term commodity futures. This body of work shows that the assumptions about mean reversion, seasonality, spikes, and persistence play a significant role in the (mis)pricing of risk in continuous time. In this paper, discrete time equivalents of diffusion models are estimated, which require that means, variances, and autocovariances be invariant and finite with respect to time.

The economic rationale of using unit-root testing for measuring market efficiency posits that with an increasing number of market participants, who can optimally choose to buy and sell in both the RT and DA markets, price series will

tend to become integrated (non-stationary) over time, even if participants are unable to arbitrage between the two markets. Benth and Schmeck (2014) argue that so long as the future contract is tradable and liquid, then producers can both buy and sell in the DA, and therefore are not constrained by the physical inability to store electricity because they can create a replicating strategy to arbitrage using only forward contracts. If this hypothesis is incorrect, then non-stationarity should be observable throughout the sample. More formally, stationarity is defined by the invariance of moments such that for generic series  $\{Y_t\}$  with realizations  $y_t$ :

$$E(y_t) = \mu \forall t,$$

$$E((y_t - \mu)^2) = \sigma^2 \forall t,$$

$$E((y_t - \mu)(y_{t-s} - \mu)) = \gamma(s) \forall t, s$$

In addition to assessing price efficiency, these conditions allow for conventional statistical inference on parametric time series models. At a cursory level these conditions do not seem realistic when considering the underlying market drivers that cause cyclical and erratic behavior in electricity prices in the short run. However, they may be plausible given that, in the long run, prices do tend to revert to the mean as they are a function of the marginal cost of production (see, e.g., Pindyck, 1999). To explore these features, we implement the Said and Dickey (1984) (referred to as the Augmented Dickey Fuller or ADF test), the Phillips and Perron (1988) (or PP test), and Kwiatkowski, Phillips, Schmidt, and Shin (1992) (or KPSS test). We focus on ascertaining whether stationarity seems to be present in our price series. This is interesting from an efficiency point of view, but also because the outcomes might emphasize the stylized facts that make electricity market data particularly challenging to model. In addition, as Fezzi (2005) has highlighted, the general question of stationarity in electricity prices remains unresolved in the literature.

Tests with the null hypothesis of a unit-root, or I(1) behavior, are more prevalent in the literature than tests for stationarity which assume stationarity as the null hypothesis, or I(0) behavior. For example, Avsar and Goss (1999, 2001)

examine the informational efficiencies of futures contracts traded on the Sydney Futures Exchange and the New York Mercantile Exchange by applying the ADF and PP tests to five markets: Victoria, New South Wales, California-Oregon Border, Palo Verde, and the Cinergy system, which is part of the PJM. This studies used four main variables: the cash settlement price (the monthly average of spot prices), the closing futures prices of contracts one month from maturity, the number of contracts traded to account for market volume, and differences between the spot and forward prices. Using monthly observations over a short span (less than two years), with sample sizes of less than 20, both the ADF and PP tests produced consistent outcomes<sup>13</sup>: that series in these markets reject the null of I(1) in favor of I(0) with the exception of the Cinergy system in which series were better represented as I(1).

The DA and RT price variables we consider more closely resemble those used in the empirical studies of Atkins and Chen (2002) and Knittel and Roberts (2005). For instance, Atkins and Chen (2002) provide results of the ADF, PP, and KPSS tests, which reject their respective null hypotheses, both with and without trend terms. The test outcomes provide conflicting evidence for whether the Albertan market data are better modelled as I(1) or I(0). The authors posit that perhaps it might be preferable to test for fractional integration when using hourly prices, as the series might be better depicted with long-memory.

The ADF test, the most commonly examined unit root test in the electricity time series literature, assumes I(1) behavior under the null against the alternative hypothesis of stationarity, or I(0). Given a series of raw prices  $\{p_t\}$ , it considers the auxiliary regression:

$$\Delta p_t = \beta' \mathbf{D}_t + (\phi - 1)p_{t-1} + \sum_{k=1}^k \psi \Delta p_{t-k} + \epsilon_t,$$

where  $\phi = 1$  under the null hypothesis and  $\mathbf{D}_t$  is a vector of deterministic terms which can include a time trend and/or a constant term. In our application, we select

---

<sup>13</sup> In addition, their results suggest that the efficient markets hypothesis (EMH) cannot be rejected for a non-storable commodity like electricity using the forecast error approach on the basis risk (the difference between future and expected spot prices).

a model with a constant, or drift, term because prices are expected to decrease over time due to increased competition and efficiency gains. A time trend variable is not included as it would indicate they decrease quadratically and substantially, which is unlikely given the short one year sub samples to which we apply the test to. Adopting this model also allows us to compare our results with those of Arciniges et al. (2003), who apply the ADF test to prices that are 8 years older than ours. In this set up, the error term is assumed to be approximately white noise so  $k$  price lags are included to augment the regression in order to account for potential autocorrelation in prices. The choice of the lag order  $k$  can be based on information criterion using the log-likelihood function of the regression, sequential significance testing of residuals, or some function of the sample size  $T$ . We use  $k=15$  lags, as suggested by Arciniges et al. (2003) based on a general to specific based sequential rule that examines the significance of autocorrelated errors with the LB test and then dropping lags that are insignificant. This choice for  $k$  is also supported by the rule of thumb suggested by Schwert (1989) who recommends

$$k = \text{Int}\{12 \times (\frac{T}{100})^{\frac{1}{4}}\}$$

which would equal 16 if  $T=365$ . The unit root test statistic is given by:

$$ADF_t = t_{\phi=1} = \frac{\hat{\phi} - 1}{s.e.(\hat{\phi})}$$

Although this statistic resembles the standard t-ratio, it follows the Dickey-Fuller distribution asymptotically, not the standard Student's t distribution<sup>14</sup>. As this test

---

<sup>14</sup> This is due to the need to attain a convergent distribution for  $(\hat{\phi} - \phi)$  as  $T \rightarrow \infty$ . In order to achieve this, it must be scaled by  $T$ , not  $\sqrt{T}$ , which also makes it superconsistent since this is a faster rate of convergence. To put this in economic terms, remembering that the assumed random walk behavior implies that the evolution of the prices over time is truly random, the price observed at any point is equal to the sum of random errors in the past such that  $p_t = \sum_0^t \epsilon_t$ . This would imply that the asymptotic variance of  $p_t$  equals  $\sigma^2 t$ , which is not convergent as  $t$  grows, therefore it must be scaled differently, a property that also applies to the OLS estimates of parameters in the regression. Since the limiting distributions of unit-root and stationarity tests are non-standard (i.e., not Gaussian, Student's t, or F) but rather can be represented using Brownian motions  $W(r)$  following, for instance, Phillips (1987) and Chan and Wie (1988). This approach relies on the Functional Central Limit theorem, of which the Central Limit is a special case. The time series is viewed as a continuous process that is observed at integer values of time, but is still defined at non-integer values of  $t$ . Graphically this can be interpreted by dividing each observation by  $\sqrt{T}$ , and then compressing the

assume the series is a random-walk under the null hypothesis, the corresponding hypothesis of the regression is:

$$H_0: \phi = 1 \Rightarrow p_t \sim I(1), \text{ against } H_1: |\phi| < 1 \Rightarrow p_t \sim I(0).$$

The critical values for the limiting distribution are derived using Monte Carlo techniques in Dickey and Fuller (1981); see also Hamilton (1994).

The PP test allows for autocorrelation in  $u_t$  making it robust in the presence of autocorrelated errors. It uses the Newey and West (1986) estimate of the long run variance<sup>15</sup> to non-parametrically modify the ADF test statistic. The regression equation used in the PP test is the same as the ADF but without the augmentation terms:

$$\Delta p_t = \beta' \mathbf{D}_t + \pi p_{t-1} + u_t.$$

The limiting null distribution of the PP test statistic is the same as that for the ADF statistic. This approach does not require choosing the number of augmentation terms (that account for autocorrelation in the error term parametrically) but allows for autocorrelation non-parametrically, which requires selecting the bandwidth for a lag truncation parameter used for computing the long run variance. The Bartlett window  $\omega_j = 1 - \frac{j}{l_q+1}$  is used in the Newey West estimator of the long-run variance (see footnote 13), which requires setting  $l_q$ . We used the Schwert (1989) rule of thumb in setting  $l_q = 16$ .

The main criticism of the ADF and PP tests is that they can falsely accept the null of I(1) when the AR parameter is close to 1, but not equal to 1. Due to this low power, particularly for near unit-root processes, we also implement the KPSS

---

entire series horizontally so that it fits into a unit interval  $r \in [0,1]$  (see, e.g., Hayashi 2000). The limiting distribution of the ADF is then :  $t_{\phi=1} \xrightarrow{d} \int_0^1 W_t dW_t / \left( \int_0^1 W_t^2 dW_t \right)^{\frac{1}{2}}$  using the continuous time sample moments from Phillips (1987).

<sup>15</sup> This test uses  $\hat{\sigma}^2$  and  $\hat{\lambda}^2$ , which are both consistent long run estimates of the variance parameters following Newey and West (1986):  $\hat{\sigma}^2 = \lim_{T \rightarrow \infty} T^{-1} \sum_{t=1}^T [E[u_t^2]]$  and  $\hat{\lambda}^2 = \lim_{T \rightarrow \infty} \sum_{t=1}^T [E[T^{-1}(\sum_{t=1}^T u_t)]]$

test, a stationarity test (instead of a unit-root test) that assumes the series is  $I(0)$  under the null hypothesis. We examine the auxiliary regression:

$$p_t = \beta' \mathbf{D}_t + \mu_t + u_t,$$

$$\mu_t = \mu_{t-1} + \varepsilon_t,$$

where  $u_t$  is assumed to be  $I(0)$ , and  $\mu_t$  is a random walk with  $\varepsilon_t$  being assumed to be white noise with variance  $\sigma_\varepsilon^2$ . Again we include only a constant term in the deterministic vector  $\mathbf{D}_t$ . The corresponding hypothesis test can be summarized by:

$H_0: \sigma_\varepsilon^2 = 0 \Rightarrow p_t \sim I(0)$ , against  $H_1: \sigma_\varepsilon^2 > 0 \Rightarrow p_t \sim I(1)$  which is one sided and right tailed. A Lagrange Multiplier test statistic<sup>16</sup> is given by:

$$KPSS = \left( T^{-2} \sum_{t=1}^T \left( \sum_{j=1}^t \hat{u}_j \right)^2 \right) / \hat{\lambda}^2$$

where  $\hat{u}$  is the residual of a regression of  $p_t$  on  $\mathbf{D}_t$ . Here again the Newey-West (1984) consistent estimator of the long run variance of  $\hat{u}_t$  is used with the same Bartlett window lag truncation parameter  $l_q$  as for the PP test.

We apply these tests to yearly sub-samples, which break the data down in line with Arciniges et al. (2003). In this set up, the daily time series for each hour from 2008 to 2015 is tested individually, in addition to testing the full samples. We also do this for a constructed series that is the average of all hours, and applied to both the RT and DA markets. In total, each of the ADF, PP, and KPSS tests are applied to 450 different hour-year combinations. Results of the ADF tests, given in Table 8, indicate that, with the exception of 2009 and 2013, there is a predominant failure to reject of the null hypothesis of a unit root. When examining the time series plots for 2009 and 2013, they had the lowest levels of volatility during the summer months when it is generally high across all hours and markets, such features support the finding of stationarity  $I(0)$  for most hours in these year. The results from the PP

---

<sup>16</sup> The equivalent continuous time limiting null distribution of the test where  $\mathbf{D}_t$  only contains a constant is given by:  $KPSS \xrightarrow{d} \int_0^1 W(r) - rW(1)dr$ .

tests, provided in Table 9, are considerably different, where rejection of the unit root null in favor of stationarity is common across all of the hour-year combination. One possible reason, for this different outcome across tests, is that significant negative autocorrelations exist in the first differences of prices (see Figure 6), an issue that Schwert (1989) has shown to cause severe size distortions for the PP test using Monte Carlo simulations.

The KPSS results in Table 10 also differ somewhat from the outcomes from the ADF and PP tests. Peak hours are more likely to be stationary than off peak hours, a counterintuitive result, as peak hours tend to exhibit greater variance than off peak hours. Most hour-year combinations are found to be  $I(0)$  using the KPSS test, with the exception of the 2014 year. Examining the time series plot for this year shows that it had extreme outliers, with prices occasionally exceeding \$1000/MWh. Perhaps the outcomes for this year are consistent with the results reported by Otero and Smith (2005), who found that the power of the KPSS falls significantly when the underlying series is  $I(1)$  but also has outliers (assuming the series are  $I(1)$  in actuality).

When applied to the whole sample, the ADF and PP tests conclude stationarity for all hours in both the RT and DA, whereas the KPSS finds that almost all hours are non-stationary. We repeat this exercise after taking the first difference of prices:  $(1 - L)P_t$ . All 1350 tests, reported in Tables 11, 12, and 13, find in favor of  $I(1)$  as opposed to  $I(2)$ , suggesting that the price series are at most  $I(1)$ . In summary, the unit root and stationarity test results are largely inconclusive for our study, findings that are in line with the literature that apply tests with an  $I(0)$  null as well as  $I(1)$  null; e.g., see Aitkins and Chen (2005), Haldrup and Nielsen (2006). To further explore this dilemma, we turn to, for example, Ballie (1996, p. 6) who notes that the binary "knife-edge distinction" between  $I(0)$  and  $I(1)$  is limiting, providing motivation to examine whether the series are better modelled as fractionally integrated, with the order of integration between 0 and 1.

## 8. Fractional Integration

The conflicting results of the ADF, PP, and KPSS tests support the findings of, for example, Avsar and Goss (2002), who note that electricity price series may be better represented as fractionally integrated (FI) processes. This behavior, also called long range dependence, implies that daily prices are neither  $I(0)$  nor  $I(1)$ , but better represented with orders of integration that lie between 0 and 1; i.e.,  $I(d)$  with  $0 < d < 1$ . A non-integer value for  $d$  means that price realizations observed far apart from each other remain strongly correlated. Dependence between observations taken far apart in time also implies that the i.i.d. assumption does not hold, violating usual central limit theorems, with the consequence being that the variance of the mean may not decrease at a rate of  $1/T$ . Should such representations be more appropriate, then significant implications result for modelling, including the use of small samples, or impulse response functions, which measure the impact of shocks to the data generating process. Unlike autoregressive moving average (ARMA) processes whose autocorrelations decay exponentially, FI processes have autocorrelations that decay hyperbolically, which means that if  $\rho(k)$  is the autocorrelation for the  $k^{\text{th}}$  lag and  $k \rightarrow \infty$ , then  $\rho(k) \rightarrow C_\rho k^{-\alpha}$  where  $C_\rho$  is some constant and  $0 < \alpha < 1$ . If this rate is slow enough, then  $\rho(k)$  may not be summable, a condition that forms the McLeod and Hipel (1978) definition of long memory, which states if

$$\lim_{T \rightarrow \infty} \sum_{k=-T}^T \rho(k) = \infty$$

holds, then the process is fractionally integrated. Avsar and Goss (2002) and Knittel and Robert (2005) both find significant correlations for at least 1000 lags using hourly series, which implies a memory of 40 days. Our daily average series show that the length of memory may be far greater, reaching at least 250 days before becoming insignificant (see Figure 7). It's important to note that the slow rate of decay in autocorrelations dictates whether the series exhibits long memory, not the length of the number of significant lags itself. If the series exhibits this property, then it might be better represented as a autoregressive fractionally

integrated moving average. This ARFIMA( $p, d, q$ ) model with  $p$  AR lags, order of integration  $d$ , and  $q$  MA lags takes the form:

$$\Phi(L)(1 - L)^d(y_t - \mu) = \Theta(L)\varepsilon_t$$

with the fractional differencing operator  $(1 - L)$  defined by:

$$\sum_{k=0}^d \binom{d}{k} (-1)^k L^{d-k}$$

and the binomial coefficients  $\binom{d}{k}$  equal to:

$$\frac{\Gamma(d + 1)}{\Gamma(k + 1)\Gamma(d - k + 1)}$$

with  $\Gamma(\cdot)$  as the gamma, or generalized fraction, function. If  $d$  takes on integer values, the structural form would simplify to an ARIMA or ARMA process.

We explore such fractional integration ideas in this section.

In order to motivate the test statistics we consider, it is important to distinguish between the time and frequency domains. The tests examined in the previous section relied on the time domain representation of the series, which states that a covariance stationarity process can be decomposed into a linear combination of lags of white noise, and that future values can be predicted by a linear function of past observations. This Wold decomposition takes the form:  $Y_t = \mu_t + \sum_{j=0}^{\infty} \theta_j \varepsilon_{t-j}$  and  $\varepsilon_t \equiv [Y_t - P(Y_t|Y_{t-1}, Y_{t-2}, \dots)] \sim WN(0, \sigma^2)$  where WN is white noise. The tests for long memory used here operate in the frequency domain instead, which interprets a univariate series as the sum of a combination of sinusoids that have random and uncorrelated coefficients (see, e.g., Hamilton, 1994). To attain this representation, the z-transform is taken of the Wold decomposition, where  $z = e^{-i\omega}$  and  $\omega = [0, 2\pi]$ , to obtain

$$y(z) = \sum_{\tau=-\infty}^{\infty} y_{\tau} z^{\tau}.$$

These two domains can be linked using the Wiener-Khintchine<sup>17</sup> theorem, which states that the spectral density of a stationary process can be expressed as the Fourier transform of the ACF. This also means that both the ACF and the spectral density function contain the same information (e.g., see Zivot, 2006). The corresponding spectral density of a stationary process is then:

$$f(\omega) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} \rho(k) e^{i\omega k},$$

which converges to  $C_f \omega^{\alpha-1}$  as  $\omega \rightarrow 0$ , commonly called the zero frequency. The  $H$  coefficient from Hurst (1951) is typically used as the measure of long memory, instead of  $\alpha$ , in this spectral density based test where

$$H = 1 - \frac{\alpha}{2}, \alpha \in \left(\frac{1}{2}, 1\right).$$

For a fractionally integrated process  $y_t$ , the differencing order  $d$  is equal to  $H-1/2$ , and as  $H$  increases, the longer the memory of the process if it is stationary.

We apply two methods for detecting long memory in this section, the first being the rescaled range or R/S test, initially proposed by Hurst (1951). The test statistic is defined as a range of partial sums of deviations from the mean, which are scaled by the standard deviation. In Figure 8, we plot the ACFs of our four representative series, which show that autocorrelations decay slowly, suggesting that there is possibly long memory. To account for this, we examine the Lo (1991) version of the R/S test, which makes the R/S test robust in the presence of short term memory by scaling the partial sums by the square root of the Newey and West (1986) estimate of the long run variance:  $\hat{\sigma}_T$  (defined in footnote 14). The test for long memory has a null hypothesis of no long range (LR) dependence and the test statistic<sup>18</sup> is defined by:

---

<sup>17</sup> See Khintchine (1934)

<sup>18</sup> Its limiting distribution is the range of a Brownian bridge on the unit interval:  $W(r), r \in [0,1]$ , see Lo (1991)

$$\tilde{Q}_T = \frac{1}{\hat{\sigma}_T} \left[ \max_{1 \leq k \leq T} \sum_{j=1}^k (p_j - \bar{p}) - \min_{1 \leq k \leq T} \sum_{j=1}^k (p_j - \bar{p}) \right]$$

$H_0: d \notin (0, \frac{1}{2}) \Rightarrow$  No LR Dependence in  $p_t$ , against  $H_1: d \in (0, \frac{1}{2}) \Rightarrow$  LR Dependence in  $p_t$

Another method to determine whether a series has long memory is to directly estimate  $d$ . For example, the approach of Geweke and Porter-Hudak (1983) (referred to hereafter by GPH) is to estimate  $d$  non-parametrically without having to specify the short term memory process (i.e., the  $p$  and  $q$  parameter in the ARFIMA( $p, d, q$ ) representation). The order of integration  $d$  is computed by estimating the regression:

$$\ln(I_y(\omega_j)) = \phi_0 + \phi_1 \ln\left(4 \sin\left(\frac{\omega_j}{2}\right)\right) + \epsilon_t, \quad j = 1, 2, \dots, T$$

over  $j$  frequencies. The estimate of  $d$  is then:

$$\hat{d} = \frac{\frac{1}{2} \sum_{j=1}^T [p_t \ln(I_y(\omega_j))]}{\sum_{j=1}^T [p_t^2]}$$

where  $I_y(\omega_j)$  is the spectral density of the series, i.e., the distribution of frequencies that the series can be broken down into  $\omega_j$ . The following conditions from Hosking (1981) can be used to interpret the GPH estimates for  $d$ :

$|d| > 1/2$ :  $p_t$  is non-stationarity

$0 < d < 1/2$ :  $p_t$  is stationary and has long memory

$-1/2 < d < 0$ :  $p_t$  is stationary and has short memory.

Examples in the literature of the modified R/S and the GPH methods applied to electricity prices include Atkins and Chen (2002) and Weron (2002). Atkins and Chen (2002) utilize hourly price series spanning 45 months, and find statistically significant evidence in favor of using fractionally integrated long memory processes to determine a series' order of integration. They obtain inconclusive unit root and stationarity test results, finding in favor of long range dependence using the modified R/S test, and report estimates for  $d$  using the GPH method of between .35 and .45 (for  $\omega_j$  ranging between .5 and .8). Weron (2002) applies the R/S and

GPH methods to average daily prices in the CALPX, Entergy, Scandinavian Nord Pool, and U.K. markets. He finds that the CALPX and Entergy exhibit short memory with values for  $d$  of -.43 and -.48 respectively<sup>19</sup>, the Nord Pool can be best modeled as a random walk, and the 265 day span of the U.K. series is too small to draw a conclusion. More interestingly, his results find that test statistics did not change significantly if weekly and annual cycles were removed from the series, indicating that the order of integration is not dependent on seasonality.

In our application we generate test statistics for Lo's modified R/S and provide estimates for  $d$  using the GPH method - these are reported in Tables 16 and 17 respectively. For the modified R/S test, 42% of the 450 hour-year combinations rejected the null hypotheses of no long range dependence in favor of long memory, however all of the full sample series all failed to reject. These are not entirely consistent with the GPH estimates for  $d$  presented in Table 17, however, this might be due to the fact that both of these methods assume the series is  $I(0)$  to begin with. The GPH estimates indicate that 1% of the hour-year combinations display short-term memory, 49% display long-term memory, and the rest are non-stationary. The presence of long-term memory in much of the samples highlights an important driver of the underlying data generating process that is persistent in nature: seasonality. This feature seems relevant to long memory, despite the findings of Weron (2002), because one definition of long range dependence offered here is that observations taken far apart in time are strongly correlated. Weather is a fundamental driver of electricity demand, and since temperature is cyclical in the long run, it is not surprising that long memory is present in many of our sub samples. We explore the cyclical nature of seasonality in the next section.

