UNIVERSITY OF CALGARY

Energy Commodity Volatility Modelling using GARCH

by

Olga Efimova

A THESIS

SUBMITTED TO THE FACULTY OF GRADUATE STUDIES

IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE

DEGREE OF MASTER OF ARTS

DEPARTMENT OF ECONOMICS

CALGARY, ALBERTA

JANUARY, 2013

Abstract

This thesis investigates the empirical properties of oil, natural gas and electricity price volatilities using a range of univariate and multivariate GARCH models and daily data from U.S. wholesale markets for the period from 2000 to 2012. The key contribution to the literature is the estimation of trivariate BEKK, CCC and DCC models that allow us to observe spillovers and interactions among energy markets. We evaluate the performance of each model with a range of diagnostic and forecast performance tests, and also include graphs for short- and long-term forecasts.

Acknowledgements

I would like to thank my supervisor Dr. Apostolos Serletis for expert guidance, the department of Economics at the University of Calgary for instruction and generous funding, and my family and friends for supporting me every step of the thesis journey.

Contents

List of Tables

List of Figures

1 Introduction

Energy markets are taking an increasingly pre-eminent place in the global economy. Energy is not only one of the most important consumer commodities, but also a major input for almost every industry. The U.S. Energy Information Administration's long-term projections suggest that total energy consumption and energy commodity prices will grow steadily over the next several decades. The price of natural gas, which has reached its minimum in 2012, is expected to increase by 60%, the price of electricity by 7%, and the price of oil by 62%, reaching $145 per barrel by 2035 (EIA, 2012). Oil, natural gas, and increasingly, electricity, are traded in competitive wholesale markets in North America and around the world.

The dynamics of energy markets are similar to those of financial markets. Prices fluctuate significantly by day and even hour. Furthermore, periods of relative tranquility are interspersed with those of extreme volatility; the latter can last for days or weeks and may be triggered by events in other energy markets, derivative markets or the macroeconomy, and also by significant supply or demand shocks. Factors like political instability in the oil-producing Middle East region and expansion of wind electricity generation are likely to keep energy markets volatile into the foreseeable future.

Energy commodity price volatility is of great concern to oil, gas and electricity market participants, and carries direct implications for derivative trading. At times when consumers become affected by swings in oil or electricity prices, volatility becomes a contentious political issue. Therefore, being able to accurately forecast energy commodity volatility is of great importance not only for derivative traders or large energy market players exercising hedging strategies, but also for policymakers. From the statistical point of view, changing price volatility is reflected as heteroskedasticity in the data; failing to account for it results in biased standard errors, invalidating any inference and tests of statistical significance in empirical studies on energy markets.

In 1982, Robert Engle introduced the Autoregressive Conditional Heteroskedasticity (ARCH) model as a way to not just correct for, but directly model heteroskedasticity in time series data by augmenting a time series regression with a separate conditional variance equation. Tim Bollerslev (1986) developed this model into GARCH (Generalized Autoregressive Conditional Heteroskedasticity), which has become an invaluable tool for researchers using heteroskedastic time series data in finance, macroeconomics and energy economics. The 1990's and early 2000's saw several empirical studies using univariate GARCH models with energy data - the most notable include Namit (1998), Morana (2001), Lin and Tamvakis (2001). More recently, the standard of practice has shifted towards multivariate GARCH, a class of vector autoregressive models that was first proposed by Bollerslev et. al. in 1988 and became much more widely used after the popularization of its simplified BEKK variant, proposed by Engle and Kroner in 1995.

Multivariate GARCH models allow the researcher to study several time series as a system, allowing for rich dynamics among series and their volatilities. All GARCH models require the use of maximum likelihood estimation (MLE) because the ordinary least squares (OLS) method is built upon the assumption of constant variance, so it cannot accommodate volatility modelling. The combination of a larger number of parameters and the need to use MLE makes multivariate GARCH models very demanding to estimate. As a result, univariate GARCH can often deliver more precise coefficient estimates and better forecasts (Andersen & Bollerslev, 1998; Wang & Wu, 2012).

This thesis contributes to the literature on energy price volatility modelling in several ways. First, it fills the gap in univariate GARCH modelling of energy commodity volatility - there have been very few such studies published since 2005. We adopt the univariate GARCH approach to model oil, natural gas and electricity price volatilities (Chapters 2, 3, 4, respectively), using daily U.S. data that is as recent as July 2012. Secondly, we construct and estimate a range of tri-variate GARCH models that explore in great detail the interdependence of wholesale oil, natural gas and electricity market prices and volatilities using recent U.S. data. These models are presented in Chapter 5. Existing studies have explored the relationships among several electricity markets (Goto & Karolyi, 2004; Worthington et.

al, 2005), between oil and natural gas markets (Ewing et. al., 2002; Serletis and Shahmoradi, 2006), and between oil markets and financial or macroeconomic indicators (Lee et al., 1995; Sadorsky, 2011, 2012; Elder & Serletis, 2010; Rahman & Serletis, 2012). However, no study has used multivariate GARCH to model oil, natural gas and electricity markets as a system, to the best of our knowledge. As an additional contribution to the literature, the use of both univariate and multivariate models over the same data set allows us to compare the performance of these models, including forecasting performance.

2 Oil Price Volatility – Univariate Models.

2.1 Introduction

Oil price volatility is of great interest to energy and financial market participants, as well as policymakers. In fact, the oil price and its volatility are widely used as leading macroeconomic indicators. At the same time, both are notoriously hard to forecast due to the complexity of the factors affecting outcomes in the oil markets. Univariate GARCH models have been used to model oil price volatility since the early 1990s, and have become standard practice. Despite the explosion of new types of GARCH models, including multivariate GARCH (Bollerslev et. al., 1988), fractionally integrated GARCH (Baillie et. al., 1996), nonparametric GARCH (Bühlmann & McNeil, 2002), and multiplicative component GARCH (Engle & Sokalska, 2012), simple models of the GARCH(1,1) type remain very useful because they converge much faster to a local maximum in quasi-maximum likelihood estimation, while delivering forecasting performance that is not obviously inferior to multivariate models (Wang & Wu, 2012). Despite this strength of univariate GARCH models, they have been largely displaced by more complicated ones from the literature on volatiltiy modelling, with a few exceptions (Gogas & Serletis, 2009; Chang, 2009). We fill this literature gap by estimating two GARCH(1,1) models of oil price volatility using very recent U.S. data. GARCH(1,1) is also a logical first step in our study of energy commodity volatility, a foundation block to precede multivariate models.

2.2 The Data

We use daily price data for wholesale crude oil at Cushing, Oklahoma, the main wholesale trading hub for West Texas Intermediate, the latter being light low-sulphur oil imported from the Middle East. The data is obtained from the U.S. Energy Information Administration (EIA) for the period from January 2, 2001 to September 27, 2012. In addition, we use daily data on the level of the Dow Jones index obtained from Yahoo! finance for the same period.

Table 2.1(A) presents summary statistics for the wholesale oil price p_t, and its log $ln\, p_t$, logarithmic first-difference $\Delta ln\, p_t$, as well as the log of the Dow Jones index $ln\, j_t$ and its first-difference $\Delta ln\, j_t$. Note that $\Delta ln\, p_t$ and $\Delta ln\, j_t$ are scaled up by a factor of 100. Table 2.1(B) reports the results of unit root tests for p_t, $ln\, p_t$ and $\Delta ln\, p_t$, and Figure 2.1 displays graphs of the three series. Due to the presence of unit roots in p_t and $ln\, p_t$, we use logged first differences $\Delta ln p_t$ to estimate all autoregressive GARCH models in the following section.