## 9. Seasonality

Energy markets in general, and electricity especially, are highly weather dependent. For instance, the use of exchange traded temperature derivatives to manage risk is

---

<sup>19</sup> These figures are converted to  $d$  as the reported figures in the paper were presented in terms of the Hurst exponent  $H$ .

an example of the influence of weather in these markets. This implies that seasonality is an important stylized fact to explore in our series. We begin by first using a simple regression to highlight the time of year effects that are present in prices. This seasonality can be shown by regressing daily price series on a time trend and dummy variables for the month that the price was observed in. We estimate the following equation for our representative prices in addition to Peak and Off-Peak load:

$$p_t = \mu_i' \mathbf{D}_{it} + \beta t + \varepsilon_t$$

where the dummy variable vector  $\mathbf{D}_{it}$  contains dummy variables for each month  $i$  at time  $t$  associated with the coefficient vector  $\mu_i$  for the corresponding month of the year; a time trend  $t$  is also included. This is equivalent to taking monthly averages of daily prices while adjusting for the fact that they decrease in the long run at a rate of  $\beta$ . The results, shown in Table 18, suggest that the DA, RT, Peak and Off-peak categorizations have distinct cyclical patterns. Unsurprisingly, prices tend to be significantly higher during peak hours and exhibit one cycle per year, being highest during the summer and lowest during the winter, with transition between these highs and lows during the spring and fall. More interestingly, off-peak prices exhibit two cycles per year in both the DA and RT. The prices are highest during the winter and summer, and lowest during spring and fall; e.g., see the load variable. However, using 12 dummy variables to describe the cyclical nature of prices is an arbitrary choice, because dummy variables implicitly place a structure on the form of seasonality and most likely lack the granularity to accurately model the various cycles in electricity prices and volumes. Despite this, the use of dummy variables to account for seasonality is one of the three main structural forms in the literature examining the long-term seasonality component (LTSC) of electricity time series.

Perhaps the other two considered approaches are better able to accommodate the intricate cycles displayed in these series: sinusoidal Fourier based methods, and maximum overlap discrete wavelet transform (MODWT), which is a type of smoothing. Both of these methods effectively interpret a time series as observations of a continuous signal, projected onto a discrete basis space (e.g., see Ramsey,

2002). Such approaches to identifying seasonality are useful in electricity time series applications because they do not impose the likely unrealistic assumption of mean changes in discrete shifts.

The same spectral decomposition used in constructing the GPH statistic can also be applied to modelling seasonality when dealing with highly cyclical time series with potentially hidden periodicities. In this section, trigonometric functions are used to represent the time series in the frequency domain in order to model the cycles exhibited in both prices and volume. With this framework, the simplest form is to represent the series by  $Y_t = R\cos(\omega t + \Phi) \epsilon_t$ , where  $R$  is the amplitude,  $f$  is the frequency, and  $\Phi$  is the phase. These parameters respectively describe the 1) magnitude of each frequency's effect on  $Y_t$ , 2) how many cycles occur per unit of time, and 3) that each cycle need not begin at the same point in time. Since the first and third parameters do not enter the equation linearly, trigonometric identities are used to attain an estimable form. In addition, there can exist many different frequencies, each motivated by patterns in the underlying market drivers. For example, weekly seasonality can arise due to a decrease in demand during weekends, or annual seasonality due to changes in temperature.

In order to incorporate these separate effects, multiple frequencies are included. The equation to be estimated for a series with a non-zero mean, and  $m$  underlying frequencies is:

$$Y_t = \eta + \sum_{j=1}^m [A_j \cos(2\pi\omega_j t) + B_j \sin(2\pi\omega_j t)] + \epsilon_t$$

where all of the frequencies  $\omega_j$  are such that  $0 < \omega_1 < \omega_2 < \dots < \omega_m < \frac{1}{2}$ , a condition which implies that the lowest frequency must reach at least one turning point in the sample in order to be visible. It is important to note that the validity of this representation (which is estimated via OLS) is dependent on the assumption that  $Y_t$  is weakly stationary, which, as we have shown, is an open question. The spectral decomposition of the series can be used to select the frequencies in the model by taking the Fourier transform of the series. However, we specify them based on the underlying seasonality drivers in the market. We use  $\omega_j = \{\frac{1}{365}, \frac{1}{180}, \frac{1}{7}\}$ , which represent annual, semi-annual, and weekly cycles. The semi-annual

frequency is motivated by heating and cooling related demand on the system, which increases further the ambient temperature away from room temperature. This is also supported by the seasonality regressions reported in Table 8, as prices increase in both the summer and winter months. Weekly seasonality is motivated by the fact that weekends have different demand profiles than weekdays. The results of the Fourier based seasonality regressions are presented in Table 19. These indicate that the time time, and weakly, semi-annual, and annual explanatory variables are statistically significant. Additionally, the model is more appropriate for the load variables than prices as demonstrated by higher adjusted  $R^2$  values. This is supported by the highly regular nature of load variation.

The plots of the estimated regressions, presented in Figures 10, 11, and 12, highlight an important drawback of this seasonality model as it assumes cycle frequencies and their magnitudes are homogenous throughout the series. For example, the peaks in 2009's prices (see Figure 10) are overestimated by the model, as actual prices were lower than for other years. In Figure 11, the peaks and troughs estimated by the model do not align with actual prices. The model might be more appropriate for estimating loads, as their cyclical nature is largely homogenous - Figure 12 shows that both the frequency and magnitude of cycles estimated by the model are closely aligned with observed loads.

In conclusion, our results suggest that the seasonality present in electricity prices and volumes is not adequately captured by discrete dummy variables as such terms assume that the series increase and decrease in a cyclical and continuous fashion, which is not a feature of the seasonality in our electricity variables. Additionally, so long as the underlying drivers of seasonality can be established (e.g., time of day effects, day of week effects, seasonal temperatures), then a corresponding trigonometric based regression can be applied to capture these cycles. However, a drawback to this method is that it assumes that the frequency of these cycles do not change, which does not seem to be the case. Given this, we find that trigonometric based seasonality functions are more applicable to modelling loads than prices, since cycles in volume are more stable over time. The plots of these seasonality functions against actual Peak and Off-Peak prices, reported in

Figures 10 and 11 respectively, also show that cycles in prices are well represented as sinusoidal. But, their apexes, which occur during the summer months, have significant spikes, resulting in a magnitude that is greater than the troughs. This indicates that prices behave distinctly in periods of high demand; i.e., that perhaps the market exists in multiple states, or regimes. Such characteristics can be formalized in an econometric framework by employing, for instance, regime switching models. We explore the applicability of such models in the next section.

## **10. Markov Regime Switching**

Using Markov regime switching (MRS) models as an approach to reconcile the departures from linearity of electricity prices is established on the premise that the market can exist in multiple states, depending on whether it is in equilibrium or not. This is fundamentally motivated by the erratic price behavior caused by the temporary inability of the market to deliver a sufficient supply. Deng, Johnson, and Sogomonian (2001) propose that such features can be tested in electricity markets by comparing the parameters of mean reverting jump diffusion models. These models posit that prices revert to a constant in the long-run but occasionally spike, or jump in the short run. In their study, price spikes are defined as large upward swings followed by large downward swings (as opposed to upward swings that remain high). This interpretation highlights the usefulness of MRS models in electricity markets in their ability to estimate the probability of transitioning into a state of erratic prices, which may be able to assist in assessing the economic value of adding capacity and infrastructure investment into the market. Although Deng et al. (2001) did not undertake an empirical application, many subsequent studies have followed this basic paradigm.

Allowing for varying transition probabilities also seems important. Time varying transition probabilities with explanatory variables are applied to the PJM by Mount, Ning, and Cai (2006), using the reserve margin to predict which state the market is in for one year periods beginning in May 1999. In this study, the state

stochastic process is governed by a logistic function that imputes the probability of regime change. Here again stylized facts about the market are used to specify the switching mechanism. Since price spikes are theoretically motivated by an intersection of an inelastic demand curve with the kinked marginal cost curve, at price levels where expensive petroleum based generators come online, the likelihood of spikes can be derived from an indicator of how much the system is stressed.

In an effort to ensure that transmission is guaranteed in the PJM, the market schedules commitments from generators well into the future, referred to as total offered commitments (TOC), which provide a measure of how much quantity is available. Once load ( $L$ ) is realized, the ratio  $R = \frac{TOC}{L} - 1$  can be used as a measure of system stress. If  $R > 0$ , then the system has sufficient capacity to meet demand, if  $R < 0$ , then the likelihood of erratic price behavior increases. Using daily peak RT price series  $y_t$ , and average daily load  $L_t$ , Mount et al. estimate:  $y_{t,i} = \alpha_i + \phi_i y_{t-1,i} + \gamma_i R_{t,i} + \omega_i \log(L_{t,i}) + \varepsilon_{t,i}$  where the  $i$  subscript denotes the state variable with two regimes,  $i=1,2$ . Their results show that high price (spike) regimes have an estimated variance that is twice that of the low price (base) regime. In addition, spikes are shown to likely occur when the reserve margin dips below 20%. Indeed, when spikes are defined as  $p_t > \$100$ , 11 out of 13 spikes were correctly detected when the model produced a switching probability (from base to spike) greater than 0.5. With a few exceptions, the majority of MRS applications to electricity markets in the literature have treated regimes as latent variables instead of specifying thresholds that select, not assume, the particular regime each observation is in for estimation.

Ethier and Mount (1998) analyze the Australian and the U.S. markets in an empirical application of MRS models to electricity prices using non-parametric Wald tests to compare whether parameters are the same across AR(1) models in both one and two-state regimes. They conclude that models allowing for shifts in both the mean and the variance work best for their studied electricity markets. The AR(1) models are discrete time analogues of the continuous time framework,

following the Ornstein and Uhlenbeck (1930) setup, which highlights another use of the MRS framework of being able to validate options pricing models. These models utilize arbitrage pricing theory, which requires some predictability in the price process, something that is usually manifested in the drift term of an AR(1) model. If, for instance, demand for derivative instruments in this market is being driven by the need to hedge marginal costs of producers and generators, then options that are priced using the no-arbitrage assumption would undervalue the derivatives with high-strike prices. This is because they would underestimate the volatility, by failing to pick up the notion the market is in a completely different state. Such ideas are taken up by Jancaruza (2014), who applies a three regime model, with a base state, upward spike state, and downward price dip state, to price European style options (which have a fixed delivery date), using an additive seasonality function in the European Electricity Exchange (EEX) market. By calibrating the MRS model to observed prices, she finds that it performs well, as it was able to generate options prices close to those observed on the actual exchange.

Huisman and Maheu (2003) empirically examine the specific question of whether regime switching improved accuracy over a stochastic jump diffusion (SJD) process. The SJD class of models, an alternative to MRS models, allow for discrete shifts in prices from what is otherwise considered a continuous sample path. Although first motivated by the need to incorporate idiosyncratic shocks in equities markets, the application of SJD type models to electricity markets is supported by the analogous nature of demand and supply of loads with news incorporation in equities prices. A key contribution of the work of Huisman and Maheu (2003) is to show that such SJD models effectively entangle the jump parameter and the mean reversion parameter, resulting in biased estimators of both parameters. This misspecification is often illustrated by the positive estimate of the mean reversion parameter generated by an SJD model when applied to an electricity price time series. Our seasonality analysis, in addition to the stylized facts of electricity being a commodity, not a financial, asset, would imply that the mean

reversion parameter be negative, i.e., in the long run, prices actually revert back to normal.<sup>20</sup>

Markov regime switching models also allow for a granular analysis of the data generating process, while being reasonably tractable to estimate. When markets behave in distinct regimes under certain circumstances, those circumstances need not be well defined nor well understood to implement MRS models. Since parameter instability is a prominent issue with price models in electricity markets, MRS models provide an intuitive solution to account for the possibly time varying nature of parameters. The two main subclasses of MRS models are largely delineated by their treatment of the switching process, specifically, whether the switching mechanism can be characterized by an observable variable (such as a threshold), or driven by an unobservable (latent) stochastic process (e.g., Franses and van Dijk, 2000).

Although measuring out of sample forecast errors can be useful in establishing the adequacy of the model, it is important to note that our objective is not to predict prices, but to explain price formation. In the risk enterprise management context, understanding a model's risk and its limitations is paramount because the goal for participants in these markets is usually to hedge exposure. Power markets are an example of a traded commodity where speculators have limited price setting capability, a fact that is changing as virtual bidding policies are being implemented and more arbitrageurs are taking advantage of them (e.g., Saravia, 2003). These market participants need to be able to estimate the likelihood of a price spike in order to justify taking a short or long positions in either the spot or forward markets.

The MRS model we implement uses the Fourier based seasonality analysis we presented earlier. Our approach is to first deseasonalize the price series, in lieu of incorporating the seasonality function in to the MRS model, which could be

---

<sup>20</sup> This rationale is analogous to Pindyck (1999), who motivates long run energy price forecasting by the theory of non-renewable resource production put forward by Hotelling (1930).

undertaken by specifying that the seasonality parameters do not switch between regimes. The general mean model we implement takes the form:

$$\begin{aligned} p(t) &= g(\sim) + y_t \\ g(\sim) &= \text{deterministic seasonality function} \\ y_t &= \text{stochastic price process} \end{aligned}$$

In this framework,  $S_t$  is a vector of the corresponding states for each observation. The joint process of  $y_t$  and  $S_t$  delineates which of the two main subclasses of MRS models is applied. The market can be in one of two states at time  $t$ :  $S_t = 0$  which represents the base regime, or  $S_t = 1$  which represents the spike or disequilibrium regime. The full model is then:

$$y_{t,i} = x'_{t,i} \beta_{S_{t,i}} + \sigma_{S_{t,i}} \varepsilon_{t,i}; \varepsilon_{t,i} \sim i. i. d. N(0,1)$$

which can be decomposed into two distinct states  $i \in \{0,1\}$ :

$$\begin{aligned} y_{t,0} &= \alpha_0 + (1 - \beta_0)y_{t-1,0} + \sigma_0 \varepsilon_{t,0} \mid S_t = 0 \\ y_{t,1} &= \alpha_1 + (1 - \beta_1)y_{t-1,1} + \sigma_1 \varepsilon_{t,1} \mid S_t = 1. \end{aligned}$$

The assumptions governing the joint distribution of  $y_t$  and  $S_t$  also governs the optimization technique used to generate parameter estimates. Here the evolution of  $y_t$  is dependent on  $S_t$ , and  $y_{t-1}$ . The two state Markov chain  $S_t$  is not observed directly, but derived from  $y_t$ . In the two regime case,  $S_t$  is assumed to evolve as an ergodic Markov chain, with the probability of remaining in a base state being  $P(S_t = 0 \mid S_{t-1} = 0) = q_0$ , and remaining in the spike state:  $P(S_t = 1 \mid S_{t-1} = 1) = q_1$ . Collecting the transition probabilities into a matrix:

$$\begin{bmatrix} P(S_t = 0 \mid S_{t-1} = 0) & P(S_t = 1 \mid S_{t-1} = 0) \\ P(S_t = 0 \mid S_{t-1} = 1) & P(S_t = 1 \mid S_{t-1} = 1) \end{bmatrix} = \begin{bmatrix} q_0 & 1 - q_0 \\ 1 - q_1 & q_1 \end{bmatrix}$$

where  $1 - q_0$  is the probability of switching from a base state to a spike state. Estimation of such models is usually undertaken using an expectation maximization (EM) algorithm in lieu of conventional maximum likelihood (ML) methods. Initially developed to estimate equations in the presence of missing observations, EM has analogous applications to models with finite mixtures, grouped, truncated, or censored data (e.g., Dempster, Laird, and Rubin, 1977). The unobserved  $S_t$

process is recovered by using the joint distribution of  $y_t$  and  $S_t$  via the chain rule. By obtaining the marginal distribution of  $y_t$ , the likelihood function<sup>21</sup> used to estimate the full parameter set  $\theta: \{\alpha_0, \alpha_1, \beta_0, \beta_1, \sigma_0, \sigma_1, q_0, q_1, \}$  in a two-regime setup is provided by:

$$f(y_t|\Omega_{t-1}) = \sum_{S_t=1}^2 f(y_t|S_t, \Omega_{t-1})f(S_t|\Omega_{t-1})$$

$$\mathcal{L}(\theta) = \sum_{t=1}^T \ln(f(y_t|\Omega_{t-1}))$$

where  $\Omega_{t-1}$  reflects the history of the series  $\{y_{t-1}, y_{t-2}, \dots, y_1\}$  in the likelihood function as its estimation is iterative, see footnote 21.

We draw inferences by comparing parameter estimates and switching probabilities between regimes in Tables 20, 21, 22 and 23. All representative series have been decomposed into base and spike regimes, with residual standard error in the spike regime significantly greater than in the base regime in line with the relevant literature. This suggests that high prices can be characterized as spikes with higher volatility and weaker model fit. In addition, the estimated transition probability of remaining in a high price regime are less than that of remaining in the base regime for all series. Some series have a base regime with a constant term less than zero. This does not imply negative prices, which can occur when producers have excess capacity, as then it is cheaper to pay consumers to take delivery than shut down power plants. These negative

---

<sup>21</sup> The structural form of the conditional distribution, and its resultant likelihood function, is why EM is used instead of ML. Although iterative ML can find local maxima, ensuring these are in fact global maxima is difficult because the likelihood function becomes an implicit equation when differentiating. Although EM methods cannot guarantee estimates come from global and not local maxima per se, the iterative technique for optimization it deploys makes it more likely to find global maxima because the approach effectively guesses which starting values are best for finding global maxima by intuitively using the Bayes' (chain) rule in combination with Jensen's inequality. By using minima of the likelihood function (E-step), the subsequent coordinate ascent optimization (M-step) guarantees that the likelihood function increases monotonically, making it a technique that is more efficient, usually, than ML. For instance, see Hamilton (1994), and Dempster et al. 1977).

estimates for prices result from the mean price being less than the seasonal average. Surprisingly, MRS models perform worse than their non-switching AR(1) counterparts using, for instance, the Akaike information criterion (AIC), despite being strongly motivated theoretically as a more appropriate model choice given price spikes during high demand.

## **11. Conclusion**

In this study we have firstly presented a survey of empirical electricity time series modelling literature, and then secondly, undertaken a thorough exploration of the features of electricity series in the PJM market compared with some representative financial series. Clearly, physical characteristics of the electricity, and its market's structure influence price formation in ways that are distinct from other financial markets. For instance, we have found that short term bursts of volatility make unit root analysis difficult to implement. This is problematic as simple ARMA representations of prices often serve as the workhorse for both forecasting as well as risk management applications that require capturing historical trends. Our results indicate that in the full sample, the Off Peak hours of 5AM to 7AM in the RT market are the only ones where prices are conclusively found to be stationary in all three tests implemented. These hours are also when the fastest increase in demand occurs during the day. This is counterintuitive, as prices would be expected to be erratic during hours of the day when demand ramps up quickly, i.e., non-stationarity.

One possible explanation is that electricity retailers may be risk averse during these hours given the challenging demand dynamics, and as a result, prefer to buy in the DA market, so as not to take on the risk in the RT market should demand overshoot supply. This would mean more load is scheduled a day-ahead than would if they were risk-neutral. This is supported by the fact that mean prices in the RT Off Peak are lower than the DA Off Peak by roughly 2.1%, i.e., there is a premium for purchasing electricity a day-ahead during hours when electricity demand is low in nominal terms, but increases rapidly.

Physical assets used in electricity production can experience various infrastructure constraints that can restrict supply. If, however, the market has become more efficient and integrated over time then it is possible that it can be better to manage negative supply shocks if deregulation has reduced barriers to entry and increased the diversity of the resource mix. This would imply that smaller market participants could fill supply shortages attributed to larger participants. In other words, a positive price shock would be less severe, and last for a shorter period of time, so that it is plausible that the memory of the price series would decrease over time. Pereira et al. (2016) suggest that as electricity markets become more integrated, the ability to better manage supply shocks should result in a decreasing persistence over time in prices, as the availability of supply pacifies price spikes. Our results for the estimate of the order of integration supports this hypothesis, as the order generally decrease over time, especially after 2012 when the PJM incorporated parts of Ohio and Kentucky into the system.

In addition to supply dynamics, long memory is also supported by seasonal effects, since temperature is predictably cyclical, correlated across long spans of time, and strongly influences energy demand. The Fourier based seasonality functions we implement are useful as they can capture cycles over multiple frequencies. Lastly, we provided an application of MRS models with Fourier based seasonality functions. Although appealing from a theoretical perspective, we find that the increased complexity of the model does not necessarily translate to better performance over simpler non-switching alternatives. However, in the case of volatility modelling we find that the distinct price spike features of electricity returns can be incorporated into the modelling framework to better capture volatility than simpler alternatives. These findings highlight the importance of establishing the features of the time series before selecting an appropriate model, and motivating it with economic rationale.

## 12. References

- Arciniegas, I., Barrett, C., & Marathe, A. (2003), Assessing the Efficiency of US Electricity Markets, *Utilities Policy*, 11(2), 75-86.
- Atkins, F. J., & Chen, J. (2002), Some Statistical Properties of Deregulated Electricity Prices in Alberta, *University of Calgary Department of Economics*, Discussion Paper.
- Avsar, G. S., & Goss, B. A. (1999), Efficiency and Liquidity in the Electricity Market: A Preliminary Analysis, *Monash University Department of Economics*, Discussion Paper.
- Avsar, S. G, & Goss, B. A. (2001). Forecast Errors and Efficiency in the US Electricity Futures Market, *Australian Economic Papers*, 40(4), 479-99.
- Ballie, R. T. (1996), Long Memory Processes and Fractional Integrations in Econometrics, *Journal of Econometrics*, 73(1), 5-59.
- Barlow, M. T. (2002), A Diffusion Model For Electricity Prices, *Mathematical Finance*, 12(4), 287-98.
- Benth, F. E., & Schmeck, D.M. (2014), Pricing Futures and Options in Electricity Markets, *The Interrelationship Between Financial and Energy Markets*, Lecture Notes in Energy, (54), 233-60.
- Bessembinder, H., & Lemmon, M. L. (2002), Equilibrium Pricing and Optimal Hedging in Electricity Forward Markets, *The Journal of Finance*, 57(3), 1347-82.
- Black, F, (1976), Studies of Stock Price Volatility Changes, *American Statistical Association*, Proceedings of the 1976 Meeting of the Business and Economic Statistics Section, 177-81.
- Bollerslev, T, (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics*, 31(3), 307-27.
- Bunn, D.W. (2000), Forecasting Loads and Prices in Competitive Power Markets, *Proceedings of the IEEE*, 88(2), 163-9.

- Cartea, Á, & Figueroa, M.G. (2005), Pricing in Electricity Markets: A Mean Reverting Jump Diffusion Model with Seasonality, *Applied Mathematical Finance*, 12(4), 313-35.
- Dacorogna, M. M. (2001), *An Introduction to High-frequency Finance*, San Diego, Academic Press.
- Dempster, A. P., Laird, N.M., & Rubin D.B. (1977), Maximum Likelihood from Incomplete Data via the EM Algorithm." *Journal of the Royal Statistical Society Series B (Methodological)*, 39(1), 1-38.
- De Jong, C., & Huisman, R. (2002), Option Formulas for Mean-Reverting Power Prices with Spikes, *Erasmus Univeristy Working Paper*.
- De Jong, C. (2006), The Nature of Power Spikes: A Regime-Switching Approach, *Studies in Nonlinear Dynamics & Econometrics*, 10(3), Article 3.
- Deng, S.-J., Johnson, B, & Sogomonian, A. (2001), Exotic Electricity Options and the Valuation of Electricity Generation and Transmission Assets, *Decision Support Systems*, 30(3), 383-92.
- Dickey, D. A. (1979), Distribution of the Estimators for Autoregressive Time Series with a Unit Root, *Journal of the American Statistical Association*, 74(366), 427-31.
- Dickey, D. A. & Fuller, W. A. (1981), Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, *Econometrica*, 49, 1057-72.
- Ding, Z., & Granger, C. W. J. (1996), Modeling Volatility Persistence of Speculative Returns: A New Approach, *Journal of Econometrics*, 73, 185-215.
- Ding, Z., Granger, C. W. J., & Engle, R.F. (1993), A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, 1(1), 83-106.
- Engle, R. F., & Ng, V. K. (1993), Measuring and Testing the Impact of News on Volatility, *The Journal of Finance*, 48(5), 1749-78.
- Engle, R. F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50(4), 987-1008.