2.3 Univariate Modelling of Crude Oil Prices

We estimate two different mean equations to model the daily change in the oil price. The first is a simple ARMA(1,1) model with no additional variables (equation 2.1); the second is an ARMA(1,1) model with additional regressors - a GARCH-in-mean parameter, contemporaneous and lagged change in the Dow Jones index, and dummy variables representing day of the week and season (equation 2.2). The Dow Jones is included to account for short-term economic shocks, as well as for the speculative component in oil price volatility. After fitting several versions of the model, we discover that a vector of seasonal dummy variables has more predictive power than monthly dummies or weather-related variables. The number of autoregressive and moving-average lags (1,1) was chosen using the Schwarz Information Criterion (SIC). The two mean equations are represented as:

$$\Delta ln\, p_t = \alpha + \beta_1 \Delta ln\, p_{t-1} + \beta_2 \varepsilon_{t-1} + \varepsilon_t \tag{2.1}$$

$$\begin{aligned} \Delta ln\, p_t = \alpha &+ \beta_1 \Delta ln\, p_{t-1} + \beta_2 \varepsilon_{t-1} + \beta_3 h_t + \beta_4 \Delta ln\, j_t + \beta_5 \Delta ln\, j_{t-1} \\ &+ \sum_{i=1}^{4} \delta_i w_t + \sum_{j=1}^{4} \gamma_j s_t + \varepsilon_t \end{aligned} \tag{2.2}$$

When estimated as Box-Jenkins equations using Maximum Likelihood, both models show strong evidence of heteroskedasticity in the error term, with Jarque-Bera statistics of 2936 and 2784, respectively. When only using data from 2001-2008, we obtained Jarque-Bera statistics of 1537 and 1454 from the two equations, indicating that heteroskedasticity was a central feature of the wholesale oil market even before the Great Recession. The standardized residual distributions are negatively skewed and have "fat tails". This heteroskedasticity is the motivation for the use of GARCH models, which correct for the non-normal error distribution by dynamically adjusting the conditional variance to take account of variations in the magnitude of the error term. Controlling heteroskedasticity by using GARCH allows us to obtain valid standard errors to assess the fit of our model, and to construct forecasts with correct prediction intervals.

We use the Schwarz Information Criterion (SIC) to select the optimal number of lagged residual "ARCH" terms p and lagged variance "GARCH" terms q in the GARCH(p, q) specification, and arrive at GARCH$(1, 1)$ as the optimal choice. SIC selects the same specification for a model in log-levels, the results of which we do not report due to bias caused by unit roots. However, SIC is still a valid test to use on such a model because it uses only a model's maximum likelihood values, and does not rely on coefficient estimates. The lagged variance term q forces the model to put a higher weight on past (squared) error terms in determining the current period variance, so the model achieves the same effect as by adding a large number of ARCH lags, with a single extra parameter:

$$
\begin{aligned}
h_t \equiv var(u_t) &= c_0 + a_1\varepsilon_{t-1}^2 + b_1 h_{t-1} \\
&= c_0 + a_1\varepsilon_{t-1}^2 + b_1(c_0 + a_1\varepsilon_{t-2}^2 + +b_1 h_{t-2}) \\
&= \sum_{k=0}^{N} b^k c_0 + a_1 \sum_{k=0}^{N} b_1^k \varepsilon_{t-k-1}^2 \qquad (2.3)
\end{aligned}
$$

Next, we add an asymmetric ARCH effect to the variance equation, motivated by a conjecture that a negative and positive shocks to the oil price may have different impacts on its volatility. We use the GJR asymmetry coefficient proposed by Glosten et. al. in 1993: $\varepsilon_{t-1}^2 \times I_{\varepsilon<0}(\varepsilon_{t-1})$.

Adding the asymmetry factor also helps correct for skewness in the error term distribution. Its "fat tails" can be helped by using a non-normal distribution assumption in estimating the GARCH model. GARCH is able to accommodate any specified error term distribution; after fitting models with normal, Student's t and GED distributions, we settle on GED as the choice of best fit.

We have estimated IGARCH, EGARCH, GARCH-M and GARCH-X models; however, only the latter two prove to be potential improvements over the standard GARCH(1,1). A GARCH-M model, also known as GARCH-in-mean, incorporates error term variance into the mean equation. GARCH-X allows for additional explanatory variables in the variance equation. We choose to include the change in the log of the Dow Jones index $\Delta ln\ j_t$, with one lag.

While both GARCH-X(1,1) and GARCH-M(1,1) perform better than the baseline GARCH(1,1), we obtain the best fit by combining the features of both. The resulting GARCH-XM(1,1) model is represented in equations 2.2 and 2.5. For reference, we also reproduce the basic asymmetric GARCH(1,1) variance expression in equation 2.4:

$$h_t = c_0 + a_1\varepsilon_{t-1}^2 + b_1 h_{t-1} + d_1\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}) \tag{2.4}$$

$$h_t = c_0 + a_1\varepsilon_{t-1}^2 + b_1 h_{t-1} + d_1\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}) + d_2\Delta ln\ j_t + d_2\Delta ln\ j_{t-1} \tag{2.5}$$

2.4 Empirical Estimates

We estimate three variations of the ARMA-GARCH setup. Model (1) combines mean equation 2.1 with variance equation 2.4, and is the baseline asymmetric ARMA(1,1)-GARCH(1,1) without additional variables. Model (2) combines the extended mean equation 2.2 (including the GARCH-in-mean term) with the baseline variance equation 2.4. Finally, model (3) uses mean equation 2.2 with variance equation 2.5, and is the fullest version of the model, an extended form of ARMA(1,1)-GARCH-XM(1,1) which contains additional variables in both the mean and variance equations. We impose the GED error term distribution to better fit the "fat-tailed" error term density, and scale logged first-differences $\Delta ln\, p_t$ and $\Delta ln\, j_t$ by 100 to improve precision of the estimates (scaling yields conditional variance estimates that are close to 1 rather than 0.01). Table 2.2 reports the resulting empirical estimates.

The estimates for mean equation coefficients indicate that the daily change in oil price is a true ARMA process, as both the autoregressive (AR) and moving-average (MA) coefficients are large and highly significant. However, the AR coefficient is negative at approximately -0.8 (with slight variations across models), indicating a great degree of instability in the oil price, which is more likely to increase today if it decreased yesterday, and vice versa. The GARCH-in-Mean coefficient is small and insignificant, suggesting that current oil price volatility does not have a significant impact on the direction or magnitude of price change. The lagged change in the Dow Jones index has a negative effect on the current change in oil price in model (2), although its significance disappears in model (3). The coefficients on weekday and seasonal variables indicate that oil price is likely to increase on tuesdays and wednesdays, and to decrease in the fall and winter.

The most striking finding from the variance equation is the contrast of very low "ARCH" coefficient on ε_{t-1}^2 (below 0.1) and very high "GARCH" coefficient on h_{t-1} (above 0.9), suggesting that volatility in the oil market is extremely persistent. There is a significant asymmetric effect in all three specifications: negative residuals (representing unexpected declines in the oil price) are associated with 7% higher variance than positive residuals of equale magnitude. Finally, model (3) estimates suggest that neither contemporaneous nor lagged change in the Dow Jones index has a significant effect on oil price.

Since the three GARCH models are nested, their log-likelihood values are directly comparable. Models (2) and (3) have significantly higher log-likelihood values than model (1) (see Table 2.2(C)). Table 2.2(C) also reports diagnostic statistics for standardized residuals $\hat{\varepsilon}_t = \frac{\varepsilon_t}{\sqrt{h_t}}$, including descriptive statistics, the Ljung-Box Q test for residual autocorrelation, and the McLeod-Li Q^2 test for squared residual autocorrelation. Both tests assume the null hypothesis that the data are independently distributed, and an alternative hypothesis of autocorrelation. Table 2.2(c) reports Q and Q^2 statistics for lags of 30 and 100 periods, with p-values in parentheses. Both tests pass at conventional significance levels, so there is no significant evidence of autocorrelation in the levels or squares of standardized residuals. This demonstrates that all three GARCH models adequately control for residual autocorrelation that causes heteroskedasticity in the raw data. Figure 2.2 presents a histogram of standardized residuals and a Gaussian Kernel Estimator plot, which both suggest a distribution that is very close to normal. Figures 2.3 and 2.4 contain graphic representations of Q and Q^2 autocorrelations over 30 periods.

2.5 Forecasting

Figures 2.5 and 2.6 present 20- and 50-day forecasts for the expected proportional change in oil price, determined by the mean model, and a 95% confidence interval for each forecast, determined by the variance model, for each of the three GARCH specifications. Both static and dynamic forecasts were constructed, but only the latter are reported here because static forecasts yielded very similar results while being of lesser conceptual interest. The quality of mean-model forecasts can be assessed with a range of forecast performance statistics, including Mean Error (ME), Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Theil's U statistics (Theil, 1971). Theil's U statistic is the ratio of RMSE from the specified model to the RMSE of a "no-change" forecast. A value of $U < 1$ indicates that our model produces forecasts that are better than the "no change" forecast, which can result from a random walk model. Table 2.3 reports all of the forecast performance statistics described above. Based on the RMSE measures, the extended ARMA(1,1) mean equation used in models (2) and (3) performs better than the simple ARMA(1,1) used in model (1). All three models produce mean price change forecasts that are easily superior to a "no change" forecast, with Theil's U statistics of 0.644-0.698. Again, the extended mean model yields slightly better forecasts. Even over a 20-period horizon, ARMA(1,1) models used in this chapter outperform the random walk, although with Theil's U values ranging only from 0.704-0.705, the difference between the simple and extended formulations largely disappears.

Unfortunately no formal statistics exist to evaluate the fit of variance estimates produced by GARCH models; however, a graphic analysis of forecasts in Figures 2.5 and 2.6 shows bounds that accurately reflect periods of high and low price volatility.

2.6 Conclusion

This chapter demonstrates that univariate GARCH models are of great practical use because they effectively correct for heteroskedasticity, which in our case stems from clusters of high- and low-volatility trading days in the wholesale oil market, yielding valid standard errors on Maximum Likelihood coefficients, and correct forecast bounds. A GARCH model for oil price variance can be combined with a sophisticated mean model estimating the oil price level or change, allowing for the creation of accurate forecasts. One disadvantage of univariate GARCH models such as the ones estimated here is the fact that they cannot reveal interactions among energy markets.

However, such interactions can uncover the strength and direction of influence among energy markets, and answer significant research questions. This consideration is the primary motivation behind the multivariate GARCH models estimated in Chapter 5.

Table 2.1: Wholesale Oil Price: Summary Statistics

A. Summary Statistics

	p_t	$ln\, p_t$	$\Delta ln\, p_t$	$ln\, j_t$	$\Delta ln\, j_t$
Mean	61.31	4.000	0.043	9.272	0.010
Standard Error	28.05	0.500	2.578	0.150	1.268
Variance	7867	0.250	6.646	0.022	1.609
Skewness	0.375	-0.312	-0.073	-1.253	0.035
Excess kurtosis	-0.672	-1.058	5.034	5.034	7.558
J-B normality	120.7	179.6	3032	20696	6808

B. Unit Root Tests

	p_t	$ln\, p_t$	$\Delta ln\, p_t$
ADF	-1.657	-1.399	-9.254
ADF - 5% c.v.	-2.863	-2.863	-2.863
Phillips-Perron	-1.537	-1.409	-55.04
PP - 5% c.v.	-2.863	-2.863	-2.863
KPSS - $\hat{\eta}_\mu$	41.35	45.32	0.039
$\hat{\eta}_\mu$ - 5% c.v.	0.463	0.463	0.463
KPSS - $\hat{\eta}_\tau$	2.151	5.139	0.0339
$\hat{\eta}_\tau$ - 5% c.v.	0.146	0.146	0.146

Table 2.2: Empirical Estimates: univariate GARCH models with daily U.S. oil price data

A. Conditional Mean Equation

Coefficient on Variable	GARCH Model (1)	(2)	(3)
Constant	0.159 (0.023)*	0.267 (0.018)	0.265 (0.107)
$\Delta ln\, p_{t-1}$	-0.886 (0.000)	-0.755 (0.000)	-0.755 (0.000)
ε_{t-1}	0.864 (0.000)	0.727 (0.000)	0.727 (0.000)
h_t	-	-0.005 (0.803)	-0.005 (0.796)
$\Delta ln\, j_t$	-	-0.021 (0.347)	-0.009 (0.735)
$\Delta ln\, j_{t-1}$	-	-0.046 (0.043)	-0.039 (0.166)
mon	-	-0.010 (0.883)	-0.009 (0.912)
tue	-	0.161 (0.000)	0.163 (0.098)
wed	-	0.101 (0.032)	0.098 (0.407)
thur	-	0.048 (0.458)	0.047 (0.561)
winter	-	-0.310 (0.017)	-0.318 (0.062)
summer	-	-0.092 (0.478)	-0.093 (0.580)
fall	-	-0.251 (0.078)	-0.227 (0.167)

B. Conditional Variance Equation

Coefficient on Variable	GARCH Model (1)	(2)	(3)
Constant	0.166 (0.000)	0.174 (0.000)	0.136 (0.000)
ε^2_{t-1}	0.019 (0.000)	0.022 (0.027)	0.017 (0.050)
h_{t-1}	0.913 (0.000)	0.908 (0.000)	0.926 (0.000)
$\varepsilon^2_{t-1} I_{\varepsilon<0}(\varepsilon_{t-1})$	0.076 (0.000)	0.076 (0.000)	0.065 (0.000)
Shape	1.410 (0.000)	1.411 (0.000)	1.407 (0.000)
$\Delta ln\, j_t$	-	-	-0.073 (0.527)
$\Delta ln\, j_{t-1}$	-	-	-0.074 (0.532)