- Engle, R.F., Ito, T., & Lin, W-L (1990), Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market, *Econometrica*, 58(3), 525-42.
- Erdogdu, E. (2016), Asymmetric Volatility in European Day-ahead Power Markets: A Comparative Microeconomic Analysis, *Energy Economics*, 56 (2016): 398-409.
- Escribano, A., Peña, J. I., & Villaplana, P. (2011), Modelling Electricity Prices: International Evidence, *Oxford Bulletin of Economics and Statistics*, 73(5), 622-50.
- Ethier, R., & Mount, T. (1998), Estimating the Volatility of Spot Prices in Restructured Electricity Markets and the Implications for Option Values, *Cornell University Department of Economics*, Working Paper.
- Eydeland, A., & Wolyniec, K. (2003), *Energy and Power Risk Management: New Developments in Modeling, Pricing, and Hedging*. Hoboken, N.J: Wiley.
- Fezzi, C. (2005), Econometric Models for the Analysis of Electricity Markets (PhD Thesis), University Di Bologna, Bologna, Italy.
- Franses, P. H., & Van Dijk, D. (2000), *Nonlinear Time Series Models in Empirical Finance*, Cambridge, UK: Cambridge University Press.
- Geman, H. (2016), Understanding the Fine Structure of Electricity Prices, *The Journal of Business*, 79(3), 1225-61.
- Geweke, J., & Porter-Hudak, S. (1983), The Estimation and Application Of Long Memory Time Series Models, *Journal of Time Series Analysis*, 4(4), 221-38.
- Ghalanos, A. (2017), Univariate GARCH Models, *Comprehensive R Archive Network*, Version 1.3-8
- Giot, P. & Laurent, S. (2003), Market Risk in Commodity Markets: A VaR Approach, *Energy Economics*, 25(5), 435-57.
- Glosten, L.R., Jagannathan, R., & Runkle, D.E. (1993), On the Relation between the Expected Value and the Volatility of the Nominal Excess Return on Stocks, *The Journal of Finance*, 48(5), 1779-801.

- Goldfeld, S. M., & Quandt, R.E. (1973), A Markov Model for Switching Regressions, *Journal of Econometrics*, 1(1), 3-15.
- Hadsell, L., Marathe, A, & Shawky, H. A. (2004), Estimating the Volatility of Wholesale Electricity Spot Prices in the US, *The Energy Journal*, 25(4), 23-40.
- Haldrup, N., Nielsen, M. O. (2006), A Regime Switching Long Memory Model for Electricity Prices, *Journal of Econometrics*, 135, 349-76.
- Hamilton, J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica*, 57(2), 357-84.
- Hamilton, J. D. (1994), *Time Series Analysis*. Princeton, NJ: Princeton University Press.
- Higgs, H. (2005), Systematic Features of High-Frequency Volatility in Australian Electricity Markets: Intraday Patterns, Information Arrival and Calendar Effects, *The Energy Journal*, 26(4), 23-41.
- Hosking, J. R. M. (1981), Fractional Differencing, *Biometrika*, 68(1), 165-76.
- Hotelling, H. (1931), The Economics of Exhaustible Resources, *Journal of Political Economy*, 39(2), 137-75.
- Huisman, R., & Mahieu, R. (2003), Regime Jumps in Electricity Prices, *Energy Economics*, 25(5), 425-34.
- Hurst, H.E. (1951), Long-term Storage Capacity of Reservoirs, *Transactions of the American Society of Civil Engineers*, (116), 770-99.
- Janczura, J. (2014), Pricing Electricity Derivatives within a Markov Regime-switching Model: A Risk Premium Approach, *Mathematical Methods of Operations Research*, 79(1), 1-30.
- Janczura, J., & Weron, R. (2009), Regime-switching Models for Electricity Spot Prices: Introducing Heteroskedastic Base Regime Dynamics and Shifted Spike Distributions, *6th International Conference on the European Energy Market*, Leuven, Belgium.
- Jarque, C. M., & Bera, A. K. (1987), A Test for Normality of Observations and Regression Residuals, *International Statistical Review*, 55(2) : 163-72.

- Kanamura, T., & Kazuhiko A. (2008), On Transition Probabilities of Regime Switching in Electricity Prices, *Energy Economics* 30(3), 1158-172.
- Karakatsani, N. V., & Bunn, D. W. (2008), Forecasting Electricity Prices: The Impact of Fundamentals and Time-varying Coefficients, *International Journal of Forecasting*, 24(4), 764-85.
- Khintchine, A. (1934), Korrelationstheorie der Stationären Stochastischen Prozesse, *Mathematische Annalen*, 109(1), 604-15.
- Knittel, C. R., & Roberts, M. R. (2005), An Empirical Examination of Restructured Electricity Prices, *Energy Economics*, 27(5), 791-817.
- Kwiatkowski, D., Phillips, P. C. B, Schmidt, P., & Shin, Y. (1992), Testing the Null Hypothesis of Stationarity against the Alternative of a Unit Root, *Journal of Econometrics*, 54, 159-78.
- Ljung, G. M., & Box, G. E. P. (1978), On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 65(2), 297-303.
- Lo, A W. (1991), Long-Term Memory in Stock Market Prices, *Econometrica*, 59(5), 1279-313
- Longstaff, F. A., & Wang, A. W. (2004), Electricity Forward Prices: A High-Frequency Empirical Analysis, *Journal of Finance*, 59(4), 1877-900.
- Lucia, J. & Schwartz, E. (2000), Electricity Prices and Power Derivatives: Evidence from the Nordic, *Review of Derivatives Research* 5(5), 5-48.
- Montero, J., Garcia-Centeno, M., & Fernandez-Aviles, G. (2011), Modelling the Volatility of the Spanish Wholesale Electricity Spot Market. Asymmetric GARCH Models vs. Threshold ARSV Model, *Estudios De Economía Aplicada*, 29(August), 597-616.
- Mount, T. D., Ning Y., & Cai, X. (2006), Predicting Price Spikes in Electricity Markets Using a Regime-switching Model with Time-Varying Parameters, *Energy Economics*, 28(1), 62-80.
- Nelson, D. B. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59(2), 347-70.

- Newey, W. & West K. (1986), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 55(3), 703-8.
- Ornstein, L. S. & Uhlenbeck, G. E. (1930), On the Theory of Brownian Motion, *Physical Review*, 36, 823-41.
- Otero, J., & Smith, J. (2005), The KPSS Test with Outliers, *Computational Economics*, 26(3), 59-67.
- Park, H., Mjelde, J. W., & Bessler, D. A. (2006), Price Dynamics among U.S. Electricity Spot Markets, *Energy Economics*, 28(1), 81-101.
- Pereira, J. P., Pesquita, V., Rodrigues, P. M. M., & Rua, A., (2016), Market Integration and the Persistence of Electricity Prices, *Banco de Portugal*, Working Paper.
- Phillips, P. C. B. (1987), Time Series Regression with a Unit Root, *Econometrica*, 55(2), 277-301.
- Phillips, P. C. B., & Perron, P. (1988), Testing for a Unit Root in Time Series Regression, *Biometrika*, 75(2), 335-46.
- Pindyck, R. S. (1999), The Long-Run Evolution of Energy Prices, *Massachusetts Institute of Technology Center for Energy and Environmental Policy Research*, Working Paper.
- Pindyck, R. S. (2001), The Dynamics of Commodity Spot and Futures Markets: A Primer, *The Energy Journal*, 22(3), 1-30.
- Rambharat, B. R., Brockwell, A. E., & Seppi, D. J. (2005), A Threshold Autoregressive Model for Wholesale Electricity Prices, *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 54(2), 287-99.
- Ramsey, J. B. (2002), Wavelets in Economics and Finance: Past and Future, *Journal of Nonlinear Dynamics and Econometrics*, 6(3), 1-29.
- Said, S. E. & Dickey, D. A. (1984), Testing for Unit Roots in Autoregressive-Moving Average Models of Unknown Order, *Biometrika*, 71(3), 599-607.
- Saravia, C. (2003), Speculative Trading and Market Performance, *University of California Energy Institute*, Working Paper.

- Schwartz, E. & Smith J. E. (2000), Short-Term Variations and Long-Term Dynamics in Commodity Prices, *Management Science* 46(7), 893-911.
- Schwartz, E. (1997), The Stochastic Behavior of Commodity Prices: Implications for Valuation and Hedging, *The Journal of Finance*, 52(3), 923-73.
- Schwert, G. W. (1989), Test for Unit Roots: A Monte Carlo Investigation, *Journal of Business & Economic Statistics*, 7(2), 147-59.
- Shenoy, S. & Gorinevsky, D. (2014), Gaussian-Laplacian Mixture Model for Electricity Market, *53rd IEEE Conference on Decision and Control*, 1720-6.
- Taylor, S. (1986), *Modelling Financial Time Series*, Hackensack, NJ: World Scientific.
- Terasvirta, T. & Malmsten H. (2010), Stylized Facts of Financial Time Series and Three Popular Models of Volatility, *European Journal of Pure and Applied Mathematics*, 3(3), 417-43.
- Weron, R. (2002), *Measuring Long-Range Dependence in Electricity Prices*, *Empirical Science of Financial Fluctuations*, Tokyo: Springer Verlag
- Weron, R. & Misiorek, A. (2008), Forecasting Spot Electricity Prices: A Comparison of Parametric and Semiparametric Time Series Models, *International Journal of Forecasting*, 24(4), 744-63.
- Worthington, A., Kay-Spratley, A, & Higgs, H. (2005), Transmission of Prices and Price Volatility in Australian Electricity Spot Markets: A Multivariate GARCH Analysis, *Energy Economics*, 27(2), 337-50.
- Zivot, E. & Wang, J. (2006), *Modeling Financial Time Series with S-plus*. New York, NY: Springer.

**Table 1: Power Production Function**

The production cost function is estimated, with specifications to account for time trend  $\beta$ , and monthly variation. A parameter estimate of  $(c - 1) > 1$  suggests convexity in the production function.

Specification 1:  $\ln(\bar{F}_{Peak}^t) = a + (c - 1)\ln(\bar{Q}^t) + \varepsilon_t$

Specification 2:  $\ln(\bar{F}_{Peak}^t) = a + (c - 1)\ln(\bar{Q}^t) + \beta t + \varepsilon_t$

Specification 3:  $\ln(\bar{F}_{Peak}^t) = \mu' D_{it} + (c - 1)\ln(\bar{Q}^t) + \beta t + \varepsilon_t$ ,

$D_{it}$ : vector of dummy variables for each month  $i$  of the year

	(1)	(2)	(3)
$(c - 1)$	1.401*** (0.055)	1.993*** (0.049)	1.920*** (0.047)
$a$	-12.026*** (0.623)	-18.320*** (0.552)	
$\beta$		-0.0003*** (0.00001)	-0.0002*** (0.00001)
$D_{Jan}$			-17.668*** (0.525)
$D_{Feb}$			-17.571*** (0.533)
$D_{Mar}$			-17.552*** (0.528)
$D_{Apr}$			-17.161*** (0.524)
$D_{May}$			-17.163*** (0.523)
$D_{Jun}$			-16.823*** (0.533)
$D_{Jul}$			-16.952*** (0.535)
$D_{Aug}$			-17.197*** (0.533)
$D_{Sep}$			-17.142*** (0.529)
$D_{Oct}$			-17.375*** (0.523)
$D_{Nov}$			-17.317*** (0.527)
$D_{Dec}$			-17.516*** (0.531)
Observations	2,922	2,922	2,922
R <sup>2</sup>	0.183	0.425	0.993
Adjusted R <sup>2</sup>	0.182	0.425	0.993
Residual Std. Error	0.423 (df = 2920)	0.355 (df = 2919)	0.329 (df = 2908)
F Statistic	651.954*** (df = 1; 2920)	1,079.590*** (df = 2; 2919)	29,091.960*** (df = 14; 2908)

Notes: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

df: degrees of freedom, k: number of parameters, T: number of observations

R<sup>2</sup>:  $1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y})^2)$

Adjusted R<sup>2</sup>:  $1 - [(1 - R^2)(T - 1)] / (T - k)$

Residual Std. Error:  $\sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / (T - k)}$

F Statistic:  $\frac{R^2 / (k-1)}{(1-R^2) / (T-k)}$  to examine the overall significance of the regression.

**Table 2:** Descriptive Statistics on Financial and Electricity Returns

Sample Statistic:		SP500	USD/EUR	Nat Gas	WTI Crude	RT Off Peak	RT Peak	DA Off Peak	DA Peak
Mean	$\hat{\mu}$	1.11E-04	-1.05E-04	-6.24E-04	-3.59E-04	-8.33E-05	-2.75E-04	-2.40E-04	-3.41E-04
Variance	$\hat{\sigma}^2$	1.95E-04	4.56E-05	1.62E-03	6.55E-04	3.77E-01	1.74E-01	2.99E-02	4.53E-02
Skewness	$\hat{\mu}_3/\hat{\sigma}^3$	-9.52E-07	5.19E-08	5.87E-05	1.22E-06	-6.90E-03	-1.31E-02	2.92E-03	-1.05E-03
Kurtosis	$\hat{\mu}_4/\hat{\sigma}^4$	4.52E-07	1.19E-08	5.11E-05	3.32E-06	2.63E+00	1.85E-01	1.64E-02	1.89E-02
Annualized Volatility	$\hat{\sigma}^2\sqrt{T/250}$	-4.87E-03	2.12E-03	2.41E-06	5.52E-07				
Annualized Volatility	$\hat{\sigma}^2\sqrt{T/365}$					-1.55E-02	9.81E+00	-7.38E-01	1.04E+01
Jarque Bera	JB	7.10E+03	6.87E+02	2.46E+04	2.02E+03	1.07E+02	4.71E+01	2.56E+03	6.74E+01
Ljung Box 30 lags	$Q_{30}$	7.61E+01	2.36E+01	1.82E+02	1.04E+02	3.93E+02	3.64E+02	3.40E+02	1.28E+03
Ljung Box 60 lags	$Q_{60}$	1.29E+02	4.60E+01	2.25E+02	1.86E+02	4.29E+02	5.10E+02	4.87E+02	2.31E+03
Ljung Box 90 lags	$Q_{90}$	1.82E+02	7.58E+01	2.51E+02	2.38E+02	4.77E+02	6.23E+02	5.45E+02	3.21E+03

Note: JB results that reject the null of normality at 1% critical value: 9.21 are outlined with a box

LB results that reject the null of no autocorrelation at the 1% critical values for 30, 60, and 90 degrees of freedom of 50.90, 88.38, and 122.95 respectively outlined with a box

**Table 3:** Descriptive Statistics on Financial and Electricity Prices

Sample Statistic:		SP500	USD/EUR	Nat Gas	WTI Crude	RT Off Peak	RT Peak	DA Off Peak	DA Peak
Mean	$\hat{\mu}$	1.46E+03	1.33E+00	4.43E+00	8.28E+01	2.76E+01	5.43E+01	2.82E+01	5.36E+01
Variance	$\hat{\sigma}^2$	2.27E+06	1.78E+00	2.39E+01	7.36E+03	1.12E+03	4.75E+03	9.97E+02	4.04E+03
Skewness	$\hat{\mu}_3/\hat{\sigma}^3$	3.73E+09	2.41E+00	1.60E+02	6.88E+05	1.15E+05	8.55E+05	6.15E+04	4.99E+05
Kurtosis	$\hat{\mu}_4/\hat{\sigma}^4$	6.42E+12	3.28E+00	1.30E+03	6.69E+07	2.94E+07	3.27E+08	8.46E+06	1.02E+08
Annualized Volatility	$\hat{\sigma}^2\sqrt{T/250}$	6.18E+01	8.29E+01	1.12E+02	1.52E+02				
Annualized Volatility	$\hat{\sigma}^2\sqrt{T/365}$					3.05E+03	2.67E+05	4.81E+07	1.84E+10
Jarque Bera	JB	1.07E+02	4.71E+01	2.56E+03	6.74E+01	1.07E+02	4.71E+01	2.56E+03	6.74E+01
Ljung Box 30 lags	$Q_{30}$	7.31E+01	2.35E+01	2.25E+02	8.00E+01	6.94E+02	9.08E+02	3.99E+02	4.38E+02
Ljung Box 60 lags	$Q_{60}$	1.05E+02	4.58E+01	2.75E+02	1.38E+02	1.15E+03	1.48E+03	7.92E+02	7.39E+02
Ljung Box 90 lags	$Q_{90}$	1.32E+02	7.51E+01	3.09E+02	1.80E+02	1.21E+03	1.64E+03	8.23E+02	8.89E+02

Note: JB results that reject the null of normality at 1% critical value: 9.21 are outlined with a box

LB results that reject the null of no autocorrelation at the 1% critical values for 30, 60, and 90 degrees of freedom of 50.90, 88.38, and 122.95 respectively outlined with box

**Table 4:** GARCH Results for Financial Returns

Parameter	Panel A: SP500				Panel B: EUR/USD			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	0.001	0.000	3.751	0.000	0.000	0.000	-0.564	0.573
$\varphi$	0.733	0.205	3.579	0.000	-0.009	1.765	-0.005	0.996
$\theta$	-0.782	0.188	-4.155	0.000	0.015	1.776	0.008	0.993
$\omega$	0.000	0.000	1.969	0.049	0.000	0.000	0.338	0.735
$\alpha$	0.117	0.015	7.636	0.000	0.036	0.004	9.101	0.000
$\beta$	0.869	0.016	55.509	0.000	0.962	0.004	264.989	0.000
Log Likelihood	6750.399				7909.318			
Sign Bias			1.942	0.052			0.286	0.775
Negative Sign Bias			1.541	0.124			1.113	0.266
Positive Sign Bias			2.322	0.020			1.705	0.088
Joint Effect			17.007	0.001			4.153	0.245
Parameter	Panel C: Natural Gas				Panel D: WTI Crude			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	-0.001	0.001	-0.969	0.332	0.000	0.000	0.790	0.430
$\varphi$	-0.640	0.141	-4.550	0.000	-0.608	0.301	-2.017	0.044
$\theta$	0.701	0.130	5.388	0.000	0.574	0.311	1.846	0.065
$\omega$	0.000	0.000	3.836	0.000	0.000	0.000	1.406	0.160
$\alpha$	0.125	0.014	9.132	0.000	0.079	0.015	5.407	0.000
$\beta$	0.868	0.013	66.143	0.000	0.918	0.015	59.808	0.000
Log Likelihood	4364.164				5351.677			
Sign bias			1.695	0.090			2.812	0.005
Negative Sign Bias			0.969	0.333			0.509	0.611
Positive Sign Bias			0.826	0.409			1.533	0.126
Joint Effect			2.932	0.402			8.916	0.030

**Table 5:** GARCH Results for Electricity Returns

Parameter	Panel A: RT Off Peak				Panel B: RT Peak			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	-0.002	0.001	-1.888	0.059	0.000	0.001	-0.312	0.755
$\varphi$	0.298	0.026	11.539	0.000	0.349	0.022	15.721	0.000
$\theta$	-0.912	0.011	-82.065	0.000	-0.910	0.010	-93.019	0.000
$\omega$	0.008	0.001	9.184	0.000	0.007	0.001	6.363	0.000
$\alpha$	0.054	0.006	8.700	0.000	0.146	0.017	8.613	0.000
$\beta$	0.915	0.007	125.442	0.000	0.811	0.020	41.194	0.000
Log Likelihood	-1972.448				-992.443			
Sign Bias			1.553	0.121			2.188	0.029
Negative Sign Bias			1.153	0.249			1.182	0.237
Positive Sign Bias			0.446	0.656			0.116	0.908
Joint Effect			6.302	0.098			19.037	0.000
	Panel C: DA Off Peak				Panel D: DA Peak			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	0.000	0.001	0.588	0.557	0.001	0.003	0.201	0.841
$\varphi$	0.690	0.022	31.291	0.000	-0.698	0.075	-9.273	0.000
$\theta$	-0.916	0.012	-79.292	0.000	0.775	0.066	11.813	0.000
$\omega$	0.001	0.000	6.225	0.000	0.001	0.000	4.007	0.000
$\alpha$	0.233	0.019	12.423	0.000	0.128	0.015	8.509	0.000
$\beta$	0.766	0.016	47.717	0.000	0.862	0.016	55.012	0.000
Log Likelihood	1906.596				839.755			
Sign bias			1.145	0.252			0.165	0.869
Negative Sign Bias			1.809	0.071			1.865	0.062
Positive Sign Bias			0.380	0.704			1.605	0.109
Joint Effect			11.479	0.009			12.309	0.006

**Table 6:** APARCH Results for Financial Returns

Parameter	Panel A: SP500				Panel B: EUR/USD			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	0.000	0.000	3.115	0.002	0.000	0.000	-3.170	0.002
$\varphi$	0.173	0.049	3.522	0.000	0.385	0.021	18.623	0.000
$\theta$	-0.234	0.048	-4.910	0.000	-0.388	0.021	-18.745	0.000
$\omega$	0.000	0.000	1.449	0.147	0.000	0.000	3.903	0.000
$\alpha$	0.100	0.014	7.341	0.000	0.041	0.003	15.768	0.000
$\beta$	0.898	0.015	61.028	0.000	0.966	0.000	2356.043	0.000
$\gamma$	1.000	0.001	1745.871	0.000	0.286	0.107	2.681	0.007
$\delta$	0.986	0.132	7.446	0.000	0.846	0.061	13.958	0.000
$\nu$	7.258	1.191	6.094	0.000	8.641	1.551	5.572	0.000
Log Likelihood	6837.749				7946.554			
Sign bias			0.262	0.794			0.239	0.811
Negative Sign Bias			2.762	0.006			1.671	0.095
Positive Sign Bias			2.643	0.008			1.896	0.058
Joint Effect			14.666	0.002			6.508	0.089
Parameter	Panel C: Natural Gas				Panel D: WTI Crude			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	-0.001	0.001	-1.547	0.122	0.000	0.000	0.109	0.913
$\varphi$	-0.561	0.040	-13.858	0.000	-0.719	0.024	-29.794	0.000
$\theta$	0.639	0.038	17.020	0.000	0.688	0.025	27.197	0.000
$\omega$	0.001	0.000	1.551	0.121	0.000	0.000	1.073	0.283
$\alpha$	0.120	0.016	7.677	0.000	0.060	0.008	7.345	0.000
$\beta$	0.885	0.016	55.016	0.000	0.948	0.007	134.492	0.000
$\gamma$	0.114	0.090	1.271	0.204	0.551	0.122	4.519	0.000
$\delta$	1.094	0.187	5.852	0.000	0.924	0.187	4.939	0.000
$\nu$	4.969	0.566	8.777	0.000	8.344	1.327	6.289	0.000
Log Likelihood	4453.334				5351.677			
Sign bias			1.742	0.082			3.032	0.002
Negative Sign Bias			1.653	0.099			0.906	0.365
Positive Sign Bias			0.537	0.591			3.392	0.001
Joint Effect			5.205	0.157			14.076	0.003