C. Standardized Residual Diagnostics and the Log Likelihood Ratio

Statistic	GARCH Model (1)	(2)	(3)
$\hat{\varepsilon}_t$ mean	-0.016	-0.016	-0.015
$\hat{\varepsilon}_t$ variance	1.010	1.010	1.010
$\hat{\varepsilon}_t$ skewness	-0.230	-0.217	-0.204
$\hat{\varepsilon}_t$ kurtosis	2.708	2.678	2.676
Jarque-Bera	901.6	879.1	875.4
$Q(30)$	25.00 (0.73)	25.86 (0.63)	24.99 (0.68)
$Q^2(30)$	37.17 (0.14)	35.28 (0.20)	35.11 (0.20)
$Q(100)$	98.68 (0.60)	94.88 (0.60)	93.65 (0.63)
$Q^2(100)$	122.1 (0.06)	121.9 (0.06)	123.7 (0.05)
Log-Likelihood	-6425	-6419	-6418

** P-values in parentheses*

Table 2.3: Performance Statistics for 20-Day Forecasts

Forecast Statistic	GARCH Model (1)	(2)	(3)
ME	0.311	0.281	0.274
MAE	1.565	1.530	1.524
RMSE	1.994	1.949	1.950
Theil's U (1-step)	0.698	0.644	0.644
Theil's U (20-step)	0.705	0.705	0.704

Figure 2.1: Wholesale Oil Price at Cushing, Oklahoma

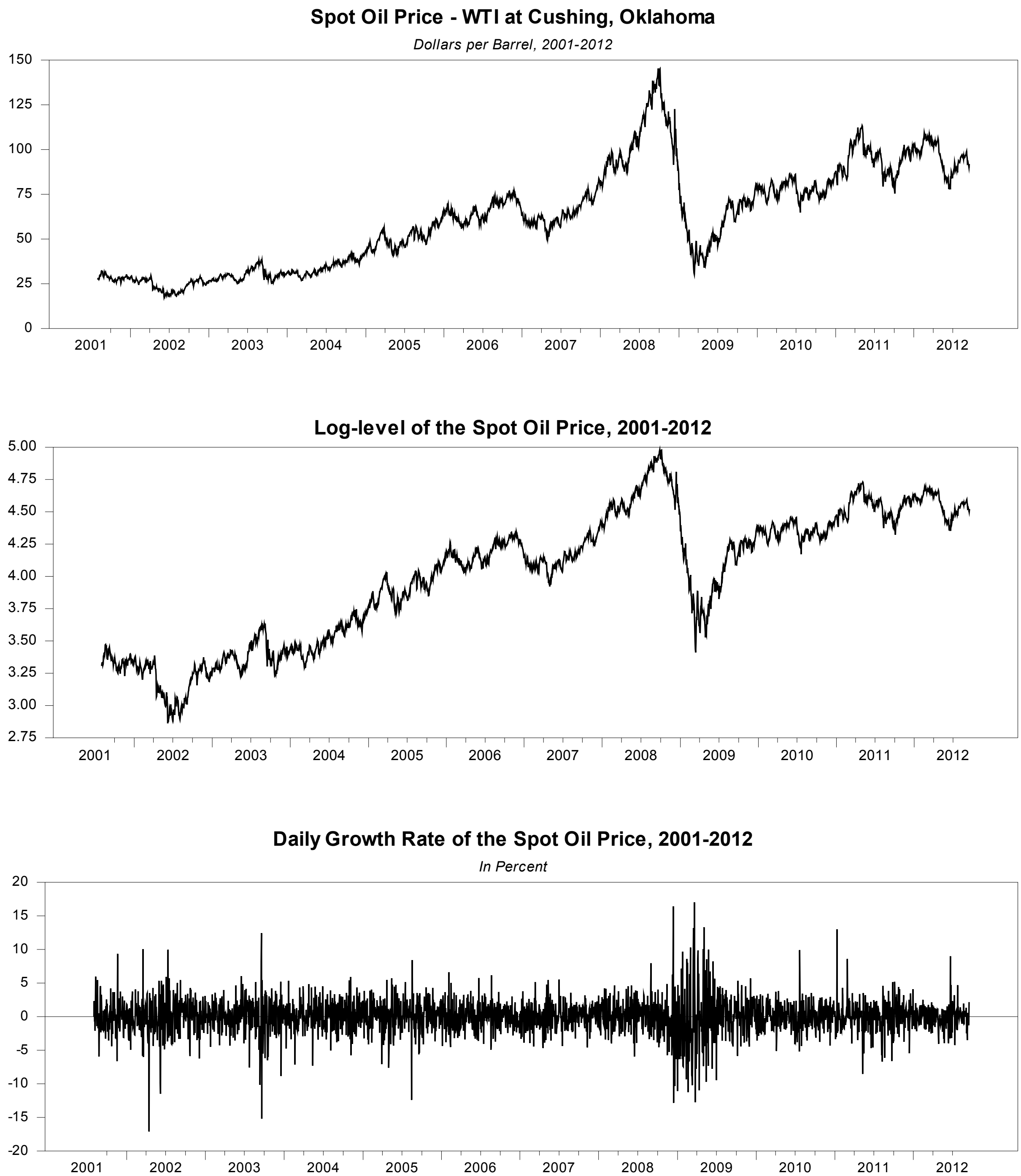
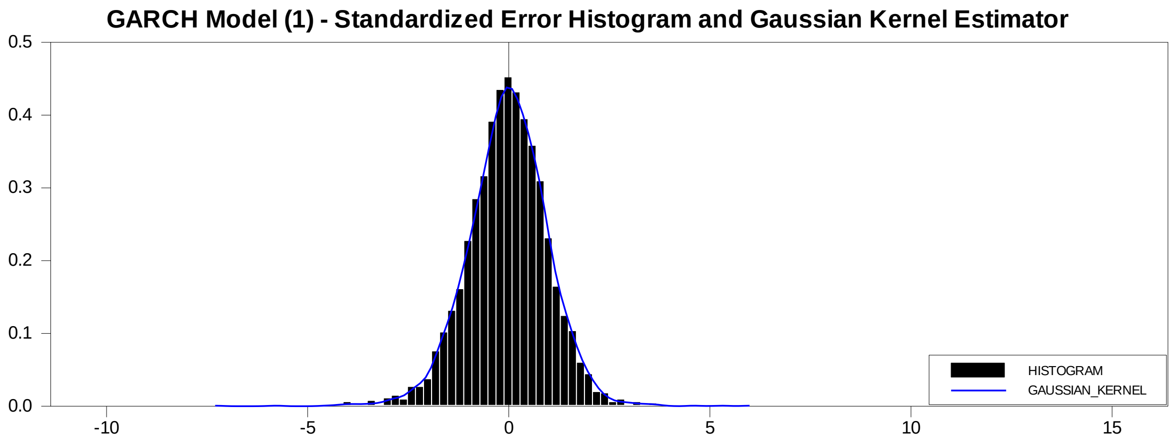
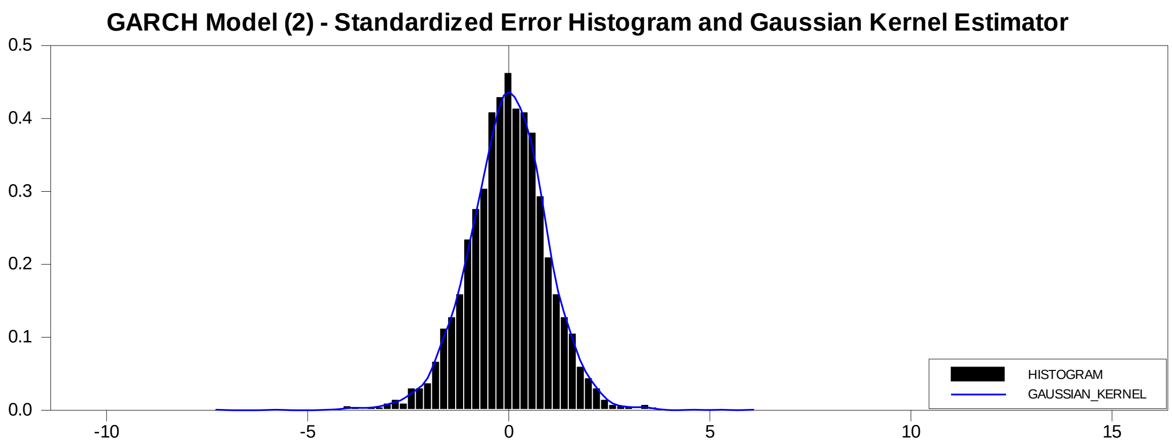