**Table 7:** APARCH Results for Electricity Returns

Parameter	Panel A: RT Off Peak				Panel B: RT Peak			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	0.003	0.001	5.836	0.000	-0.004	0.000	-8.896	0.000
$\varphi$	0.479	0.025	19.517	0.000	0.343	0.019	18.146	0.000
$\theta$	-0.940	0.010	-90.566	0.000	-0.907	0.009	-97.567	0.000
$\omega$	0.210	0.042	5.005	0.000	0.007	0.000	15.483	0.000
$\alpha$	1.000	0.178	5.629	0.000	0.087	0.017	5.233	0.000
$\beta$	0.308	0.071	4.343	0.000	0.926	0.019	47.739	0.000
$\gamma$	0.224	0.061	3.665	0.000	-0.786	0.118	-6.666	0.000
$\delta$	1.140	0.148	7.685	0.000	0.666	0.079	8.409	0.000
$\nu$	2.179	0.045	48.421	0.000	4.124	0.306	13.463	0.000
Log Likelihood	-585.027				-782.836			
sign bias			0.735	0.462			0.883	0.377
Negative Sign Bias			2.101	0.036			1.090	0.276
Positive Sign Bias			1.133	0.257			0.133	0.894
Joint Effect			6.052	0.109			5.356	0.147
Parameter	Panel C: DA Off Peak				Panel D: DA Peak			
	Estimate	Std. Error	t value	p	Estimate	Std. Error	t value	p
$\mu$	0.000	0.000	-0.840	0.401	0.005	0.001	3.957	0.000
$\varphi$	0.688	0.021	32.807	0.000	-0.583	0.020	-28.768	0.000
$\theta$	-0.916	0.011	-83.420	0.000	0.678	0.018	36.692	0.000
$\omega$	0.004	0.002	2.259	0.024	0.004	0.001	3.144	0.002
$\alpha$	0.290	0.038	7.705	0.000	0.090	0.009	9.616	0.000
$\beta$	0.723	0.037	19.301	0.000	0.921	0.010	89.335	0.000
$\gamma$	0.127	0.044	2.894	0.004	-1.000	0.001	-921.644	0.000
$\delta$	1.396	0.177	7.882	0.000	0.930	0.114	8.141	0.000
$\nu$	4.626	0.373	12.401	0.000	5.130	0.504	10.176	0.000
Log Likelihood	2117.478				998.150			
sign bias			1.213	0.225			1.056	0.291
Negative Sign Bias			0.664	0.507			0.436	0.663
Positive Sign Bias			0.068	0.946			1.630	0.103
Joint Effect			4.323	0.229			2.847	0.416

**Table 8: ADF Unit-Root Test Results**

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	-3.394	-2.374	-4.648	-4.387	-3.288	-2.966	-3.129	-3.039	-2.494	-3.105	-3.94	-5.163	-2.154	-2.355	-2.837	-2.975	-6.734	-6.049
2	-2.805	-2.565	-3.737	-4.579	-3.685	-3.093	-3.026	-2.54	-2.848	-1.987	-3.566	-3.663	-2.288	-1.945	-2.206	-2.793	-6.739	-5.604
3	-3.587	-2.613	-4.625	-4.204	-3.381	-3.036	-3.091	-2.547	-2.895	-1.984	-2.698	-2.829	-2.299	-1.9	-2.38	-2.658	-7.089	-5.463
4	-3.141	-2.904	-4.013	-3.859	-3.345	-2.968	-2.953	-2.484	-2.403	-2.035	-1.986	-2.22	-2.43	-1.85	-2.375	-2.548	-6.585	-5.45
5	-3.062	-2.65	-5.156	-3.607	-3.731	-2.943	-2.919	-2.459	-2.76	-1.986	-2.377	-1.863	-2.365	-1.879	-2.463	-2.458	-6.93	-5.509
6	-2.967	-2.703	-3.191	-3.593	-3.983	-2.942	-2.807	-2.378	-2.396	-1.89	-2.62	-1.763	-2.701	-1.913	-2.912	-2.592	-7.434	-5.739
7	-2.536	-2.898	-3.132	-3.514	-4.306	-3.029	-2.004	-2.317	-1.937	-1.815	-2.834	-1.877	-2.368	-1.81	-2.562	-2.831	-7.061	-5.911
8	-2.838	-2.385	-2.954	-3.03	-4.871	-3.007	-2.427	-2.372	-2.347	-1.669	-3.219	-2.046	-2.891	-1.798	-2.113	-2.584	-7.305	-5.787
9	-3.079	-2.12	-3.878	-3.28	-4.023	-2.924	-2.656	-2.448	-2.347	-1.776	-3.706	-2.393	-3.138	-1.664	-3.091	-2.434	-7.798	-5.386
10	-3.061	-2.072	-3.69	-4.241	-3.735	-3.086	-3.073	-2.422	-3.015	-1.982	-4.723	-2.54	-3.366	-2.133	-3.124	-2.666	-8.045	-6.009
11	-2.664	-2.1	-4.589	-4.483	-3.423	-3.002	-3.12	-2.482	-3.235	-2.231	-4.749	-3.575	-2.921	-2.31	-3.118	-2.716	-7.21	-5.982
12	-2.315	-2.03	-4.544	-4.431	-2.756	-2.938	-2.86	-2.451	-3.595	-2.434	-4.304	-5.147	-2.799	-2.275	-3.599	-2.791	-6.367	-5.769
13	-1.743	-1.854	-4.652	-4.502	-2.716	-2.267	-2.585	-2.38	-2.635	-2.105	-4.133	-4.955	-2.395	-2.528	-3.526	-3.023	-5.703	-5.77
14	-1.561	-1.752	-4.555	-4.547	-2.163	-1.82	-2.214	-2.271	-3.161	-1.96	-4.022	-4.193	-2.477	-2.595	-4.215	-3.253	-6.102	-5.478
15	-1.871	-1.703	-4.232	-4.13	-2.001	-1.522	-2.404	-2.406	-3.179	-1.885	-2.996	-4.013	-2.535	-2.644	-3.324	-3.353	-5.596	-5.617
16	-2.105	-1.612	-3.536	-3.441	-1.945	-1.387	-3.425	-2.334	-3.02	-1.882	-3.39	-3.755	-3.084	-2.6	-3.468	-3.096	-6.582	-5.519
17	-2.042	-1.576	-3.65	-2.955	-2.046	-1.268	-3.073	-2.127	-2.719	-1.797	-3.523	-3.645	-2.517	-2.621	-3.483	-2.873	-6.316	-5.513
18	-2.175	-1.622	-2.929	-3.419	-2.415	-1.43	-3.383	-2.213	-3.379	-1.822	-4.266	-3.775	-2.614	-2.687	-3.876	-2.964	-6.756	-5.87
19	-2.541	-1.919	-2.87	-4.069	-3.302	-1.917	-3.132	-2.646	-2.889	-2.037	-4.825	-4.324	-1.999	-2.37	-3.522	-3.594	-6.455	-6.335
20	-2.933	-2.241	-3.147	-4.098	-3.371	-2.44	-3.506	-2.834	-2.354	-2.129	-5.027	-4.695	-2.378	-2.094	-2.728	-3.02	-6.384	-6.155
21	-2.368	-2.622	-3.666	-4.266	-3.769	-2.911	-3.285	-3.04	-3.783	-2.022	-4.855	-5.762	-2.187	-2.341	-3.402	-2.842	-6.458	-6.243
22	-2.082	-2.196	-3.28	-4.388	-2.708	-2.885	-2.755	-2.666	-2.952	-2.886	-4.249	-4.92	-2.462	-2.266	-3.323	-2.775	-5.877	-5.563
23	-3.175	-1.924	-4.345	-4.392	-3.029	-2.68	-2.875	-2.539	-2.753	-2.367	-3.843	-4.242	-2.389	-2.422	-2.991	-2.937	-5.703	-5.488
24	-3.17	-2.251	-3.85	-4.484	-2.927	-2.807	-2.905	-2.563	-3.021	-2.399	-4.494	-4.327	-2.347	-2.402	-2.797	-2.935	-6.197	-5.784
All Hours	-2.374	-2.365	-4.387	-4.284	-2.966	-2.859	-3.039	-2.508	-3.105	-2.279	-5.163	-4.217	-2.355	-2.187	-2.975	-2.802	-6.049	-5.499

Notes: Hours which reject the null in favor of the alternative indicating stationarity at the 1% level are outlined below.

$$\Delta p_t = \beta' D_t + (\phi - 1)p_{t-1} + \sum_{k=1}^K \psi \Delta p_{t-k} + \epsilon_t, H_0: \phi = 1 \Rightarrow p_t \sim I(1), \text{ against } H_1: |\phi| < 1 \Rightarrow p_t \sim I(0), \text{ with } k \text{ set to } 15 \text{ roughly equal to } k = \text{Int}\{12 \times (\frac{T}{100})^{\frac{1}{4}}\}$$

Critical values for test statistics: 5% 1%  
 -2.86 -3.43

**Table 9: PP Unit-Root Test Results**

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	-14.68	-5.95	-10	-4.83	-12.89	-4.9	-12.24	-7.12	-12.31	-6.5	-15.43	-5.76	-12.76	-9.11	-12	-7.29	-42.25	-25.41
2	-14.11	-6.35	-12.35	-4.83	-14.1	-4.87	-11.29	-7.35	-14.45	-6.27	-16.64	-6.07	-11.14	-8.02	-12.53	-6.4	-39.45	-22.26
3	-12.62	-7.2	-11.76	-4.62	-13.03	-4.96	-10.47	-7.59	-16.48	-6.21	-15.03	-5.77	-11.81	-8.07	-11.09	-6.07	-38.6	-22.52
4	-14.04	-6.85	-12.34	-4.33	-11.23	-4.83	-10.43	-7.51	-14.32	-5.7	-14.44	-5.55	-13.62	-8.58	-12.76	-6.24	-43.69	-23.68
5	-14.48	-7.43	-11.42	-4.56	-12.91	-4.77	-13.06	-7.54	-17.21	-5.95	-15.14	-5.74	-13.96	-8.73	-16.31	-6.11	-46.19	-24.34
6	-14.52	-8.66	-13.48	-5.59	-16.32	-5.45	-16.88	-8.76	-13.85	-8.12	-15.15	-6.87	-16.82	-9.79	-14.13	-7.6	-52.57	-28.12
7	-14.71	-10.61	-13.87	-6.83	-15.62	-7.2	-17.35	-10.6	-16.43	-9.62	-13.97	-8.87	-15.54	-10.32	-13.86	-7.56	-49.07	-30.17
8	-14.37	-10.25	-13.62	-6.62	-15.13	-7.35	-19.22	-10.8	-16.27	-10.39	-14.06	-8.59	-16.48	-11.24	-15.57	-8.23	-52.2	-33.88
9	-14.84	-9.23	-13.35	-5.6	-15.42	-6.16	-15.49	-9.07	-16.18	-9.28	-13.27	-8.75	-17.16	-11.48	-14.06	-6.8	-53.89	-32.54
10	-13.78	-8.24	-12.64	-5.45	-14.75	-6.05	-17.32	-8.81	-15.47	-10.1	-14.95	-8.25	-16.8	-11.65	-11.78	-7.74	-52.45	-31.78
11	-14.32	-7.56	-12.75	-6.09	-14.47	-6.53	-14.37	-9.44	-15	-10.87	-16.24	-8.38	-16.37	-10.96	-12.26	-8.88	-51.4	-28.84
12	-14.19	-5.8	-15.55	-7.07	-14.16	-6.38	-14.63	-8.14	-17.67	-10.28	-13.74	-7.25	-15.34	-11.46	-12.6	-9.12	-49.17	-28.31
13	-12.95	-4.77	-14.98	-7.89	-14.95	-5.89	-12.58	-7.25	-16.2	-9.94	-9.87	-6.18	-12.82	-10.35	-15.69	-9.33	-41.99	-23.46
14	-11.47	-4.34	-13.92	-8.07	-12.5	-5.61	-12.82	-7.23	-13.08	-10.49	-11.46	-5.82	-12.23	-10.43	-14.23	-9.86	-39.94	-23.09
15	-10.15	-3.79	-14.57	-7.85	-13.81	-5.13	-12.94	-7.18	-15.85	-10.07	-17.13	-5.5	-12.55	-9.98	-15.9	-9.99	-43.51	-21.55
16	-9.77	-3.49	-14.05	-7.27	-15.04	-4.68	-12.01	-6.97	-13.73	-9.8	-14.4	-5.43	-14.34	-9.77	-13.68	-9.71	-41.17	-20.66
17	-8.6	-3.46	-15.57	-6.95	-13.66	-4.55	-13.8	-7.06	-16.7	-9.56	-12.02	-5.49	-13.49	-10.12	-17.17	-9.39	-42.49	-21.51
18	-10.89	-4.22	-14.26	-4.82	-14.15	-4.91	-14.6	-7.54	-16.29	-10.12	-12.68	-5.5	-11.65	-9.39	-15.08	-9.19	-41.37	-22.74
19	-13.29	-4.94	-14.1	-4.41	-15.28	-5.23	-14.46	-7.72	-17.83	-11.04	-11.75	-6.01	-11.53	-9.47	-13.59	-9.9	-41.66	-26.07
20	-12.97	-5.62	-12.98	-4.77	-13.78	-5.52	-15.56	-8.08	-16.93	-8.92	-11.1	-7.25	-13.57	-9.66	-11.33	-10.23	-42.96	-26.94
21	-15.87	-5.39	-14.12	-5.31	-13.06	-5.61	-11.74	-7.62	-14.64	-9.71	-12.97	-7.49	-12.79	-9.43	-11.68	-10.21	-41.86	-24.98
22	-13.22	-4.66	-13.85	-5.33	-11.5	-5.34	-11.71	-7.13	-15.56	-8.43	-13.05	-7.21	-9.42	-8.85	-13.94	-9.51	-36.34	-22.38
23	-13.72	-5.02	-10.31	-4.67	-13.52	-4.76	-10.69	-6.84	-13.23	-6.73	-10.67	-7.09	-11.14	-8.14	-13.61	-11.27	-38.08	-22.55
24	-13.38	-5.51	-8.83	-4.28	-14.29	-4.64	-9.46	-6.19	-11.7	-6.71	-12.97	-7.05	-11.79	-7.31	-11.12	-10.48	-38.81	-20.49
All Hours	-5.95	-4.6	-4.83	-4.34	-4.9	-3.4	-7.12	-3.71	-6.5	-4.61	-5.76	-4.61	-9.11	-4.91	-7.29	-4.15	-25.41	-12.1

Notes: Hours which reject the null in favor of the alternative indicating stationarity at the 1% level are outlined.

$$\Delta p_t = \beta' D_t + \pi p_{t-1} + u_t, H_0: \phi = 1 \Rightarrow p_t \sim I(1), \text{ against } H_1: |\phi| < 1 \Rightarrow p_t \sim I(0)$$

Critical values for test statistics: 5% 1%  
 -2.86 -3.43

Table 10: KPSS Stationary Test Results

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	0.358	0.577	0.853	0.757	0.116	0.136	0.836	0.562	0.508	0.439	0.501	0.112	1.126	1.363	0.794	0.738	1.488	1.816
2	0.504	0.389	0.812	0.789	0.146	0.198	0.833	0.864	0.451	0.696	0.561	0.399	1.158	1.121	0.954	0.839	1.147	1.172
3	0.499	0.441	0.784	0.778	0.246	0.268	0.863	0.898	0.344	0.577	0.621	0.488	1.114	1.139	0.968	0.911	0.922	1.197
4	0.446	0.551	0.841	0.775	0.299	0.342	0.829	0.955	0.408	0.522	0.591	0.527	1.225	1.154	0.973	0.933	0.79	0.994
5	0.521	0.672	0.92	0.77	0.343	0.397	0.818	0.972	0.377	0.498	0.512	0.571	1.233	1.124	0.869	0.922	0.81	0.79
6	0.708	0.761	0.948	0.796	0.36	0.427	0.837	1.007	0.517	0.492	0.526	0.561	1.148	1.1	0.794	0.851	0.576	0.746
7	1.048	1.09	0.848	0.84	0.385	0.417	1.014	1.023	0.396	0.498	0.527	0.487	1.197	1.08	0.822	0.74	0.469	0.774
8	1.234	1.293	0.832	0.831	0.495	0.458	1.105	1.103	0.444	0.471	0.551	0.465	1.197	1.091	0.92	0.755	0.645	0.781
9	1.58	1.46	0.798	0.868	0.225	0.482	1.244	1.188	0.387	0.429	0.501	0.519	1.158	1.136	0.716	0.827	1.355	0.52
10	1.496	1.7	0.705	0.846	0.12	0.455	0.992	1.282	0.375	0.464	0.229	0.502	1.16	1.041	0.676	0.743	1.627	0.952
11	1.059	1.481	0.736	0.852	0.142	0.36	0.854	1.266	0.298	0.56	0.363	0.394	1.264	1.09	0.699	0.736	1.781	1.441
12	0.723	1.054	0.535	0.843	0.185	0.233	0.565	1.03	0.404	0.671	0.192	0.245	1.277	1.091	0.544	0.714	1.837	1.833
13	0.617	0.686	0.732	0.862	0.242	0.208	0.438	0.648	0.385	0.636	0.228	0.223	1.554	1.035	0.574	0.674	2.19	1.978
14	0.505	0.546	0.564	0.862	0.291	0.261	0.404	0.45	0.348	0.518	0.259	0.283	1.421	1.061	0.532	0.625	1.6	2.017
15	0.459	0.496	0.413	0.703	0.387	0.32	0.372	0.362	0.446	0.451	0.345	0.291	1.188	1.008	0.368	0.56	1.582	1.72
16	0.43	0.468	0.481	0.516	0.407	0.361	0.311	0.351	0.399	0.432	0.327	0.307	0.923	0.965	0.373	0.471	1.215	1.392
17	0.422	0.452	0.264	0.388	0.393	0.384	0.352	0.358	0.437	0.428	0.365	0.31	0.741	0.861	0.311	0.411	1.183	1.088
18	0.283	0.424	0.237	0.267	0.284	0.359	0.184	0.341	0.289	0.423	0.268	0.305	0.782	0.742	0.241	0.352	1.217	1.031
19	0.218	0.302	0.429	0.315	0.106	0.218	0.451	0.206	0.296	0.398	0.046	0.23	1.111	0.688	0.395	0.319	1.614	1.336
20	0.331	0.31	0.557	0.487	0.101	0.136	0.42	0.372	0.566	0.374	0.078	0.099	1.098	0.873	0.512	0.453	1.516	1.194
21	0.636	0.537	0.748	0.614	0.091	0.129	0.486	0.536	0.344	0.728	0.157	0.128	1.486	0.932	0.706	0.49	1.964	1.388
22	0.643	0.859	0.823	0.781	0.148	0.134	0.538	0.914	0.262	0.463	0.358	0.153	1.478	1.157	0.767	0.786	2.124	1.849
23	0.457	0.739	0.843	0.798	0.125	0.15	0.741	0.812	0.316	0.471	0.291	0.253	1.485	1.119	0.761	0.749	2.342	1.974
24	0.348	0.626	0.808	0.86	0.164	0.158	0.778	0.734	0.335	0.568	0.274	0.221	1.316	1.064	0.839	0.718	2.126	1.768
All Hours	0.577	0.523	0.757	0.825	0.136	0.212	0.562	0.808	0.439	0.577	0.112	0.221	1.363	1.083	0.738	0.773	1.816	1.902

Notes: Hours which support the null indicating stationarity at the 1% level are outlined.

$$p_t = \beta' D_t + \mu_t + u_t, \mu_t = \mu_{t-1} + \varepsilon_t, H_0: \sigma_\varepsilon^2 = 0 \Rightarrow p_t \sim I(0), \text{ against } H_1: \sigma_\varepsilon^2 > 0 \Rightarrow p_t \sim I(1)$$

Critical values of the test statistic: 5% 1%  
0.463 0.739

**Table 11:** ADF Unit-Root Test Results of Differenced Series

	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
Hour	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	-8.343	-6.995	-7.913	-6.583	-6.531	-6.456	-7.351	-6.726	-6.378	-7.007	-7.293	-6.489	-11.3	-12.95	-5.736	-5.795	-24.773	-25.469
2	-8.575	-9.315	-7.739	-6.722	-6.83	-6.595	-7.498	-7.21	-5.923	-5.776	-7.711	-7.472	-13.099	-8.207	-6.468	-5.514	-27.473	-23.636
3	-8.278	-8.833	-7.967	-6.571	-6.309	-6.442	-7.531	-7.229	-6.923	-5.826	-8.082	-7.537	-13.524	-8.249	-6.32	-5.638	-28.219	-23.155
4	-8.356	-8.812	-7.921	-6.672	-6.606	-6.249	-7.577	-7.341	-6.675	-5.801	-8.197	-7.486	-10.008	-8.485	-6.613	-5.651	-25.509	-23.27
5	-7.69	-9.185	-8.395	-6.411	-6.376	-6.182	-6.748	-7.388	-7.032	-5.724	-7.346	-7.273	-12.885	-8.261	-5.926	-5.654	-26.786	-22.964
6	-8.221	-9.136	-8.509	-6.362	-6.501	-6.149	-7.117	-7.484	-5.935	-5.647	-7.263	-7.214	-9.776	-7.883	-6.582	-5.733	-25.719	-22.549
7	-8.493	-8.522	-8.193	-6.401	-6.384	-6.202	-8.881	-7.785	-7.873	-5.649	-8.428	-7.507	-11.996	-7.767	-6.916	-6.177	-26.404	-22.622
8	-9.05	-8.1	-6.77	-5.814	-6.502	-5.764	-7.737	-7.759	-7.402	-5.863	-7.338	-7.279	-15.313	-8.338	-7.001	-6.474	-28.43	-22.938
9	-8.89	-7.706	-8.311	-5.924	-6.142	-5.613	-8.095	-7.37	-6.719	-6.208	-6.903	-6.791	-18.937	-9.523	-6.429	-6.707	-27.373	-24.205
10	-7.054	-7.639	-7.412	-6.405	-7.51	-6.028	-6.935	-7.404	-6.921	-5.739	-7.062	-6.691	-12.621	-7.834	-6.193	-6.036	-24.955	-22.244
11	-6.684	-7.775	-7.101	-6.322	-6.343	-6.15	-7.208	-7.195	-7.206	-5.718	-7.09	-6.899	-12.907	-7.73	-6.54	-5.903	-24.991	-21.859
12	-6.569	-7.46	-6.062	-6.14	-7.246	-6.511	-7.925	-6.719	-7.344	-6.091	-7.02	-7.024	-11.352	-7.369	-6.045	-5.703	-24.97	-21.126
13	-6.093	-7.187	-6.665	-6.003	-7.486	-6.535	-6.996	-6.4	-7.584	-6.307	-6.468	-7.071	-9.957	-7.348	-6.916	-5.907	-22.021	-21.072
14	-6.051	-7.098	-6.1	-5.975	-7.655	-6.763	-7.298	-6.212	-6.702	-6.939	-6.869	-6.969	-10.552	-7.391	-6.98	-5.639	-22.056	-21.081
15	-5.782	-6.931	-6.527	-5.913	-7.218	-7.131	-7.162	-6.365	-7.501	-7.296	-6.511	-7.058	-8.614	-7.393	-8.106	-5.852	-21.258	-20.795
16	-5.591	-6.856	-6.143	-5.891	-7.425	-7.227	-5.567	-6.363	-6.722	-7.383	-6.251	-6.78	-8.209	-7.37	-6.994	-6.007	-19.12	-20.519
17	-5.605	-6.886	-6.28	-6.103	-6.471	-7.228	-5.943	-6.272	-7.243	-7.453	-7.299	-6.692	-8.452	-7.203	-6.604	-6.172	-19.796	-20.283
18	-6.353	-6.879	-6.139	-6.03	-6.512	-6.924	-6.282	-6.246	-7.904	-7.017	-6.738	-6.673	-10.586	-7.149	-6.448	-6.311	-22.896	-20.283
19	-6.366	-6.872	-5.479	-5.717	-6.466	-6.475	-6.541	-6.133	-7.591	-6.602	-6.817	-6.792	-9.502	-7.243	-5.885	-6.388	-22.302	-20.53
20	-7.05	-7.049	-6.463	-5.744	-6.985	-6.172	-7.627	-6.176	-8.221	-6.939	-6.574	-6.794	-10.406	-7.469	-7.692	-6.193	-24.847	-20.886
21	-7.143	-7.318	-7.133	-5.899	-6.576	-6.251	-7.144	-6.604	-8.178	-6.718	-6.231	-7.026	-11.445	-7.633	-6.766	-6.064	-24.363	-21.335
22	-7.162	-7.604	-7.386	-6.242	-6.822	-6.143	-6.861	-6.546	-8.342	-6.502	-6.526	-6.577	-9.053	-7.82	-6.678	-6.27	-22.898	-21.719
23	-7.492	-7.56	-7.406	-6.24	-6.955	-5.99	-7.031	-6.5	-6.047	-5.901	-6.819	-6.819	-10.516	-7.59	-6.209	-6.077	-25.304	-21.581
24	-8.453	-8.194	-7.187	-6.564	-6.526	-6.217	-7.039	-6.697	-6.297	-5.605	-7.292	-7.173	-11.267	-7.46	-5.895	-6.344	-26.05	-21.533
<i>All Hours</i>	-6.995	-9.02	-6.583	-6.132	-6.456	-6.337	-6.726	-7.034	-7.007	-5.695	-6.489	-7.548	-12.95	-7.419	-5.795	-6.553	-25.469	-21.725

Notes:  $\Delta^2 p_t = \beta' D_t + (\phi - 1)\Delta p_{t-1} + \sum_{k=1}^k \psi \Delta^2 p_{t-k} + \epsilon_t$ ,  $H_0: \phi = 1 \Rightarrow p_t \sim I(2)$ , against  $H_1: |\phi| < 1 \Rightarrow p_t \sim I(1)$  with k set to 15 roughly equal  $tok = \text{Int}\{12 \times (\frac{\tau}{100})^{\frac{1}{2}}\}$

Critical values for test statistics: 5% 1%  
 -2.86 -3.43

**Table 12:** PP Unit-Root Test Results of Differenced Series

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	-58.74	-30.76	-38.49	-28.07	-43.9	-30.85	-41.18	-36.03	-39.85	-42.66	-69.34	-34.01	-48.71	-46.46	-35.37	-29.73	-163.07	-127.11
2	-55.23	-21.9	-50.42	-19.22	-52	-16.2	-36.36	-22.9	-51.68	-29.13	-74.28	-19.89	-41.97	-39.02	-40.46	-29.99	-145.57	-106.64
3	-47.1	-22	-43.01	-20.22	-44.51	-16.85	-31.73	-24.07	-62.22	-29.26	-63.3	-22.78	-47.12	-34.14	-38.12	-26.15	-143.11	-93.21
4	-55.85	-26.38	-49.16	-20.62	-40.86	-17.41	-37.31	-26.62	-56.26	-29.44	-60.88	-23.12	-55.42	-35.79	-43.91	-25.76	-168.24	-96.92
5	-55.72	-24.5	-41.53	-20.34	-49.53	-17.23	-47.12	-27.53	-66.55	-30.29	-57.66	-23.66	-58	-38.18	-57.07	-26.74	-177.26	-103.53
6	-55.95	-26.32	-51.85	-20.59	-62.92	-16.43	-72.9	-27.39	-47.9	-32.29	-61.37	-23.85	-73.73	-37.47	-50.07	-25.59	-225.06	-104.76
7	-52.11	-28.91	-48.64	-22.42	-50.62	-16.99	-70.66	-32.9	-58.35	-34.62	-53.4	-24.81	-69.39	-39.45	-49.98	-29.79	-211.49	-116
8	-52.78	-30.7	-48.86	-23.64	-52.64	-20.99	-79.08	-37.08	-57.1	-30.88	-53.16	-28.8	-76.69	-38.93	-58.72	-25.39	-226.24	-121.06
9	-61.67	-29.87	-54.42	-23.14	-52.23	-22.97	-55.06	-36.73	-57.74	-32.23	-49.08	-27.28	-80.26	-42.1	-50.04	-27.89	-240.96	-130.67
10	-50.29	-30.21	-47.42	-24.42	-54.47	-21.02	-65.01	-33.06	-55.62	-32.76	-60.14	-31.73	-75.44	-46.81	-43.26	-25.96	-230.1	-140.59
11	-49.79	-32.37	-47.75	-24.94	-53	-21.06	-52.89	-32.56	-57.22	-34.41	-65.01	-30.75	-70.2	-48.17	-43.86	-33.14	-220.49	-143.3
12	-53.61	-31.78	-55.05	-25.39	-52.44	-22.06	-55.27	-33.62	-72.85	-34.93	-53.43	-30.25	-65.75	-45.22	-43.55	-33.84	-206.51	-135.11
13	-50.98	-29.29	-53.2	-25.96	-57.68	-22.11	-42.66	-28.26	-61.36	-33.91	-38.88	-26.36	-53.09	-45.77	-59.4	-32.5	-179.5	-134.28
14	-48.69	-26.86	-45.98	-25.99	-44.96	-20.9	-47.47	-22.78	-46.15	-34.04	-40.02	-21.27	-50.94	-41.17	-56.43	-32.97	-164.3	-115.23
15	-49.84	-25.26	-49.76	-24.67	-54.43	-20.65	-49.45	-20.8	-60.78	-34.84	-66.27	-20.73	-48.82	-39.81	-64.04	-33.32	-192.16	-106.72
16	-48.26	-23.77	-45.82	-24.29	-64.12	-20.66	-40.1	-20.57	-49.28	-32.7	-52.28	-16.1	-56.37	-36.78	-50.74	-33.99	-176.86	-95.22
17	-36.87	-22.49	-55.78	-23.9	-55.06	-20.25	-48.68	-20.71	-62.78	-32.41	-45.52	-14.66	-51.96	-35.61	-63.2	-33.68	-178.77	-89.86
18	-43.53	-21.11	-53.6	-23.43	-51.13	-19.65	-50.1	-21.01	-65.07	-31.16	-47.23	-14.54	-44.86	-38.89	-53.86	-31.92	-163.93	-93.14
19	-47.16	-20.2	-50.6	-20.73	-48.88	-17.94	-50.95	-22.25	-76.34	-33.1	-41.03	-14.82	-45.59	-36.16	-50.72	-32.1	-166.8	-95.8
20	-43.1	-20.23	-45.5	-18.99	-45.62	-18.32	-62.57	-24.05	-67.74	-37.34	-38.47	-18.28	-55.42	-35.28	-38.59	-35.96	-166.59	-105.84
21	-56.8	-20.43	-49.35	-19.58	-46.71	-18.08	-44.07	-25.64	-57.47	-28.64	-46.4	-24.61	-53.35	-33.22	-37.99	-35.11	-166.57	-104.63
22	-48.21	-21.65	-49.7	-20.08	-39.27	-17.48	-42.27	-25.36	-63.67	-29.37	-47.87	-25.47	-30.65	-33.76	-46.17	-35.78	-139.67	-104.78
23	-47.92	-21.03	-39.38	-20.63	-43.68	-17.6	-36.18	-24.55	-46.78	-26.13	-37.42	-25.43	-44.1	-32.68	-45.12	-32.83	-143.48	-98.82
24	-44.35	-20.53	-34.24	-20.5	-47.07	-15.6	-28.18	-24.39	-37.43	-23.93	-52.41	-25.03	-45.78	-30	-31.87	-37.8	-136.26	-95.28
<i>All Hours</i>	-30.76	-20.66	-28.07	-20.47	-30.85	-16.55	-36.03	-23.56	-42.66	-27.03	-34.01	-26.55	-46.46	-26.81	-29.73	-35.91	-127.11	-87.79

Notes:  $\Delta^2 p_t = \beta' D_t + \pi \Delta p_{t-1} + u_t$ ,  $H_0: \phi = 1 \Rightarrow p_t \sim I(2)$ , against  $H_1: |\phi| < 1 \Rightarrow p_t \sim I(1)$

Critical values for test statistics: 5% 1%  
 -2.86 -3.43

**Table 13:** KPSS Stationary Test Results of Differenced Series

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	0.04	0.065	0.044	0.031	0.026	0.043	0.027	0.051	0.026	0.041	0.037	0.029	0.032	0.039	0.024	0.035	0.008	0.008
2	0.032	0.045	0.041	0.115	0.028	0.029	0.033	0.048	0.03	0.039	0.032	0.041	0.036	0.035	0.027	0.026	0.006	0.006
3	0.034	0.039	0.035	0.133	0.025	0.029	0.028	0.055	0.028	0.039	0.037	0.04	0.038	0.037	0.026	0.026	0.007	0.007
4	0.031	0.037	0.043	0.124	0.027	0.029	0.028	0.054	0.036	0.041	0.058	0.053	0.031	0.037	0.026	0.027	0.007	0.007
5	0.032	0.039	0.034	0.13	0.027	0.029	0.027	0.051	0.031	0.047	0.081	0.06	0.034	0.036	0.026	0.028	0.007	0.007
6	0.037	0.041	0.032	0.113	0.025	0.029	0.035	0.055	0.045	0.049	0.035	0.056	0.031	0.034	0.028	0.028	0.006	0.007
7	0.043	0.043	0.028	0.09	0.035	0.029	0.044	0.058	0.031	0.041	0.036	0.044	0.037	0.033	0.032	0.027	0.008	0.007
8	0.047	0.043	0.029	0.064	0.032	0.034	0.034	0.057	0.031	0.037	0.033	0.037	0.036	0.032	0.034	0.031	0.007	0.008
9	0.058	0.053	0.035	0.051	0.066	0.036	0.096	0.054	0.051	0.037	0.041	0.035	0.031	0.033	0.03	0.033	0.007	0.007
10	0.053	0.077	0.032	0.061	0.039	0.035	0.046	0.062	0.047	0.041	0.036	0.039	0.029	0.032	0.033	0.031	0.006	0.008
11	0.059	0.097	0.029	0.067	0.038	0.034	0.045	0.061	0.036	0.042	0.029	0.037	0.029	0.033	0.034	0.031	0.007	0.008
12	0.059	0.107	0.046	0.066	0.04	0.034	0.039	0.059	0.037	0.047	0.034	0.035	0.029	0.033	0.031	0.031	0.007	0.009
13	0.071	0.118	0.026	0.047	0.039	0.045	0.043	0.053	0.031	0.049	0.029	0.033	0.042	0.032	0.036	0.031	0.011	0.009
14	0.084	0.123	0.021	0.033	0.042	0.058	0.044	0.054	0.031	0.047	0.027	0.034	0.04	0.034	0.033	0.032	0.01	0.011
15	0.081	0.119	0.023	0.027	0.049	0.072	0.04	0.046	0.036	0.04	0.03	0.031	0.039	0.033	0.038	0.034	0.01	0.011
16	0.068	0.12	0.025	0.027	0.045	0.083	0.032	0.046	0.031	0.04	0.031	0.03	0.033	0.034	0.041	0.038	0.009	0.011
17	0.067	0.115	0.031	0.031	0.041	0.09	0.034	0.048	0.034	0.042	0.031	0.029	0.041	0.034	0.031	0.04	0.008	0.011
18	0.039	0.103	0.033	0.037	0.035	0.078	0.033	0.046	0.034	0.041	0.026	0.028	0.033	0.033	0.029	0.037	0.008	0.009
19	0.033	0.055	0.028	0.194	0.032	0.045	0.038	0.035	0.031	0.036	0.025	0.023	0.035	0.029	0.033	0.028	0.006	0.006
20	0.03	0.04	0.027	0.156	0.03	0.036	0.043	0.038	0.033	0.032	0.025	0.022	0.034	0.031	0.036	0.028	0.006	0.006
21	0.037	0.04	0.028	0.128	0.032	0.031	0.042	0.039	0.031	0.036	0.023	0.022	0.035	0.03	0.031	0.03	0.006	0.006
22	0.059	0.066	0.03	0.129	0.033	0.031	0.038	0.044	0.034	0.032	0.028	0.025	0.037	0.031	0.031	0.032	0.007	0.007
23	0.032	0.077	0.033	0.092	0.034	0.034	0.042	0.051	0.033	0.031	0.034	0.026	0.036	0.032	0.03	0.03	0.006	0.007
24	0.037	0.058	0.062	0.085	0.029	0.031	0.066	0.049	0.055	0.03	0.026	0.025	0.036	0.032	0.028	0.027	0.007	0.007
<i>All Hours</i>	0.065	0.054	0.031	0.142	0.043	0.031	0.051	0.056	0.041	0.04	0.029	0.025	0.039	0.033	0.035	0.027	0.008	0.007

Notes:  $\Delta p_t = \beta' D_t + \mu_t + u_t, \mu_t = \mu_{t-1} + \varepsilon_t, H_0: \sigma_\varepsilon^2 = 0 \Rightarrow p_t \sim I(1)$ , against  $H_1: \sigma_\varepsilon^2 > 0 \Rightarrow p_t \sim I(2)$

Critical values of the test statistic: 5% 1%  
0.463 0.739

**Table 14:** Lo's Modified R/S Long Range Dependency Test Results

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	1.252	1.679	1.82	1.749	1.189	1.272	1.269	1.497	1.681	1.679	1.422	1.237	1.725	1.817	1.472	1.329	2.3	2.626
2	1.197	1.499	1.814	1.798	1.335	1.493	1.246	1.285	1.805	1.758	1.582	1.581	1.747	1.651	1.532	1.413	2.06	2.163
3	1.34	1.154	1.738	1.812	1.586	1.655	1.393	1.339	1.759	1.852	1.892	1.804	1.687	1.673	1.546	1.459	1.747	2.132
4	1.587	1.361	1.761	1.808	1.763	1.778	1.404	1.421	1.913	1.982	1.965	1.916	1.809	1.684	1.582	1.503	1.696	1.93
5	1.692	1.596	1.733	1.791	1.865	1.836	1.449	1.506	1.956	2.047	1.907	1.952	1.805	1.667	1.53	1.525	1.88	1.76
6	1.594	1.691	1.747	1.789	1.712	1.87	1.408	1.542	1.862	2.057	1.861	1.987	1.819	1.66	1.538	1.484	2.129	1.778
7	1.914	1.861	1.837	1.814	1.797	1.881	1.831	1.569	1.75	2.033	1.934	2.026	1.871	1.658	1.563	1.425	2.443	2.046
8	1.938	2.015	1.869	1.84	1.871	1.905	1.786	1.763	1.886	2.009	1.729	2.052	1.848	1.709	1.665	1.503	2.362	2.278
9	1.984	2.021	1.78	1.892	1.421	1.877	1.734	1.788	1.447	2.033	1.475	1.945	1.651	1.738	1.428	1.581	2.356	2.169
10	1.976	1.969	1.862	1.834	1.109	1.835	1.678	1.732	1.252	1.88	1.348	1.864	1.694	1.614	1.317	1.439	2.494	2.235
11	1.955	2.176	1.693	1.815	1.257	1.662	1.75	1.645	1.425	1.656	1.597	1.576	1.762	1.62	1.29	1.398	2.478	2.429
12	1.904	2.073	1.539	1.753	1.51	1.253	1.59	1.721	1.69	1.651	1.391	1.621	1.697	1.6	1.196	1.307	2.523	2.644
13	1.967	1.908	1.38	1.634	1.643	1.404	1.735	1.573	1.513	1.944	1.365	1.585	1.756	1.519	1.328	1.23	2.745	2.695
14	1.952	1.86	1.353	1.42	1.678	1.669	1.803	1.597	1.579	1.835	1.489	1.603	1.648	1.484	1.362	1.122	2.387	2.676
15	1.986	1.894	1.426	1.33	1.95	1.796	1.815	1.643	1.894	1.72	1.846	1.574	1.607	1.378	1.668	1.334	2.396	2.446
16	1.931	1.897	1.516	1.396	1.976	1.857	1.692	1.717	1.829	1.75	1.801	1.614	1.553	1.297	1.639	1.56	2.19	2.232
17	1.973	1.891	1.207	1.393	1.917	1.912	1.867	1.801	2.014	1.805	1.924	1.606	1.556	1.265	1.602	1.619	2.137	2.05
18	1.856	1.901	1.696	1.144	1.299	1.859	1.559	1.789	1.378	1.796	1.125	1.615	1.502	1.203	1.226	1.554	2.203	1.987
19	1.739	1.812	1.875	1.622	1.096	1.412	1.224	1.45	1.336	1.491	0.929	1.253	1.813	1.397	1.311	1.129	2.477	2.296
20	1.376	1.609	1.816	1.827	1.207	1.2	1.178	1.169	1.767	1.57	0.889	0.961	1.84	1.563	1.376	1.315	2.582	2.239
21	1.959	1.534	1.791	1.792	1.118	1.178	1.367	1.205	1.6	1.821	1.459	0.727	1.853	1.587	1.421	1.253	2.728	2.425
22	1.891	1.891	1.765	1.803	1.374	1.173	1.569	1.647	1.401	1.594	1.742	1.415	1.822	1.617	1.266	1.417	2.706	2.712
23	1.725	1.905	1.827	1.807	1.27	1.247	1.606	1.68	1.475	1.594	1.69	1.732	1.858	1.578	1.374	1.286	2.844	2.716
24	1.433	1.736	1.86	1.802	1.234	1.173	1.509	1.466	1.591	1.741	1.566	1.566	1.816	1.557	1.411	1.326	2.552	2.556
All Hours	1.679	1.579	1.749	1.757	1.272	1.348	1.497	1.411	1.679	1.769	1.237	1.58	1.817	1.58	1.329	1.349	2.626	2.634

Notes: Hours which reject the null in favor of the alternative indicating long memory are outlined.

$$H_0: d \notin (0, \frac{1}{2}) \Rightarrow \text{No LR Dependence in } y_t, \text{ against } H_1: d \in (0, \frac{1}{2}) \Rightarrow \text{LR Dependence in } y_t$$

Critical values for test statistics: 5% 1%  
1.747 1.620

**Table 15:** GPH Estimate of the Order of Fractional Integration

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	0.151	0.507	0.604	0.678	0.164	0.353	0.593	0.542	0.71	0.526	0.082	0.031	0.301	0.293	0.54	0.429	0.599	0.582
2	0.166	0.325	0.686	0.804	0.273	0.311	0.741	0.762	0.679	0.841	0.119	0.085	0.289	0.372	0.712	0.527	0.498	0.512
3	0.102	0.21	0.557	0.803	0.275	0.346	0.566	0.814	0.579	0.981	0.062	-0.019	0.3	0.369	0.608	0.538	0.446	0.526
4	0.289	0.331	0.603	0.789	0.305	0.364	0.724	0.846	0.583	0.877	0.447	0.182	0.418	0.377	0.67	0.547	0.543	0.502
5	0.395	0.39	0.629	0.788	0.462	0.394	0.755	0.882	0.459	0.824	0.387	0.323	0.287	0.371	0.639	0.559	0.482	0.476
6	0.518	0.391	0.490	0.812	0.282	0.396	0.569	0.918	0.627	0.806	0.359	0.364	0.289	0.357	0.493	0.545	0.616	0.458
7	0.576	0.311	0.551	0.811	0.325	0.365	0.729	1.004	0.791	0.783	0.335	0.331	0.331	0.353	0.490	0.512	0.493	0.442
8	0.544	0.459	0.658	0.812	0.384	0.343	0.683	0.810	0.471	0.799	0.132	0.345	0.302	0.371	0.589	0.522	0.360	0.47
9	0.347	0.524	0.579	0.769	0.332	0.342	0.705	0.796	0.843	0.774	0.140	0.253	0.239	0.367	0.421	0.504	0.518	0.456
10	0.521	0.558	0.657	0.847	0.008	0.327	0.5	0.793	0.431	0.809	0.003	0.357	0.172	0.343	0.451	0.508	0.518	0.472
11	0.347	0.404	0.546	0.883	0.307	0.327	0.395	0.749	0.297	0.787	0.093	0.318	0.244	0.331	0.438	0.506	0.539	0.539
12	0.476	0.442	0.742	0.911	0.452	0.109	0.446	0.732	0.329	0.757	0.182	0.116	0.246	0.333	0.321	0.462	0.54	0.589
13	0.571	0.494	0.609	0.877	0.428	0.435	0.581	0.676	0.561	0.725	0.137	0.138	0.323	0.3	0.57	0.413	0.694	0.598
14	0.707	0.515	0.39	0.782	0.609	0.746	0.483	0.658	0.727	0.809	0.139	0.145	0.286	0.28	0.231	0.402	0.638	0.649
15	0.694	0.526	0.265	0.632	0.422	0.66	0.467	0.538	0.194	0.893	0.407	0.156	0.28	0.249	0.271	0.4	0.65	0.662
16	0.709	0.56	0.621	0.665	0.444	0.683	0.435	0.595	0.309	0.904	0.308	0.151	0.332	0.236	0.365	0.469	0.586	0.663
17	0.75	0.601	0.572	0.712	0.626	0.707	0.38	0.696	0.277	0.894	0.405	0.166	0.317	0.219	0.302	0.498	0.602	0.629
18	0.486	0.575	0.85	0.542	0.643	0.679	0.193	0.665	0.495	0.929	0.209	0.153	0.267	0.186	0.06	0.466	0.42	0.59
19	0.498	0.53	0.9	0.854	0.285	0.555	0.445	0.576	0.473	0.816	-0.011	0.086	0.292	0.279	0.521	0.178	0.493	0.509
20	0.473	0.374	0.749	0.938	0.129	0.31	0.386	0.572	0.714	0.746	0.003	-0.127	0.277	0.348	0.456	0.438	0.63	0.465
21	0.553	0.42	0.589	0.949	0.175	0.235	0.465	0.581	0.38	0.678	0.098	-0.339	0.245	0.326	0.349	0.447	0.604	0.512
22	0.612	0.383	0.657	0.875	0.512	0.101	0.352	0.649	0.663	0.426	0.236	0.175	0.297	0.329	0.34	0.451	0.63	0.602
23	0.434	0.467	0.56	0.866	0.311	0.262	0.497	0.679	0.676	0.786	0.356	0.273	0.368	0.324	0.479	0.477	0.633	0.656
24	0.318	0.419	0.748	0.892	0.327	-0.085	0.525	0.699	0.608	0.814	0.096	0.236	0.272	0.317	0.426	0.463	0.645	0.559
All Hours	0.507	0.398	0.678	0.985	0.353	0.182	0.542	0.715	0.526	0.858	0.031	0.173	0.293	0.325	0.429	0.478	0.582	0.558

Notes: Hours for which the estimate of  $d$  indicates stationarity with long memory are outlined.