Figure 2.2: Residual Density Diagnosic Plots for Univariate GARCH Models of the Oil Price





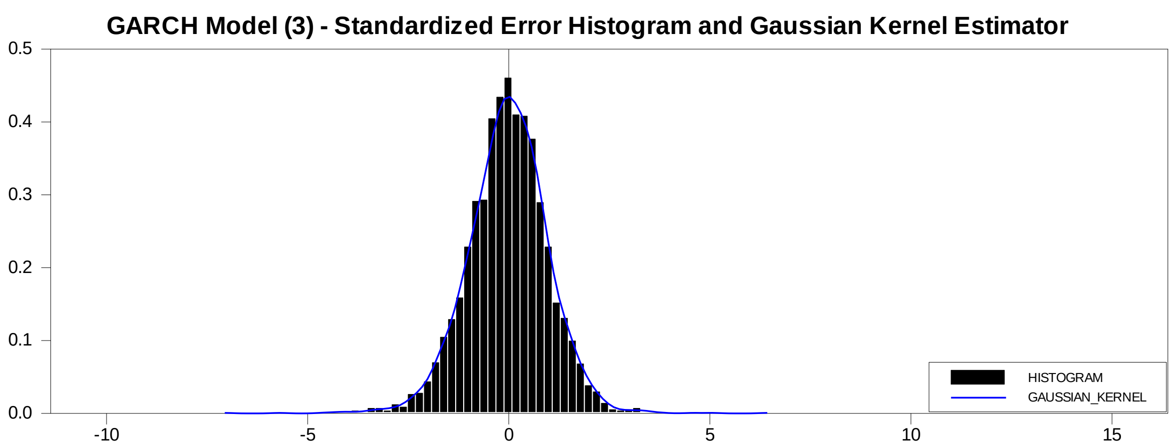
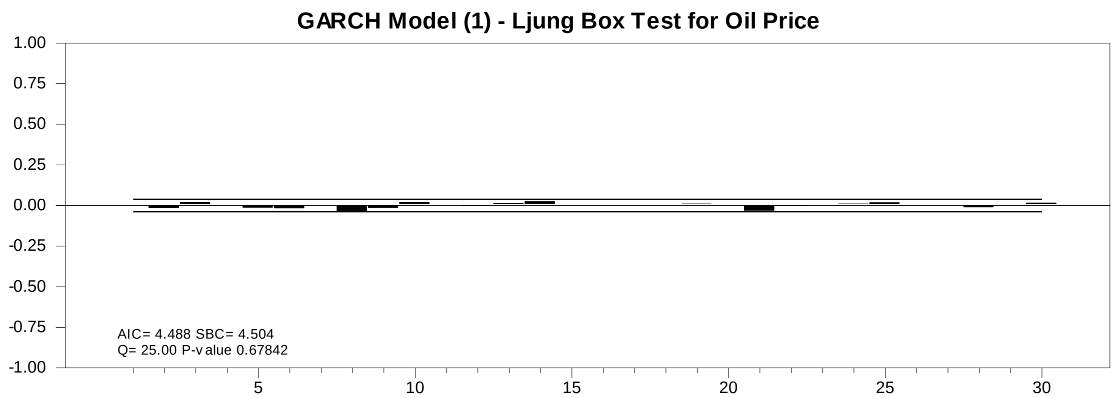
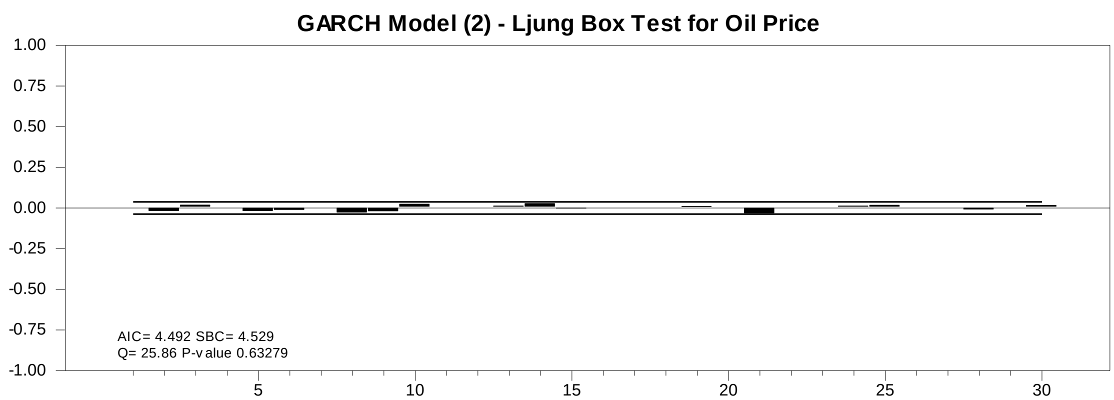


Figure 2.3: Residual Autocorrelation Tests for Univariate GARCH Models of the Oil Price