$|d| > 1/2$ :  $p_t$  is non-stationary

$0 < d < 1/2$ :  $p_t$  is stationary and has long memory

$-1/2 < d < 0$ :  $p_t$  is stationary and has short memory

**Table 16:** Lo's Modified R/S Long Range Dependency Test Results of Differenced Series

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	1.709	1.039	1.829	1.219	1.613	1.087	2.309	1.668	1.304	1.42	2.162	1.222	2.081	2.531	1.978	1.49	1.39	1.855
2	1.683	1.193	1.724	1.573	2.05	1.27	2.199	1.911	1.331	0.911	2.668	1.115	2.016	1.917	1.701	1.492	1.443	1.654
3	1.558	1.297	1.964	1.553	2.24	1.258	2.553	1.918	1.301	1.049	2.612	1.33	1.985	1.847	1.755	1.495	1.474	1.529
4	1.699	1.317	1.67	1.569	1.786	1.435	1.857	2.021	1.294	1.143	1.805	1.477	1.987	1.727	1.618	1.502	1.559	1.403
5	1.537	1.383	2.192	1.547	1.433	1.455	2.262	2.091	2.343	1.325	1.987	1.612	2.054	1.782	1.813	1.549	1.556	1.455
6	1.353	1.342	2.287	1.556	2.233	1.485	2.435	2.085	1.516	1.359	2.082	1.537	2.155	1.813	1.296	1.502	1.822	1.494
7	1.071	1.296	1.529	1.455	1.244	1.518	1.251	2.189	1.79	1.315	1.593	1.47	2.195	1.93	1.524	1.851	1.573	1.59
8	1.015	1.177	1.205	1.174	1.391	1.341	1.047	1.837	1.382	1.134	2.136	1.301	2.876	1.697	1.717	1.578	2.346	1.33
9	1.21	1.123	1.559	1.177	0.921	1.292	1.651	1.698	1.606	1.077	1.433	1.212	3.081	1.843	2.098	1.676	2.737	1.597
10	1.003	1.17	1.092	1.168	1.023	1.163	1.408	1.739	1.183	1.212	2.115	1.159	2.924	2.169	1.726	1.709	2.566	1.904
11	0.961	1.326	1.221	1.231	1.056	1	1.245	1.597	1.511	1.11	2.268	1.064	2.843	2.242	1.767	1.791	2.423	1.922
12	0.991	1.446	1.346	1.218	1.324	0.872	1.896	1.427	2.216	1.073	1.941	1.031	2.715	2.042	1.355	1.781	2.204	1.674
13	0.991	1.614	1.114	1.173	1.097	1.075	1.551	1.428	2.486	1.428	1.802	1.346	1.76	2.109	1.293	1.831	0.821	1.697
14	1.36	1.675	0.847	1.16	1.285	1.318	1.397	1.64	1.892	1.898	2.095	1.635	2.041	1.961	1.763	1.62	0.914	1.43
15	1.371	1.651	1.291	1.035	1.416	1.459	1.494	1.761	1.839	1.869	1.745	1.665	1.429	1.976	1.687	1.567	0.683	1.287
16	1.751	1.691	0.903	0.939	1.568	1.407	2.097	1.719	1.338	1.841	1.802	1.516	2.001	1.82	1.347	1.414	1.133	1.041
17	1.18	1.668	1.187	1.086	1.151	1.502	2.344	1.626	1.996	1.834	1.686	1.514	1.244	1.77	2.109	1.261	1.216	0.934
18	1.282	1.576	1.468	1.039	1.295	1.416	2.201	1.573	1.212	1.802	1.305	1.48	2.175	1.844	1.206	1.383	1.408	1.024
19	1.116	1.479	1.238	1.167	1.131	1.221	1.068	1.505	2.082	1.78	1.326	1.598	1.386	1.929	1.311	1.586	0.875	1.318
20	1.038	1.372	0.974	1.234	1.278	1.119	1.029	1.495	1.866	1.862	1.493	1.71	1.869	1.976	1.586	1.878	1.407	1.635
21	1.285	1.336	1.102	1.223	1.037	1.105	1.459	1.204	1.611	1.224	1.24	1.051	2.204	2.081	1.703	1.894	1.462	1.781
22	1.09	1.431	1.222	1.255	1.167	1.093	1.199	1.142	1.832	1.046	1.693	0.937	1.628	1.931	2.092	1.859	0.961	1.609
23	0.999	1.395	2.034	1.333	1.289	1.162	1.639	1.228	1.681	1.028	1.366	1.109	2.024	2.001	1.925	1.879	1.452	1.653
24	1.407	1.352	1.959	1.43	1.772	1.133	2.014	1.467	1.668	0.981	2.268	1.308	1.756	1.902	1.922	2.186	1.281	1.558
All Hours	1.039	1.113	1.219	1.456	1.087	1.168	1.668	1.701	1.42	0.958	1.222	1.326	2.531	1.758	1.49	2.171	1.855	1.371

Notes: Hours which reject the null in favor of the alternative indicating long memory are outlined.

$$H_0: d \notin (0, \frac{1}{2}) \Rightarrow \text{No LR Dependence in } \Delta p_t, \text{ against } H_1: d \in (0, \frac{1}{2}) \Rightarrow \text{LR Dependence in } \Delta p_t$$

Critical values for test statistics: 5% 1%  
1.747 1.620

**Table 17:** GPH Estimate of the Order of Fractional Integration of Differenced Series

Hour	2008		2009		2010		2011		2012		2013		2014		2015		All Years	
	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA	RT	DA
1	-0.594	-0.481	-0.547	-0.453	-0.855	-0.533	-0.639	-0.496	-0.298	-0.484	-0.584	-0.833	-0.753	-0.745	-0.537	-0.627	-0.407	-0.417
2	-0.689	-0.576	-0.459	-0.326	-0.675	-0.522	-0.288	-0.352	-0.301	-0.181	-0.609	-0.671	-0.746	-0.665	-0.368	-0.537	-0.51	-0.491
3	-0.82	-0.486	-0.567	-0.352	-0.74	-0.509	-0.599	-0.314	-0.423	-0.029	-0.529	-0.707	-0.724	-0.654	-0.471	-0.537	-0.565	-0.464
4	-0.666	-0.43	-0.584	-0.346	-0.701	-0.525	-0.34	-0.245	-0.465	-0.146	-0.389	-0.677	-0.621	-0.649	-0.388	-0.521	-0.367	-0.479
5	-0.511	-0.434	-0.499	-0.354	-0.547	-0.518	-0.415	-0.197	-0.399	-0.195	-0.273	-0.675	-0.812	-0.659	-0.401	-0.492	-0.435	-0.497
6	-0.484	-0.425	-0.618	-0.394	-0.714	-0.532	-0.681	-0.206	-0.355	-0.191	-0.55	-0.672	-0.74	-0.668	-0.55	-0.503	-0.332	-0.537
7	-0.316	-0.467	-0.575	-0.456	-0.624	-0.558	-0.257	-0.102	-0.212	-0.379	-0.645	-0.696	-0.684	-0.669	-0.548	-0.522	-0.436	-0.598
8	-0.399	-0.274	-0.494	-0.463	-0.62	-0.597	-0.331	-0.203	-0.523	-0.214	-0.754	-0.634	-0.709	-0.647	-0.468	-0.493	-0.64	-0.53
9	-0.026	-0.286	-0.582	-0.504	-0.312	-0.633	-0.133	-0.223	-0.229	-0.242	-0.69	-0.707	-0.772	-0.647	-0.614	-0.507	-0.538	-0.535
10	-0.155	-0.324	-0.563	-0.404	-0.76	-0.657	-0.61	-0.277	-0.609	-0.21	-0.769	-0.698	-0.843	-0.676	-0.569	-0.503	-0.609	-0.551
11	-0.325	-0.399	-0.623	-0.336	-0.658	-0.649	-0.705	-0.348	-0.791	-0.229	-0.847	-0.75	-0.772	-0.691	-0.575	-0.501	-0.535	-0.55
12	-0.326	-0.438	-0.423	-0.262	-0.527	-0.695	-0.604	-0.403	-0.679	-0.417	-0.805	-0.976	-0.78	-0.694	-0.7	-0.54	-0.6	-0.523
13	-0.358	-0.456	-0.532	-0.314	-0.537	-0.482	-0.475	-0.415	-0.453	-0.308	-0.86	-0.804	-0.731	-0.732	-0.496	-0.588	-0.322	-0.462
14	-0.285	-0.463	-0.811	-0.41	-0.371	-0.256	-0.529	-0.418	-0.271	-0.208	-0.857	-0.847	-0.803	-0.753	-0.766	-0.597	-0.38	-0.388
15	-0.279	-0.465	-0.792	-0.64	-0.57	-0.286	-0.559	-0.469	-0.805	-0.138	-0.58	-0.843	-0.851	-0.789	-0.724	-0.595	-0.369	-0.367
16	-0.29	-0.438	-0.472	-0.375	-0.55	-0.267	-0.568	-0.411	-0.691	-0.12	-0.706	-0.85	-0.758	-0.811	-0.626	-0.531	-0.454	-0.375
17	-0.252	-0.404	-0.301	-0.306	-0.332	-0.264	-0.622	-0.313	-0.719	-0.121	-0.583	-0.837	-0.702	-0.84	-0.71	-0.5	-0.44	-0.409
18	-0.343	-0.419	-0.171	-0.358	-0.218	-0.284	-0.812	-0.339	-0.513	-0.091	-0.903	-0.85	-0.751	-0.877	-0.944	-0.526	-0.475	-0.437
19	-0.441	-0.462	-0.157	-0.171	-0.55	-0.296	-0.595	-0.404	-0.535	-0.201	-0.975	-0.953	-0.744	-0.77	-0.497	-0.632	-0.449	-0.503
20	-0.6	-0.658	-0.263	-0.101	-0.589	-0.348	-0.631	-0.476	-0.278	-0.334	-1.077	-0.993	-0.76	-0.687	-0.571	-0.561	-0.263	-0.478
21	-0.497	-0.616	-0.377	-0.044	-0.583	-0.409	-0.537	-0.516	-0.587	-0.345	-0.824	-0.971	-0.811	-0.709	-0.685	-0.559	-0.329	-0.45
22	-0.407	-0.626	-0.322	-0.097	-0.377	-0.365	-0.629	-0.465	-0.344	-0.8	-0.755	-0.73	-0.754	-0.719	-0.708	-0.564	-0.35	-0.388
23	-0.609	-0.544	-0.545	-0.134	-0.582	-0.46	-0.557	-0.403	-0.295	-0.389	-0.538	-0.75	-0.708	-0.719	-0.566	-0.541	-0.323	-0.358
24	-0.614	-0.59	-0.288	-0.171	-0.426	-0.625	-0.662	-0.427	-0.404	-0.269	-0.764	-0.767	-0.774	-0.713	-0.651	-0.568	-0.398	-0.437
<i>All Hours</i>	-0.481	-0.583	-0.453	-0.114	-0.533	-0.515	-0.496	-0.381	-0.484	-0.236	-0.833	-0.817	-0.745	-0.708	-0.627	-0.558	-0.417	-0.467

Notes:  $|d| > 1/2$ :  $\Delta p_t$  is non-stationary  
 $0 < d < 1/2$ :  $\Delta p_t$  is stationary and has long memory  
 $-1/2 < d < 0$ :  $\Delta p_t$  is stationary and has short memory

**Table 18:** Seasonality Regression with Dummy Variable

$$p_t = \mu_i' D_{it} + \beta t + \varepsilon_t, \quad i: \text{month of the year}$$

Parameter	RT Off Peak	RT Peak	DA Off Peak	DA Peak	Load Off Peak	Load Peak
$\mu_{Jan}$	29.594*** (0.895)	58.790** (1.946)	30.090*** (0.660)	57.436*** (1.431)	62,785.460*** (465.328)	85,989.880*** (742.346)
$\mu_{Feb}$	45.914*** (3.513)	73.493*** (7.637)	44.757*** (2.589)	73.026*** (5.616)	75,066.890*** (1,825.845)	90,923.670*** (2,912.800)
$\mu_{Mar}$	46.028*** (3.514)	61.517*** (7.637)	45.570*** (2.589)	61.758*** (5.617)	66,962.670*** (1,825.912)	78,501.710*** (2,912.907)
$\mu_{Apr}$	34.582*** (3.514)	79.574*** (7.638)	40.209*** (2.589)	74.226*** (5.617)	58,488.230*** (1,826.012)	75,662.070*** (2,913.067)
$\mu_{May}$	18.983*** (3.514)	75.714*** (7.638)	27.377*** (2.590)	76.976*** (5.617)	56,010.260*** (1,826.146)	75,520.270*** (2,913.281)
$\mu_{Jun}$	33.752*** (3.514)	159.993*** (7.639)	33.844*** (2.590)	150.137*** (5.618)	64,601.350*** (1,826.313)	102,393.900*** (2,913.547)
$\mu_{Jul}$	37.967*** (3.515)	145.160*** (7.640)	39.641*** (2.590)	151.713*** (5.619)	67,881.350*** (1,826.513)	109,306.500*** (2,913.867)
$\mu_{Aug}$	35.700*** (3.515)	110.237*** (7.641)	34.091*** (2.590)	109.517*** (5.619)	65,438.910*** (1,826.747)	104,167.000*** (2,914.240)
$\mu_{Sep}$	34.394*** (3.516)	96.943*** (7.642)	34.635*** (2.591)	86.723*** (5.620)	60,706.890*** (1,827.014)	92,210.380*** (2,914.666)
$\mu_{Oct}$	30.220*** (3.516)	53.279*** (7.643)	32.176*** (2.591)	55.244*** (5.621)	55,850.030*** (1,827.314)	74,951.930*** (2,915.144)
$\mu_{Nov}$	37.923*** (3.517)	80.905*** (7.644)	37.856*** (2.592)	74.647*** (5.622)	63,110.600*** (1,827.648)	83,737.870*** (2,915.676)
$\mu_{Dec}$	37.522*** (3.518)	77.837*** (7.646)	37.604*** (2.592)	75.982*** (5.623)	70,317.040*** (1,828.014)	91,865.890*** (2,916.261)
$\beta$	-0.002*** (0.001)	-0.003*** (0.001)	-0.002*** (0.0004)	-0.003*** (0.001)	4.571*** (0.260)	6.379*** (0.415)
Observations	2,922	2,922	2,922	2,922	2,922	2,922
R <sup>2</sup>	0.675	0.674	0.802	0.783	0.980	0.973
Adjusted R <sup>2</sup>	0.673	0.672	0.801	0.782	0.980	0.973
Residual Std. Error (df = 3153)	19.243	41.828	14.181	30.762	10,000.350	15,953.730
F Statistic (df = 13; 3153)	464.383***	461.659***	907.333***	808.578***	10,927.830***	8,115.955***

Notes: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

df: degrees of freedom, k: number of parameters, T: number of observations

R<sup>2</sup>:  $1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y})^2)$

Adjusted R<sup>2</sup>:  $1 - [(1 - R^2)(T - 1) / (T - k)]$

Residual Std. Error:  $\sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / (T - k)}$

F Statistic:  $\frac{R^2 / (k-1)}{(1-R^2) / (T-k)}$

**Table 19: Seasonality Regression with Sinusoids**

$$p_t = c_1 + c_2 \cos\left(\frac{2\pi t}{360}\right) + c_3 \sin\left(\frac{2\pi t}{360}\right) + c_4 \cos\left(\frac{2\pi t}{180}\right) + c_5 \sin\left(\frac{2\pi t}{180}\right) + c_6 \cos\left(\frac{2\pi t}{7}\right) + c_7 \sin\left(\frac{2\pi t}{7}\right) + \beta t + \varepsilon_t$$

Parameter	RT Off Peak	RT Peak	DA Off Peak	DA Peak	Load Off Peak	Load Peak
$c_2$	-4.542*** (0.459)	17.038*** (0.946)	-3.844*** (0.329)	15.021*** (0.730)	-3,065.010*** (193.587)	7,275.431*** (269.702)
$c_3$	-5.478*** (0.457)	10.929*** (0.942)	-5.382*** (0.327)	9.080*** (0.727)	-3,321.445*** (192.695)	8,739.797*** (268.459)
$c_4$	-1.429*** (0.456)	3.518*** (0.941)	-0.981*** (0.327)	3.399*** (0.726)	-891.450*** (192.414)	227.831 (268.067)
$c_5$	2.609*** (0.458)	8.796*** (0.945)	2.326*** (0.328)	10.705*** (0.729)	7,859.736*** (193.299)	13,188.420*** (269.300)
$c_6$	-0.008 (0.457)	-7.099*** (0.943)	-0.514 (0.328)	-6.722*** (0.727)	-1,309.701*** (192.789)	-5,751.097*** (268.589)
$c_1$	-1.662*** (0.457)	-0.723 (0.942)	-0.819** (0.327)	-1.977*** (0.727)	-1,685.224*** (192.720)	-1,991.182*** (268.494)
$\beta$	-0.003*** (0.0004)	-0.011*** (0.001)	-0.003*** (0.0003)	-0.010*** (0.001)	4.058*** (0.150)	5.007*** (0.208)
$c_1$	32.746*** (0.647)	72.002*** (1.336)	33.322*** (0.464)	69.054*** (1.030)	63,132.720*** (273.269)	85,929.580*** (380.713)
Observations	3,166	3,166	3,166	3,166	3,166	3,166
R <sup>2</sup>	0.106	0.224	0.162	0.290	0.502	0.620
Adjusted R <sup>2</sup>	0.104	0.222	0.160	0.288	0.501	0.619
Residual Std. Error (df = 3158)	18.171	37.496	13.032	28.917	7,669.094	10,684.440
F Statistic (df = 7; 3158)	53.217***	129.909***	87.147***	184.271***	454.470***	736.810***

Notes: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

df: degrees of freedom, k: number of parameters, T: number of observations

R<sup>2</sup>:  $1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y}))$

Adjusted R<sup>2</sup>:  $1 - [(1 - R^2)(T - 1) / (T - k)]$

Residual Std. Error:  $\sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / (T - k)}$

F Statistic:  $\frac{R^2 / (k-1)}{(1-R^2) / (T-k)}$

**Table 20: RT Off Peak MRS**

	Estimate:	Std. Error	t-value	p
<b>Regime 1:</b>				
$\alpha_0$	0.717	0.112	-6.432	0.000
$(1 - \beta_0)$	0.629	0.016	39.566	0.000
Residual Standard Error	5.324			
R <sup>2</sup>	.476			
<b>Regime 2:</b>				
$\alpha_1$	11.408	2.894	3.941	0.000
$(1 - \beta_1)$	.317	.054	5.867	0.000
Residual Standard Error	45.987			
R <sup>2</sup>	.097			
$P(S_t = 0   S_{t-1} = 0)$	0.962			
$P(S_t = 1   S_{t-1} = 1)$	0.672			
<b>Full Model:</b>				
Log Likelihood	-10884.27			
AIC	21776.54			
BIC	21832.98			

Notes: df: degrees of freedom, L: log likelihood, k: number of parameters, T: number of observations

$$\text{Residual Std. Error: } \sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / \text{df}}$$

$$R^2: 1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y})^2)$$

$$\text{AIC: } 2k - 2\ln(L)$$

$$\text{BIC: } 2(T)k - 2\ln(L)$$

**Table 21: RT Peak MRS**

	Estimate:	Std. Error	t-value	p
<b>Regime 0:</b>				
$\alpha_0$	-2.541	0.292	-8.6882	0.000
$(1 - \beta_0)$	0.767	0.019	40.793	0.000
Residual Standard Error	8.773			
R <sup>2</sup>	0.701			
<b>Regime 1:</b>				
$\alpha_1$	14.795	1.990	7.437	0.000
$(1 - \beta_1)$	0.233	0.032	7.268	0.000
Residual Standard Error	54.711			
R <sup>2</sup>	0.054			
$P(S_t = 0   S_{t-1} = 0)$	0.908			
$P(S_t = 1   S_{t-1} = 1)$	0.816			
<b>Full Model:</b>				
Log Likelihood	-13921.39			
AIC	27850.78			
BIC	27907.22			

Notes: df: degrees of freedom, L: log likelihood, k: number of parameters, T: number of observations

$$\text{Residual Std. Error: } \sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / \text{df}}$$

$$R^2: 1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y})^2)$$

$$\text{AIC: } 2k - 2\ln(L)$$

$$\text{BIC: } 2(T)k - 2\ln(L)$$

**Table 22: DA Off Peak MRS**

	Estimate:	Std. Error	t-value	p
<b>Regime 1:</b>				
$\alpha_0$	-0.281	0.059	-4.737	0.000
$(1 - \beta_0)$	0.869	0.0108	80.435	0.000
Residual Standard Error	2.909			
R <sup>2</sup>	0.808			
<b>Regime 2:</b>				
$\alpha_1$	9.685	2.398	4.040	0.000
$(1 - \beta_1)$	0.574	0.065	8.804	0.000
Residual Standard Error	29.639			
R <sup>2</sup>	0.332			
$P(S_t = 0   S_{t-1} = 0)$	0.983			
$P(S_t = 1   S_{t-1} = 1)$	0.780			
<b>Full Model:</b>				
Log Likelihood	-8615.99			
AIC	17239.99			
BIC	17296.43			

Notes: df: degrees of freedom, L: log likelihood, k: number of parameters, T: number of observations

$$\text{Residual Std. Error: } \sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / df}$$

$$R^2: 1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y}))$$

$$\text{AIC: } 2k - 2\ln(L)$$

**Table 23: DA Peak MRS**

	Estimate:	Std. Error	t-value	p
<b>Regime 1:</b>				
$\alpha_0$	-0.914	0.155	-5.897	0.000
$(1 - \beta_0)$	0.856	0.010	85.630	0.000
Residual Standard Error	6.744			
Multiple R-Squared	0.787			
<b>Regime 2:</b>				
$\alpha_1$	9.685	2.398	4.040	0.000
$(1 - \beta_1)$	0.574	0.065	8.804	0.000
Residual Standard Error	29.639			
Multiple R-Squared	0.332			
$P(S_t = 0   S_{t-1} = 0)$	0.966			
$P(S_t = 1   S_{t-1} = 1)$	0.824			
<b>Full Model:</b>				
Log Likelihood	-11798.68			
AIC	23605.37			
BIC	23661.81			

Notes: df: degrees of freedom, L: log likelihood, k: number of parameters, T: number of observations

$$\text{Residual Std. Error: } \sqrt{\sum_{t=1}^T \hat{\varepsilon}_t^2 / df}$$

$$R^2: 1 - (\sum_{t=1}^T \hat{\varepsilon}_t^2 / \sum_{t=1}^T (y_t - \bar{y}))$$

$$\text{AIC: } 2k - 2\ln(L)$$

$$\text{BIC: } 2(T)k - 2\ln(L)$$

**Table 24:** Non-Switching AR(1)
$$p_t = g(\sim) + y_t$$

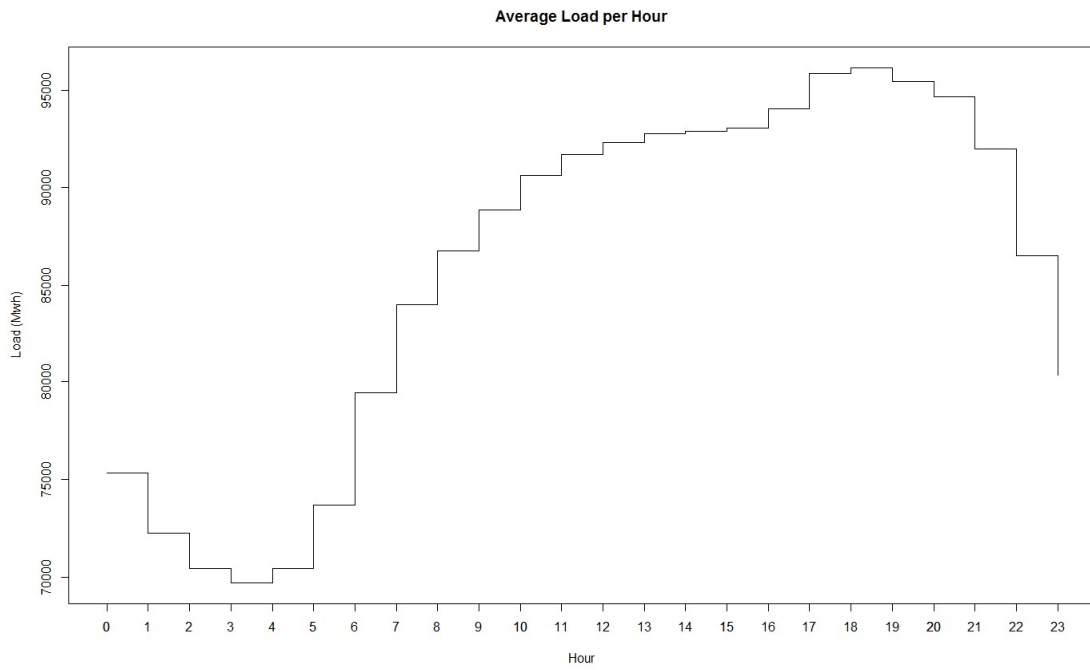
$$g(\sim) = c_1 + c_2 \cos\left(\frac{2\pi t}{360}\right) + c_3 \sin\left(\frac{2\pi t}{360}\right) + c_4 \cos\left(\frac{2\pi t}{180}\right) + c_5 \sin\left(\frac{2\pi t}{180}\right) + c_6 \cos\left(\frac{2\pi t}{7}\right) + c_7 \sin\left(\frac{2\pi t}{7}\right) + \delta t$$

$$y_t = \alpha + (1 - \beta)y_{t-1} + \epsilon_t$$

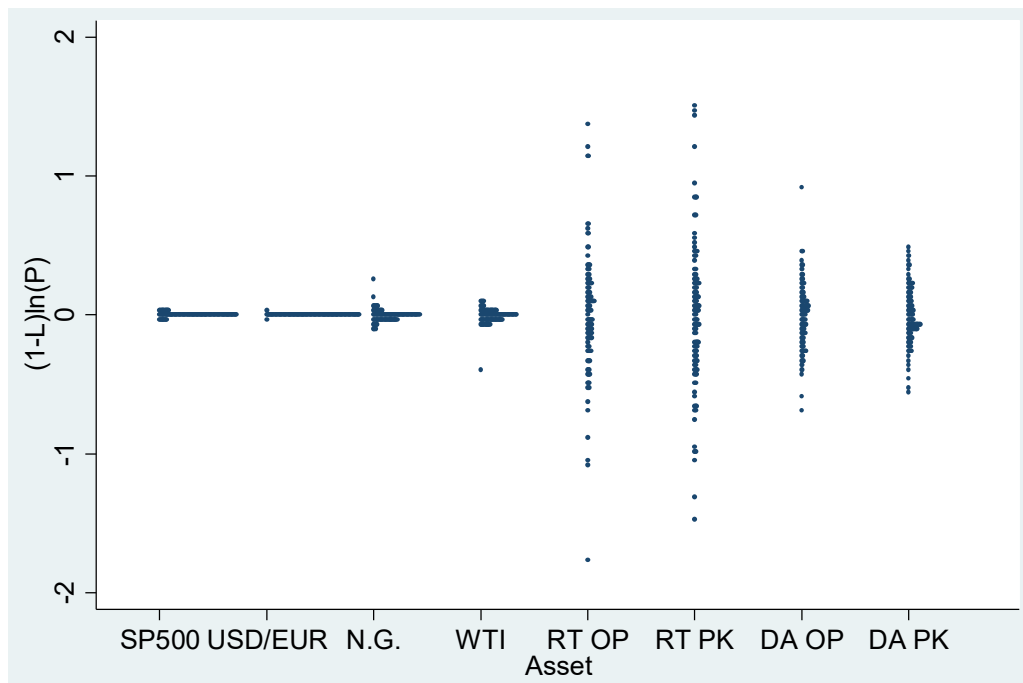
Parameter	RT Off Peak	RT Peak	DA Off Peak	DA Peak
$(1 - \beta)$	0.439*** (0.016)	0.406*** (0.016)	0.738*** (0.012)	0.757*** (0.012)
$\alpha$	0.029 (0.289)	0.017 (0.611)	0.018 (0.157)	0.013 (0.337)
Observations	3,150	3,150	3,150	3,150
Log Likelihood	-13,248.440	-15,606.310	-11,319.120	-13,732.040
AIC	26,500.890	31,216.610	22,642.250	27,468.080

Notes: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01  
L: log likelihood, k: number of parameters  
AIC:  $2k - 2\ln(L)$

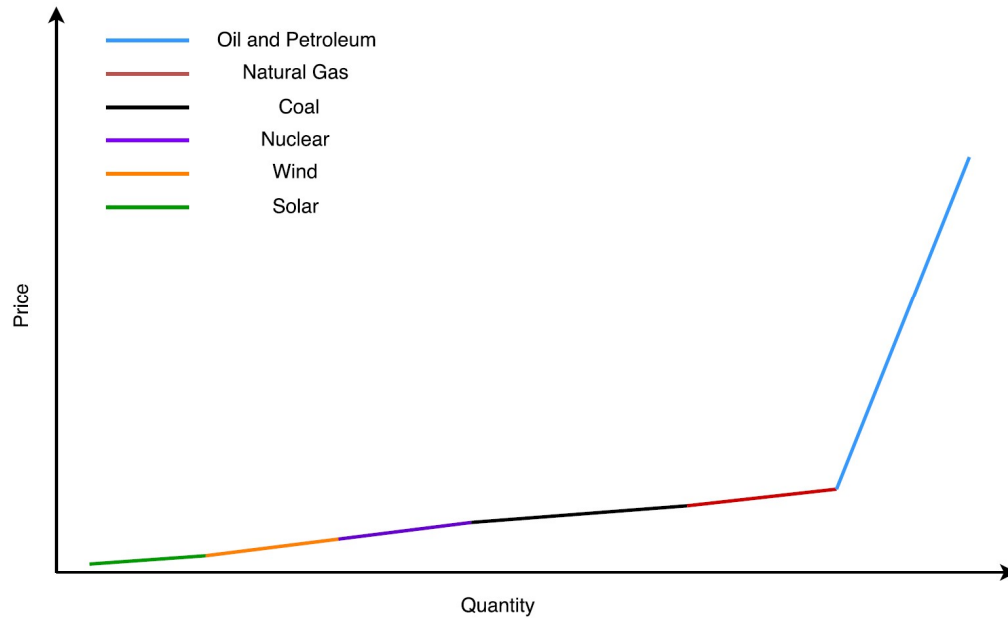
**Figure 1: Average Hourly Loads**



**Figure 2: Distributions of Financial and Electricity Returns**

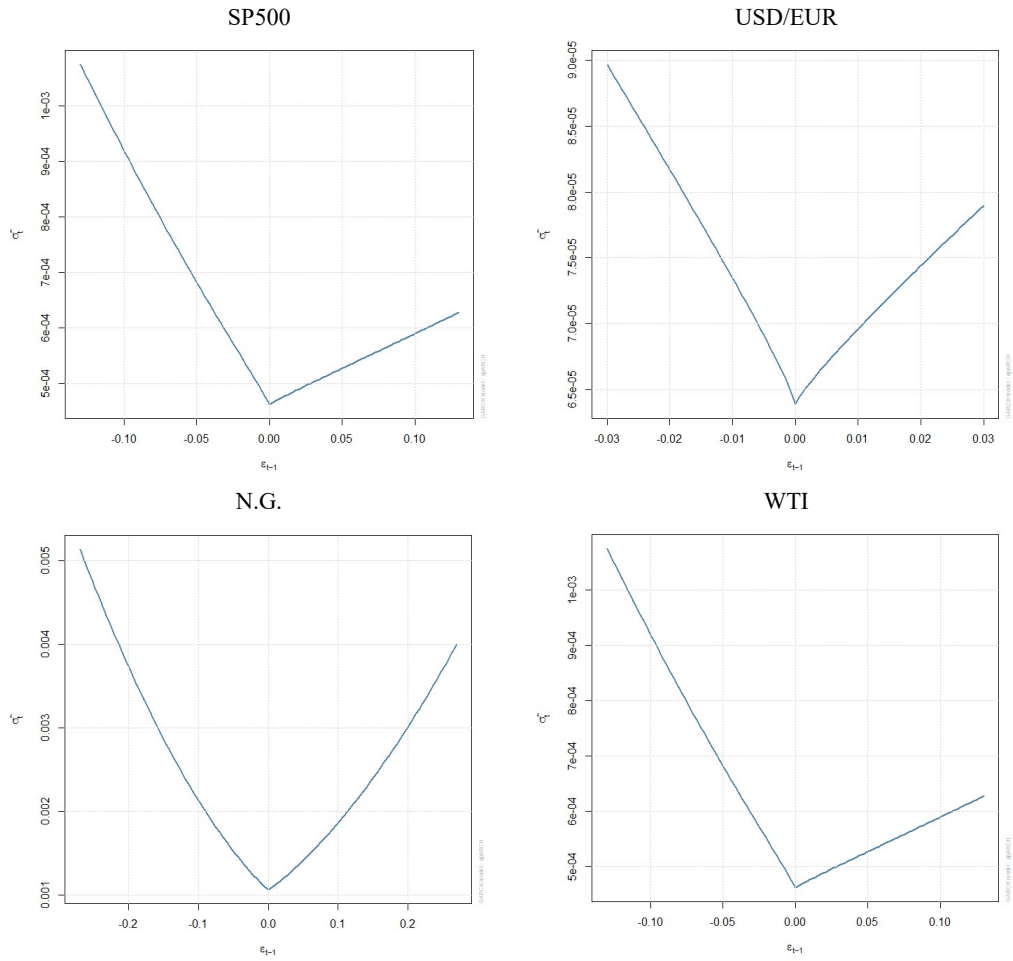


**Figure 3:** Hypothetical PJM Generation Stack

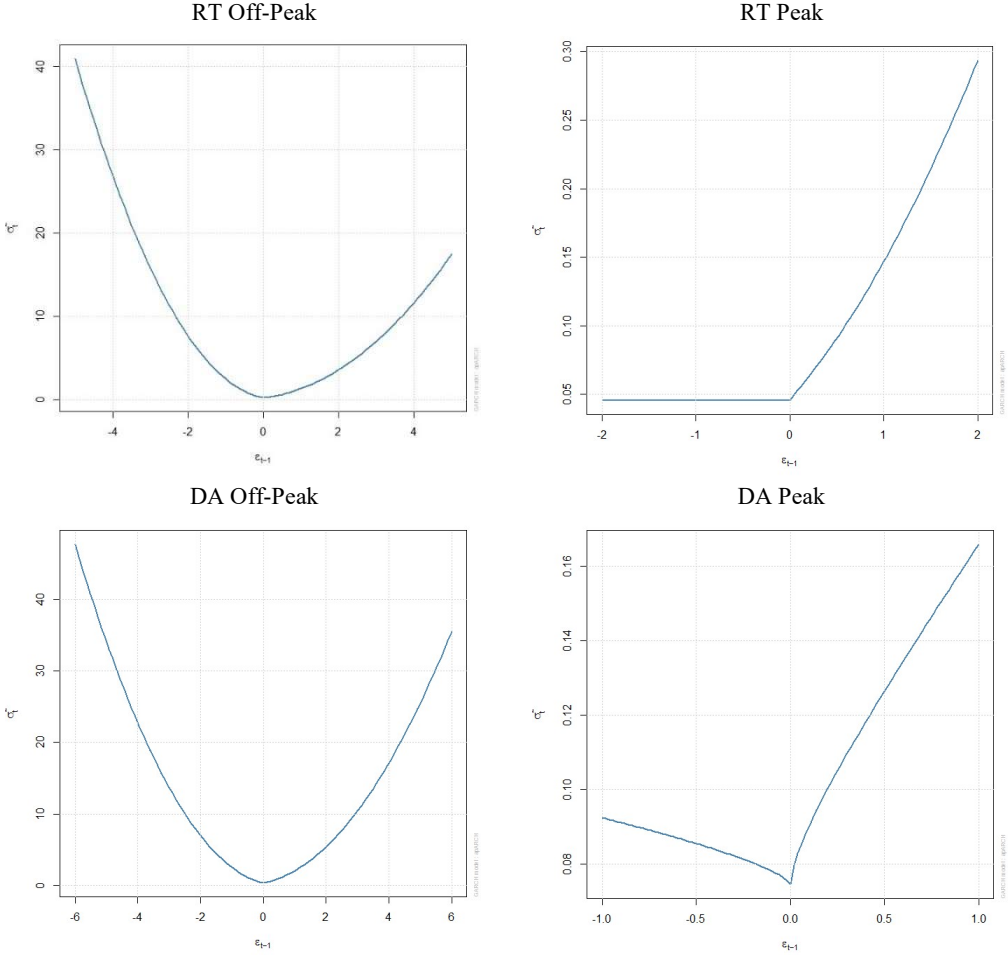


Source: Recreated from SNL Financial / S&P Global Market Intelligence, PJM Generation Supply Curve

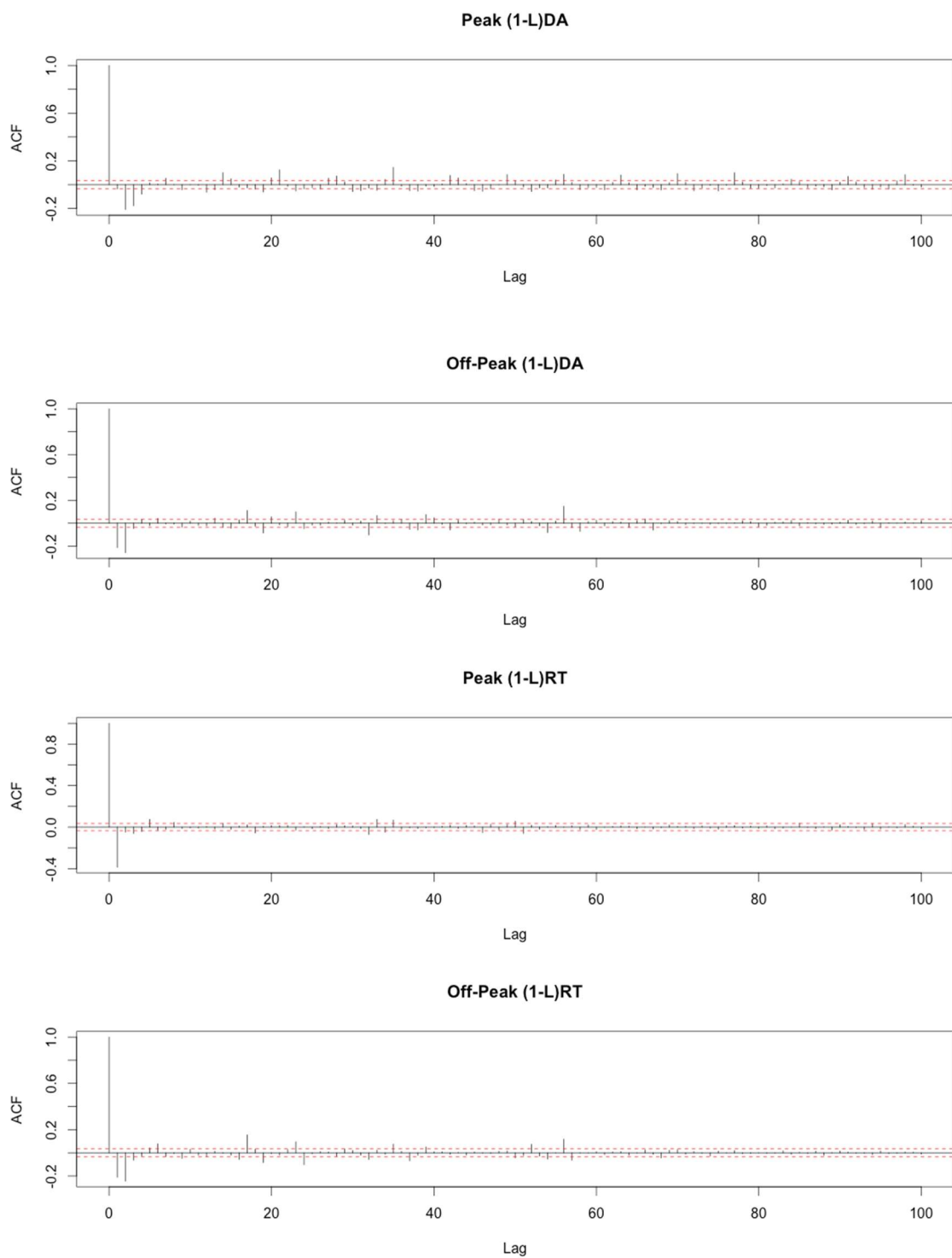
**Figure 4:** News Impact Curves for Financial Returns  $\sigma_t^2$  as a discontinuous function of  $\varepsilon_{t-1}$



**Figure 5: News Impact Curves for Electricity Returns**  
 $\sigma_t^2$  as a discontinuous function of  $\varepsilon_{t-1}$