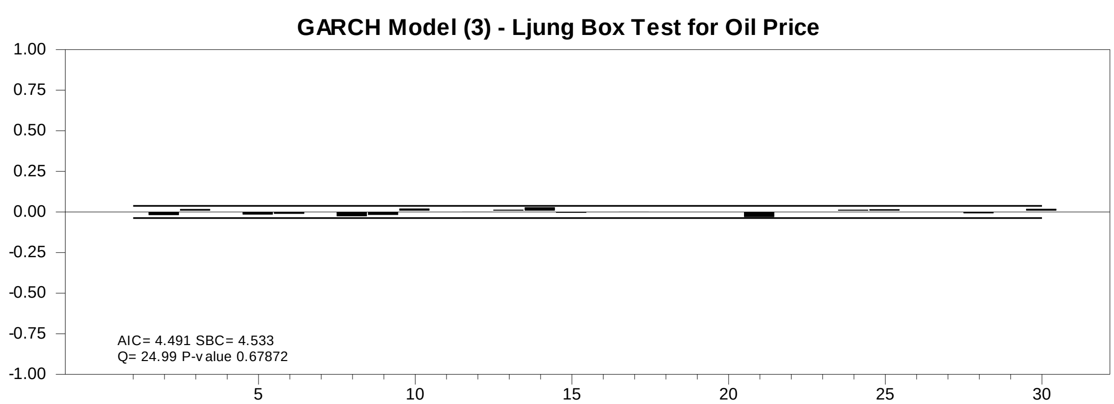
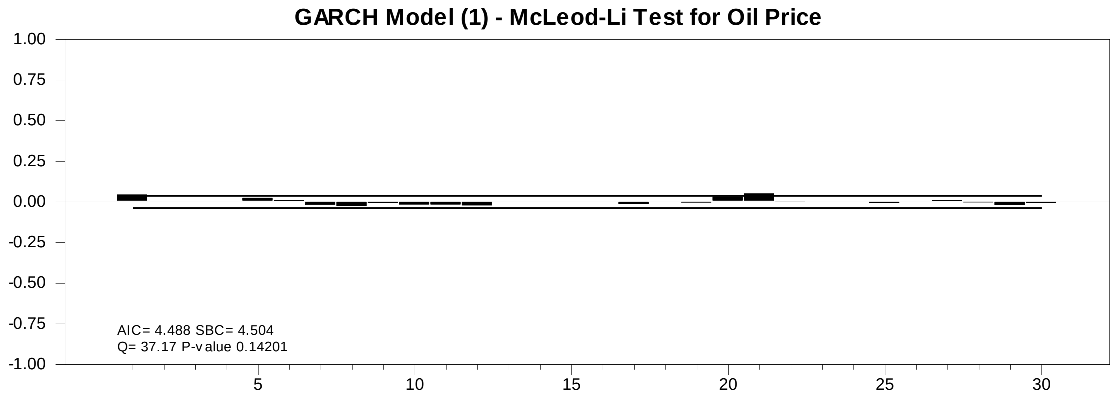
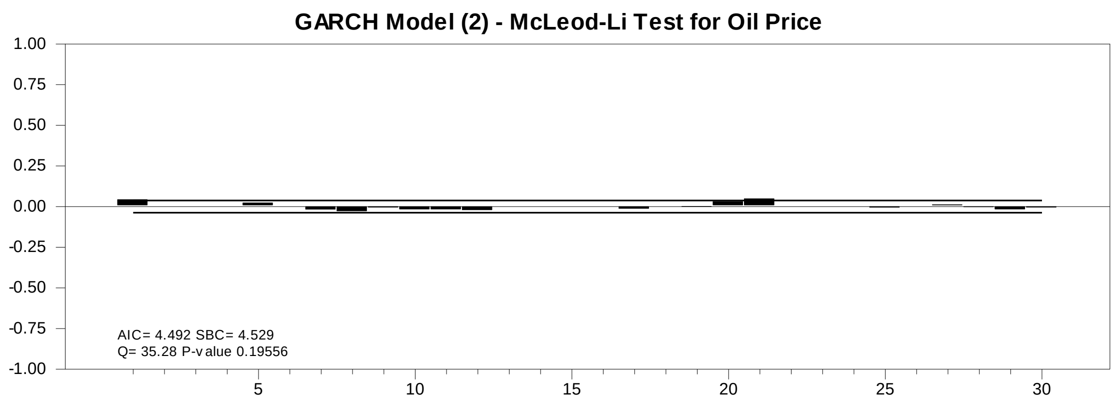


Figure 2.4: Squared Residual Autocorrelation Plots for Univariate GARCH Models of the Oil Price





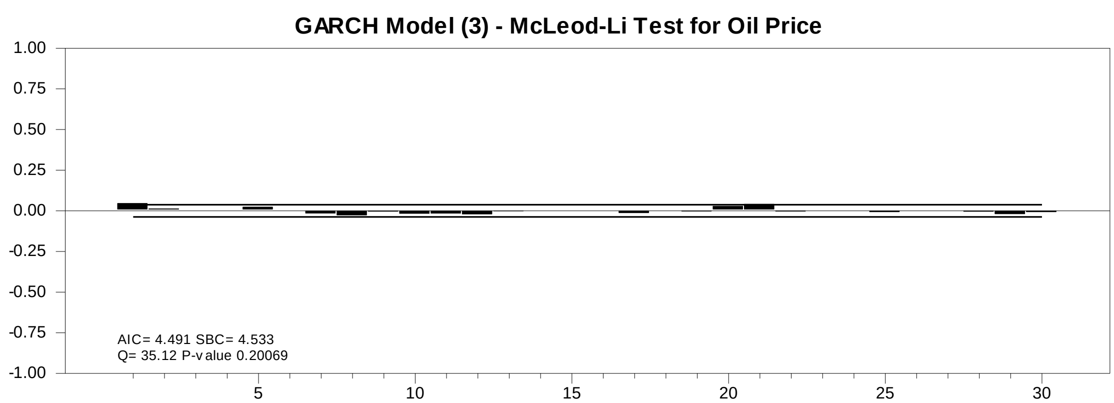


Figure 2.5: 20-Day Forecasts for Percentage Daily Change in Oil Prices by Univariate GARCH Models

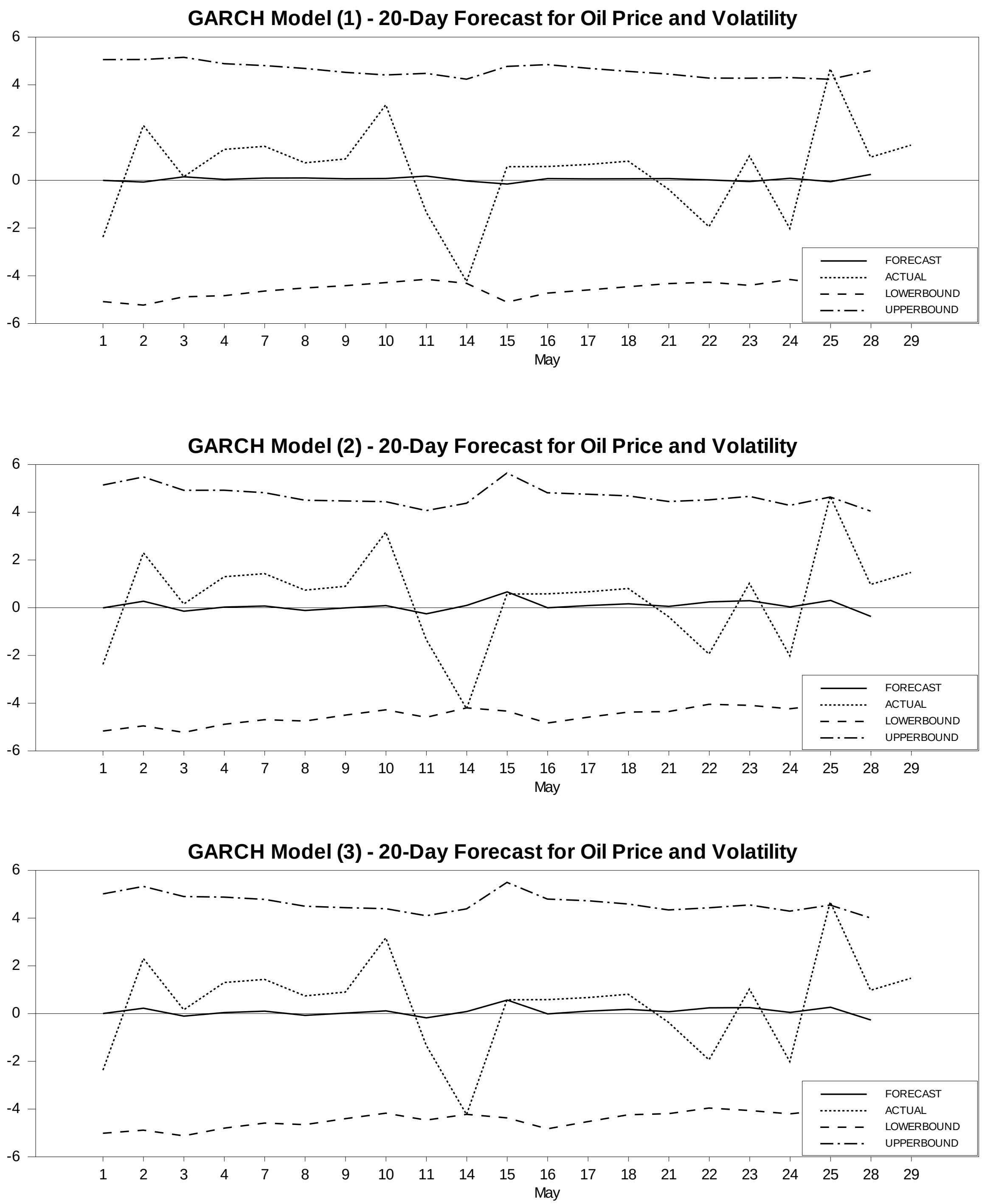
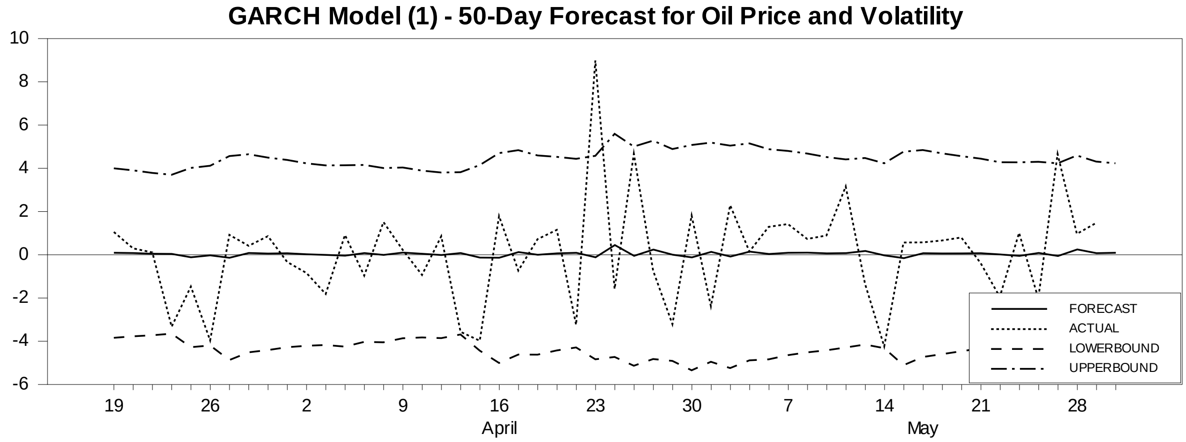
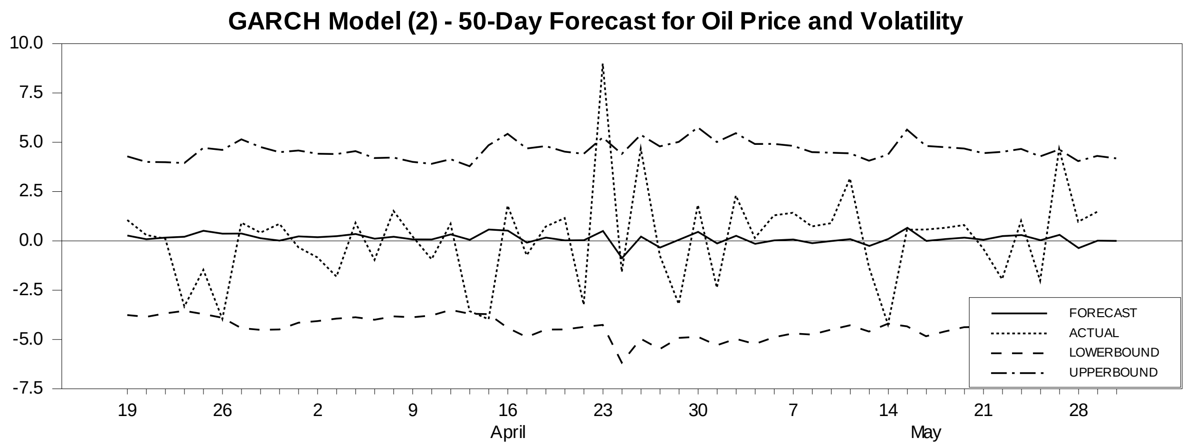
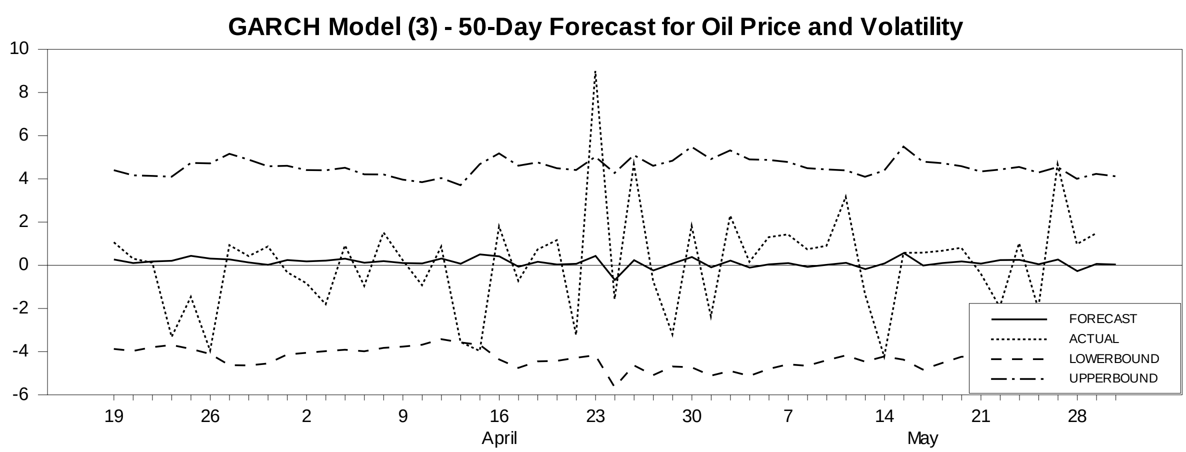


Figure 2.6: 50-Day Forecasts for Percentage Daily Change in Oil Prices by Univariate GARCH Models







3 Natural Gas Price Volatility – Univariate Models.

3.1 Introduction

Volatility modelling of natural gas prices receives almost as much attention as that of oil prices, primarily because being able to construct accurate forecasts has tremendous implications for hedging and derivatives trading in financial markets. At the same time, the natural gas market is influenced in a much larger extent by fundamental factors, such as predictable fluctuations in demand driven by weather variables, storage and transportation conditions, and seasonal production patterns, which make it much easier to construct models with significant predictive power (EIA, 2012). A third of natural gas in the United States is delivered to residential and commercial consumers (32.4%); 31.2% is used to generate electricity; and 27.8% is used in industrial sectors (EIA, 2011). On cold days, more natural gas is demanded to provide heating to households and businesses, both directly and as an input to electricity production. The same happens on hot days when energy is demanded for air conditioning. There is considerable variation in energy input mix among states; overall, Texas and California are the largest consumers of natural gas ($3,345$ and $2,274$ billion cubic feet, respectively), followed by Louisiana, New York and Florida.