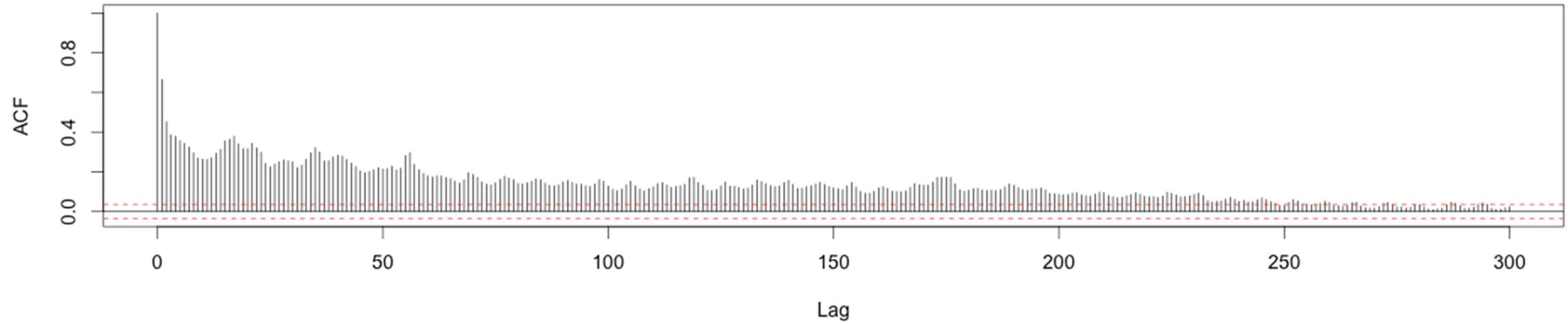


**Figure 6:** Autocorrelation Function Plots of First Difference of Representative Price Series 3AM and 6PM in RT and DA

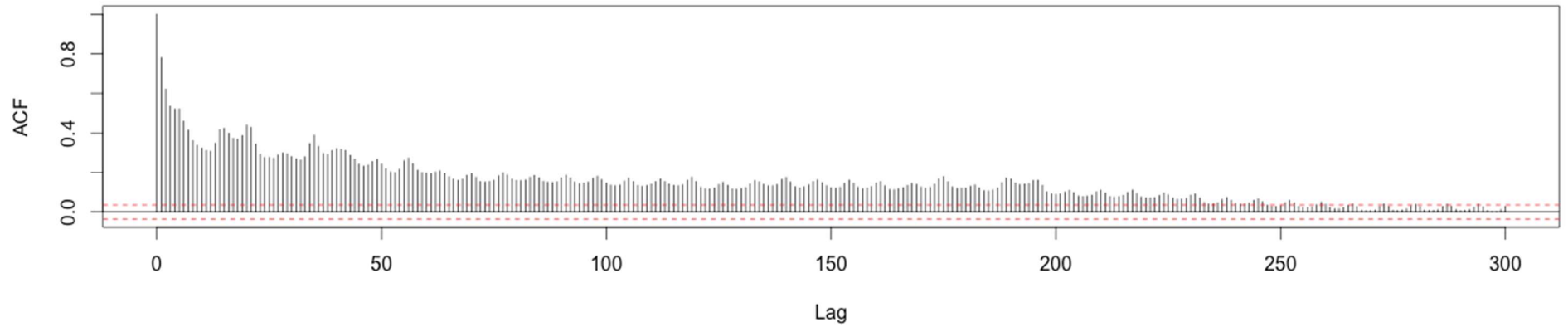


**Figure 7:** Long Run Autocorrelations Function Plots of the Average Daily Prices in RT and DA

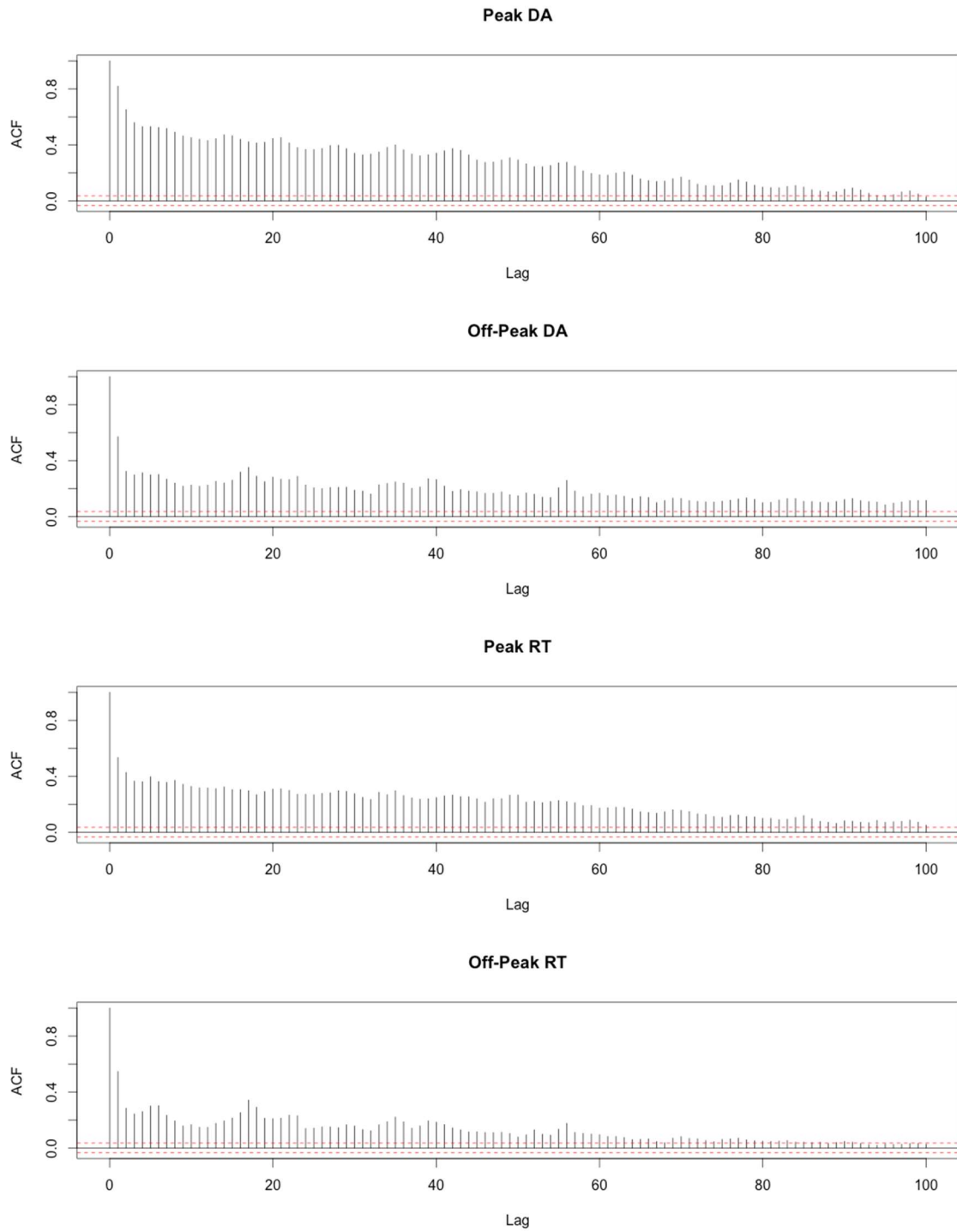
**ACF of Average Daily RT Prices**



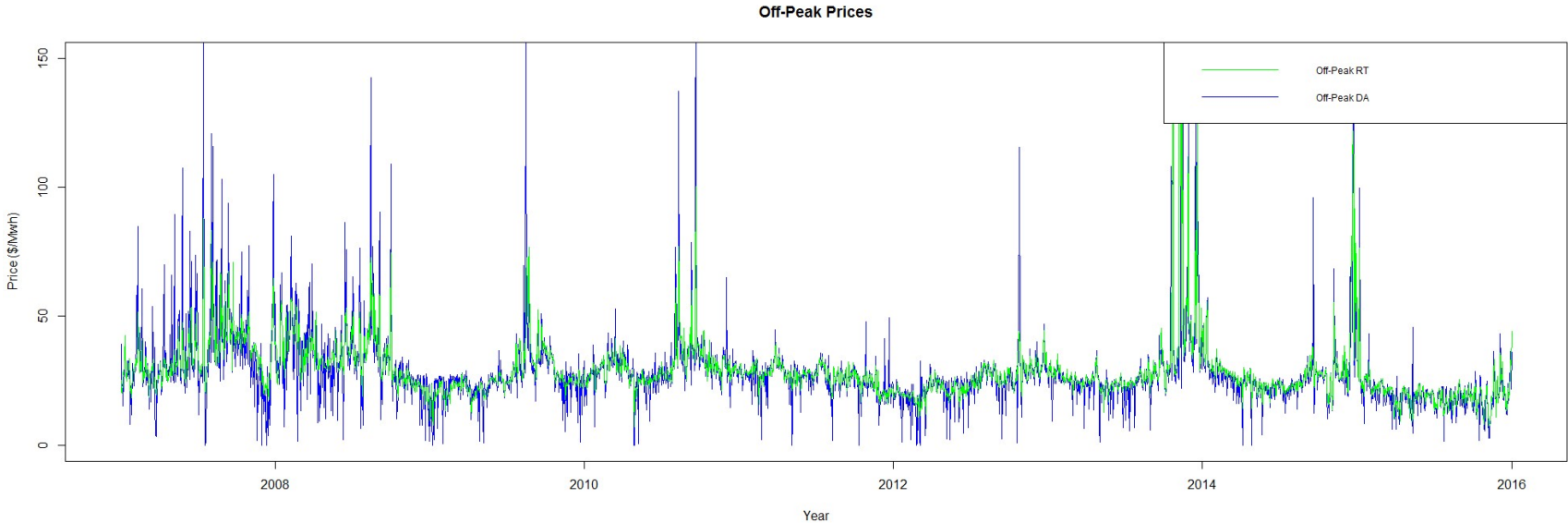
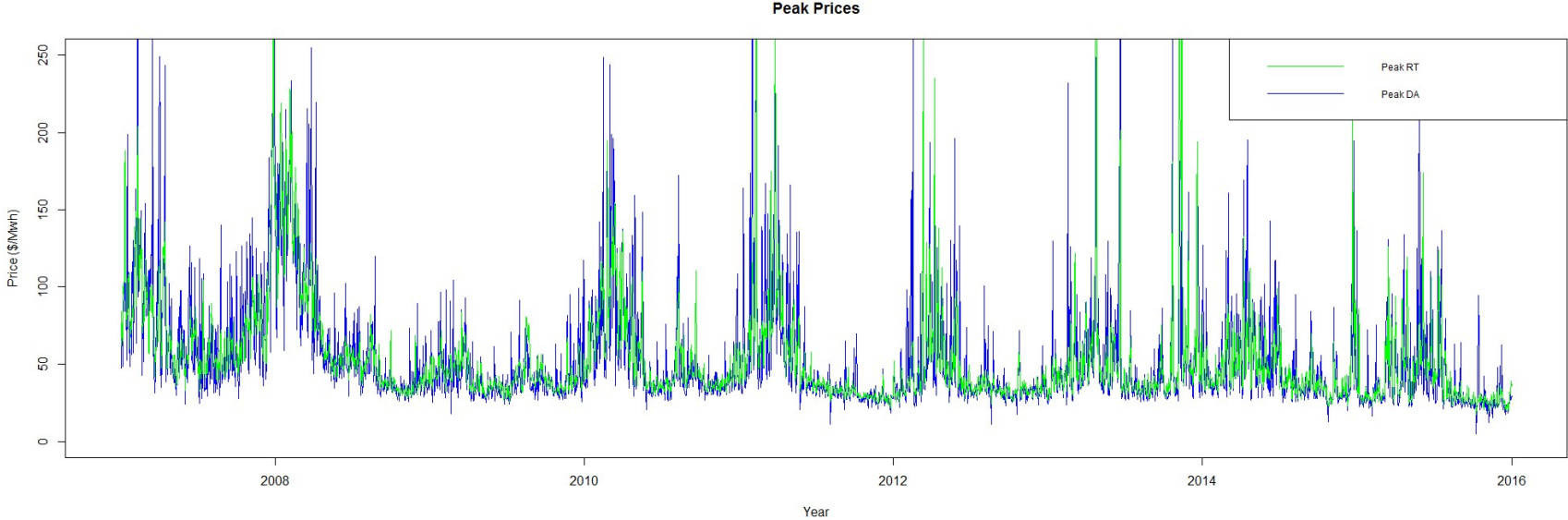
**ACF of Average Daily DA Prices**



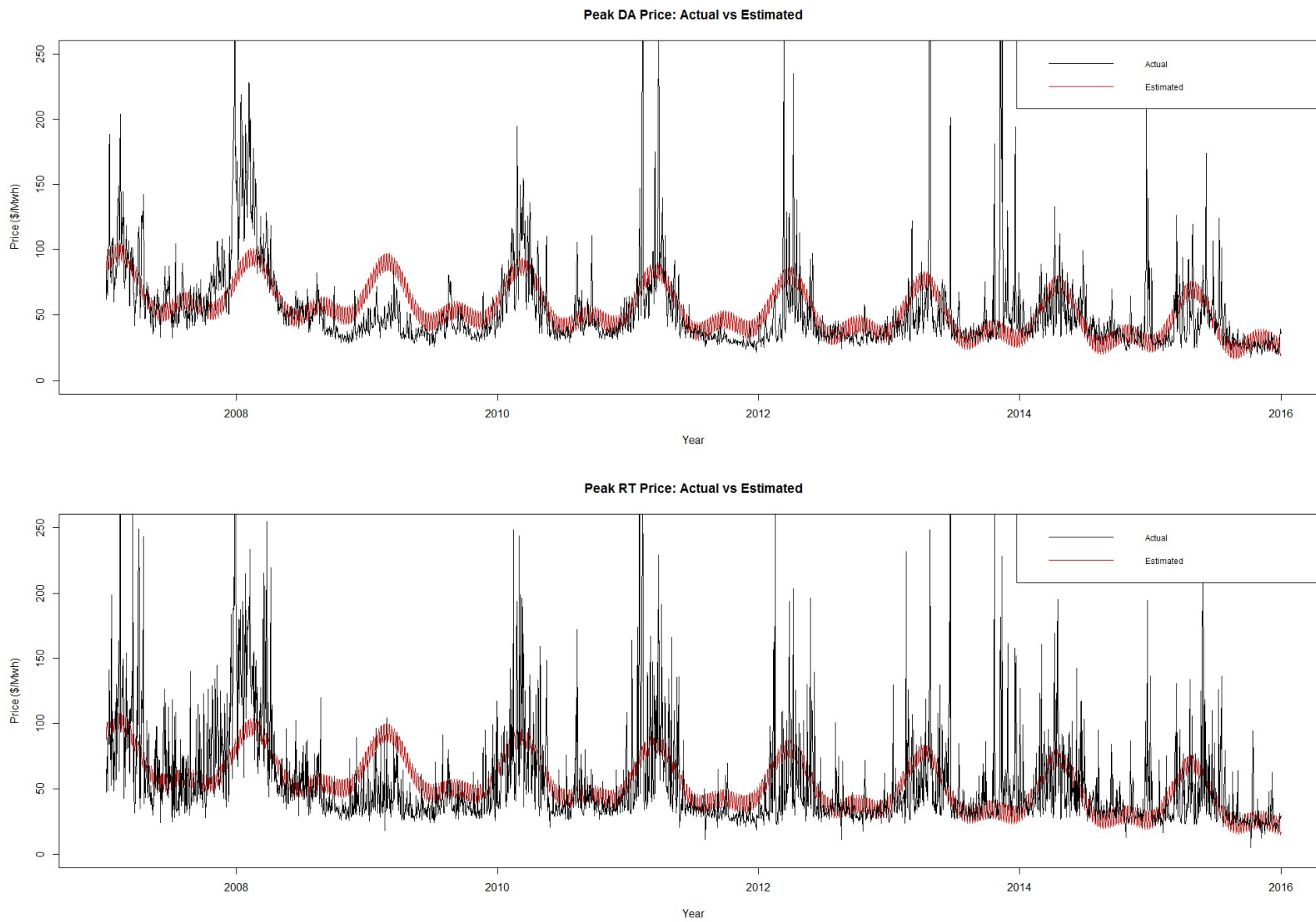
**Figure 8:** Autocorrelation Functions Plots of Representative Price Series  
3AM and 6PM in RT and DA



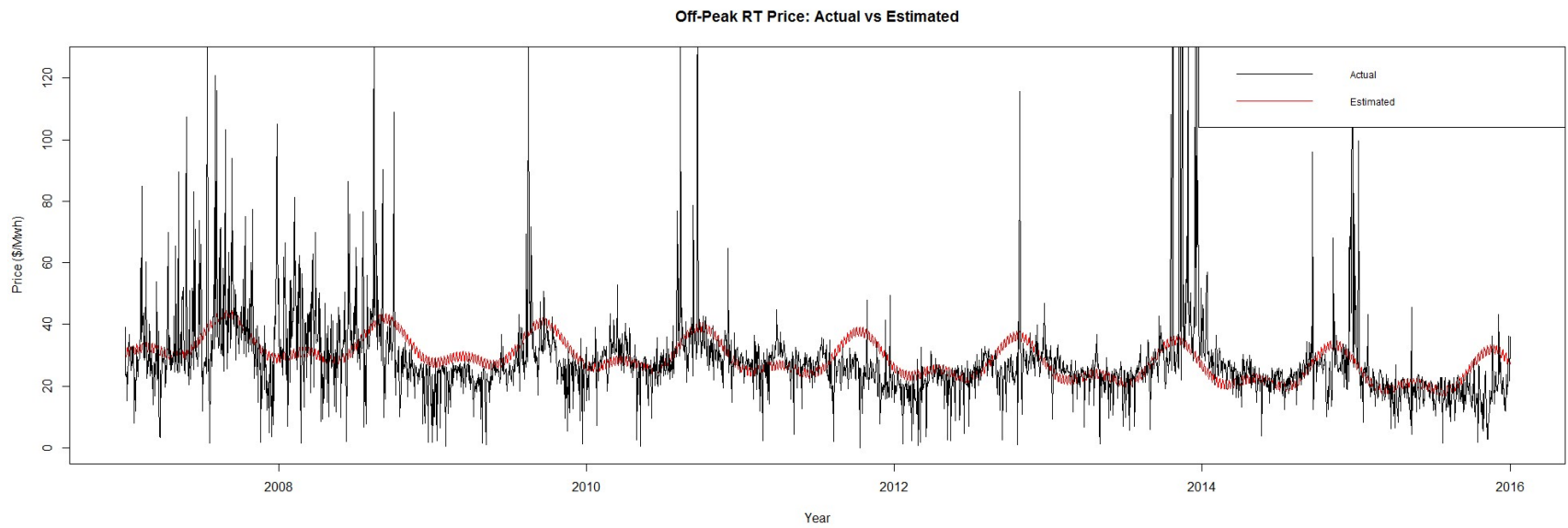
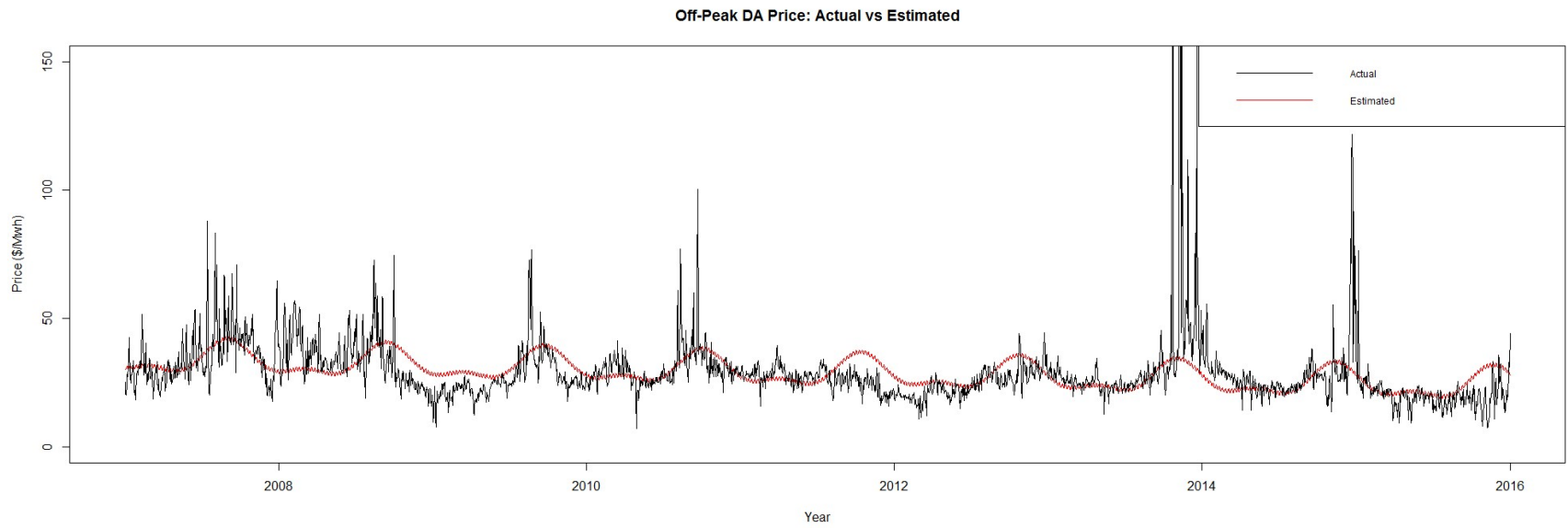
**Figure 9: Electricity Price Series**



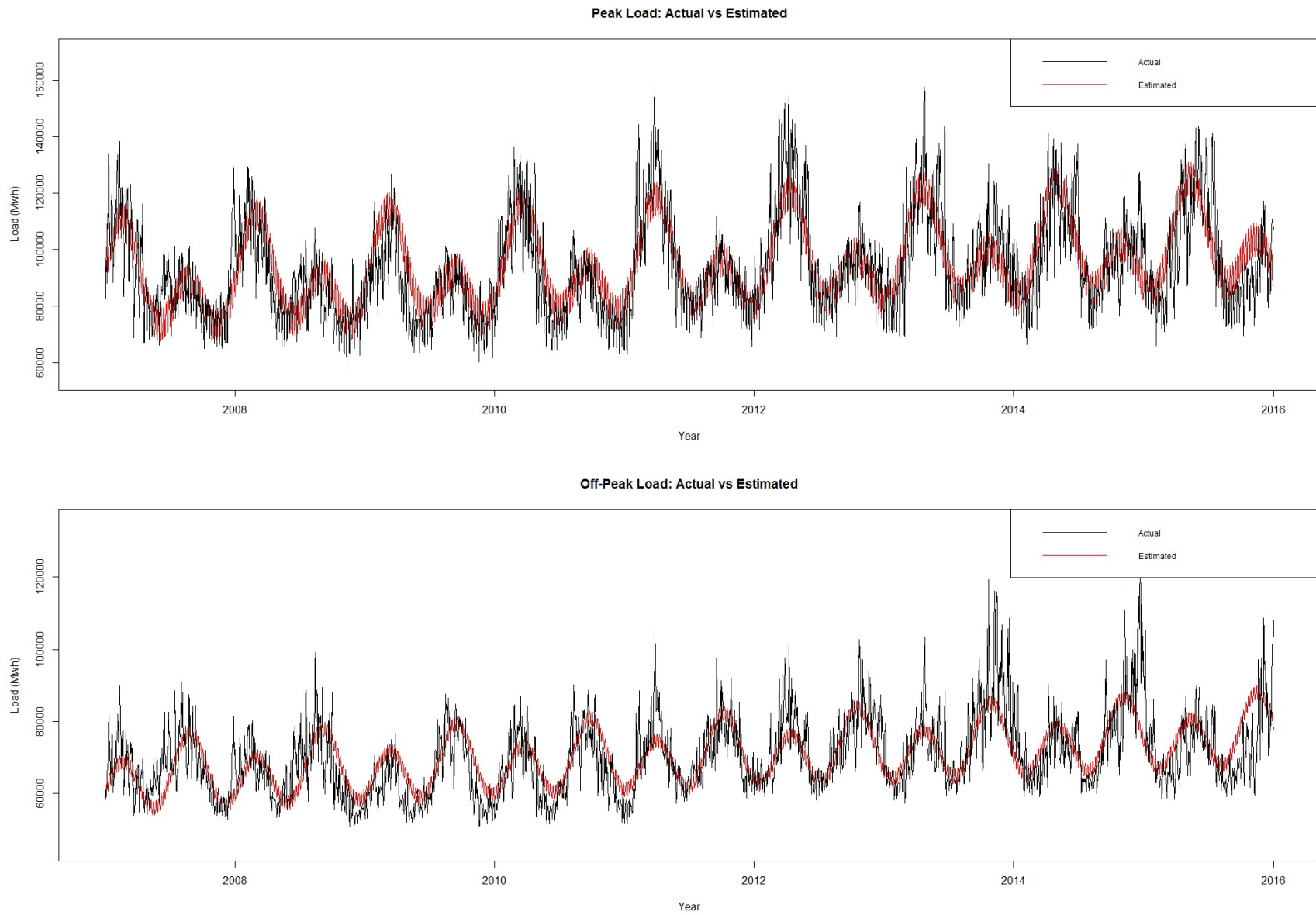
**Figure 10: Peak Hour Fourier Seasonality**



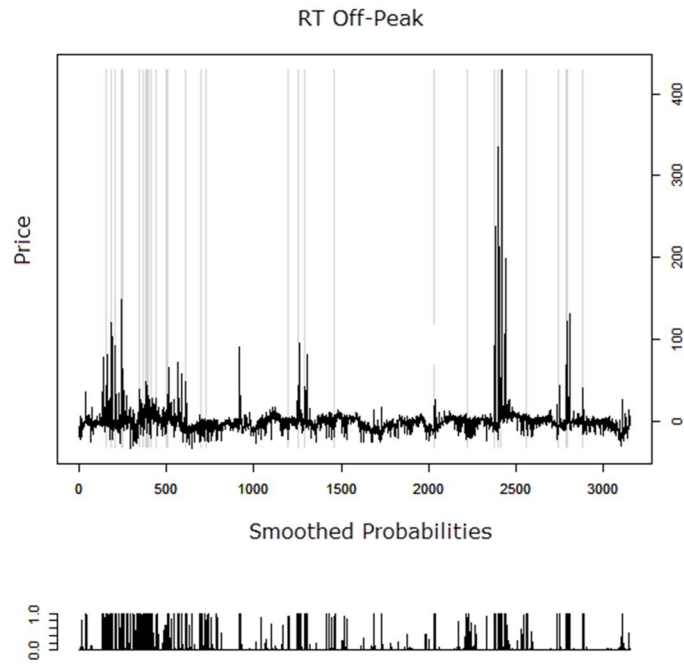
**Figure 11: Off-Peak Fourier Seasonality**



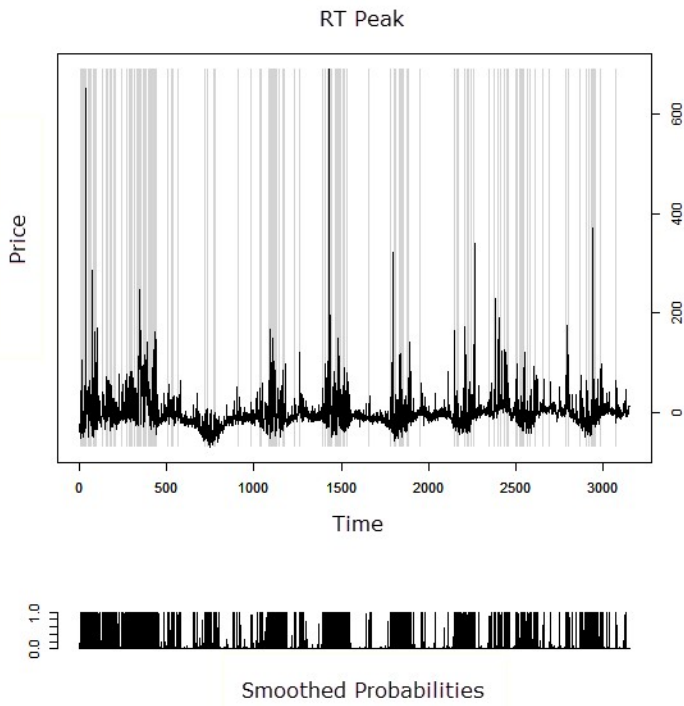
**Figure 12: Load Fourier Seasonality**



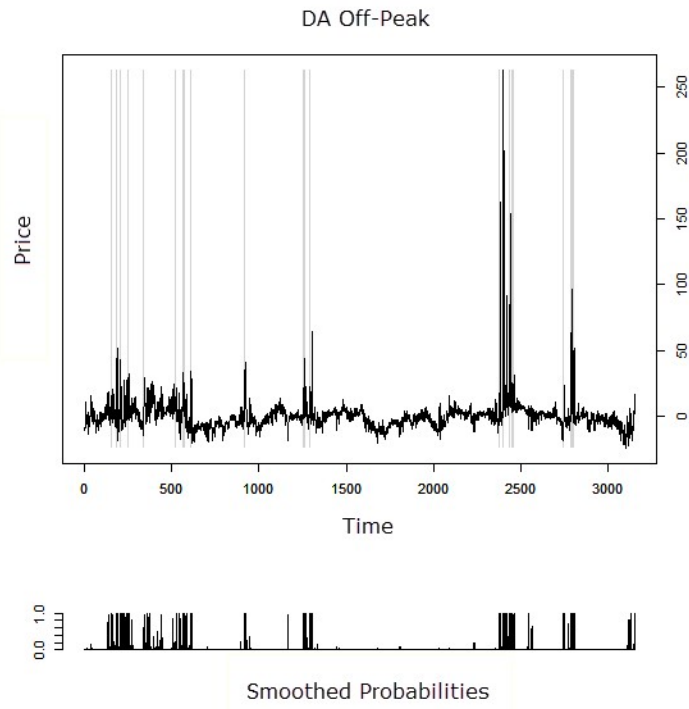
**Figure 13: RT Off Peak Regime Switching Probabilities**



**Figure 14: RT Peak Regime Switching Probabilities**



**Figure 15: DA Off Peak Regime Switching Probailities**



**Figure 16: DA Peak Regime Switching Probailities**