3.2 The Data

We use daily price data on the wholesale price of natural gas at Henry Hub for the period from January 2, 2001 to September 10, 2012. The data is provided by the U.S. Energy Information Administration (EIA). Henry Hub is located in Erath, Louisiana, and is the largest natural gas trading hub in the United States, with direct connections to 13 major pipelines. The EIA is also our source for data on the total storage inventory of natural gas in the continental U.S. We transform weekly storage data into daily data in the following way. Assuming a constant rate of inventory change during each week, we calculate the average daily rate of change as $(v_w - v_{w-1})/5$, where v_w represents the observed storage level on Wednesday of week w, then use this value to fill in the four missing observations in each week. Further, we take a logarithm of each daily inventory level to obtain series $ln\, v_t$. We also use daily Dow Jones index data from Yahoo! finance in logarithmic form ($ln\, j_t$) and daily data on mean daily temperature from the National Oceanic and Atmospheric Administration's National Climatic Data Center. The latter is transformed into Degree Days (DD), a common measure used in empirical finance literature (Mu, 2007). Total degree days (DD) is a sum of heating degree days (HDD) and cooling degree days (CDD):

$$DD_t = CDD_t + HDD_t \tag{3.1}$$

$$CDD_t = Max(0,\ T - 65^oF) \tag{3.2}$$

$$HDD_t = Max(0,\ 65^oF - T) \tag{3.3}$$

Table 3.1(A) presents summary statistics for levels (p_t), log-levels ($ln\, p_t$) and logged first-differences ($\Delta ln\, p_t$) of the gas price series. The $\Delta ln\, p_t$ series is scaled up by 100 to make its scale similar to that of the other variables. Figure 3.1 presents the series graphically. Table 3.1(B) reports the results of several unit root tests, including augmented Dickey-Fuller (ADF), Phillips-Perron (PP) and Kwiatowki, Phillips, Schmidt and Shin (1992) (KPSS) tests. The ADF and PP tests evaluate the null hypothesis of a unit root against an alternative of stationarity, while the KPSS test assumes a null hypothesis of stationarity and an alternative of a unit root. KPSS statistic $\widehat{\eta}_\mu$ represents a test with a null hypothesis of stationarity around a constant, and $\widehat{\eta}_\tau$ assumes a null of stationarity around a trend. For p_t and $ln\, p_t$, all three tests point to the existence of a unit root: the ADF and PP (both augmented with 30 lags) do not reject the null hypothesis of a unit root at the 5% level, while the KPSS rejects the null of stationarity at the 5% level. This leads us to difference the natural gas price series before estimating autoregressive GARCH models in the following section.

3.3 Univariate Modelling of Natural Gas Prices

We use two models for the mean daily change in natural gas price - a baseline ARMA$(2,2)$ (equation 3.4), and an extended AR(2) with additional variables (equation 3.5). For both models, the optimal lag order was selected using the Schwarz Information Criterion (SIC). We find that changes in the gas price are to a large extent driven by observable fundamental factors. When we add variables representing weather, day of the week, season, level of the Dow Jones index, and gas storage conditions, the moving-average terms, which represent the impact of recent unexpected price shocks, become small and insignificant, making an AR specification preferable to ARMA. The AR(2) model includes the following additional regressors: logged first difference in the Dow Jones index, $\Delta ln\, j_t$; log of the nation-wide storage inventories $ln\, v_t$; DD measures for California and Texas, d_{ct} and d_{xt}; and sets of dummy variables representing day of the week w_{jt} and season s_{kt}. In addition, the AR(2) includes a set of interaction terms between storage inventory and the season, because the two are highly correlated: inventory levels are the highest in the fall after the summer off-shore drilling season, and the lowest in the spring, as offshore drilling is rarely done in the winter due to storms and otherwise unfavourable weather conditions.

$$\Delta ln\, p_t = \alpha + \sum_{i=1}^{2} \beta_i \Delta ln\, p_{t-i} + \beta_3 \varepsilon_{t-1} + \beta_4 \varepsilon_{t-2} + \varepsilon_t \tag{3.4}$$

$$\begin{aligned} \Delta ln\, p_t = \; & \alpha + \sum_{i=1}^{2} \beta_i \Delta ln\, p_{t-i} + \beta_3 \Delta ln\, j_t + \beta_4 ln\, v_t + \beta_5 d_{ct} + \beta_6 d_{xt} + \\ & \sum_{j=1}^{4} \beta_{6+j} w_t + \sum_{k=1}^{3} \beta_{10+k} s_t + \sum_{k=1}^{3} \beta_{13+k} s_t \times ln\, v_t + \varepsilon_t \end{aligned} \tag{3.5}$$

We originally estimated the extended model with various combinations of CDD and HDD measures for east- and west-coast states that are the largest consumers of natural gas, including Louisiana, Florida, New York, Pennsylvania. However, being situated on the coast of the Gulf of Mexico, Louisiana and Florida do not experience great temperature fluctuations, making it difficult to pinpoint the feedback from weather shocks to natural gas demand.

Likewise, temperature in New York and East Coast in general had very little predictive power on the natural gas price. Further, we did not find decomposition of DD into HDD and CDD helpful, and finally chose d_{ct} and d_{xt} as the most useful variables to include. We estimated a version of the model where the three seasonal dummies were replaced with a dummy variable f_t to represent the "fall" months of August, September and October – the months of the Atlantic hurricane season which may be associated with supply disruptions. However, we found f_t to yield small and insignificant coefficients in both mean and variance equations, which may reflect the fact that vast inventories of natural gas provide a sufficient buffer against temporary supply disruptions. As Figure 3.1 demonstrates, natural gas inventories were never close to zero during the past decade, and have followed an upward trend against the background of seasonal fluctuations (inventories are built up during the summer season of offshore production, and depleted during the winter). In fact, only unanticipated shocks could affect the natural gas price due to transportation constraints; however, data on the discrepancy between forecasted and actual weather, or on the actual days of unexpected hurricane-caused shutdowns is not available for the United States.

We choose two variance equation specifications - a basic GARCH(1,1) (equation 3.6), and a GARCH-X with additional variables representing the contemporaneous change in the Dow Jones index and the day of the week (equation 3.7). Both equations are symmetric and assume normal error distributions. In testing various models, we found no reasons for using asymmetry or changing the error distribution as we did with oil price models in Chapter 2. Furthermore, we do not include a GARCH-in-mean term because we found its coefficient to be small and statistically insignificant in trial estimations. The resulting two variance equations are:

$$h_t = c_0 + a_1 u_{t-1}^2 + b_1 h_{t-1} \tag{3.6}$$

$$h_t = c_0 + a_1 u_{t-1}^2 + b_1 h_{t-1} + d_1 \Delta ln\, j_t + \sum_{j=1}^{4} d_{1+j} w_t \tag{3.7}$$

3.4 Empirical Estimates

Table 3.2 reports the empirical estimates from three GARCH specifications: model (1) combines the baseline mean and variance equations (3.4) and (3.6); model (2) uses the extended mean equation (3.5) with the baseline variance equation (3.6); model (3) uses the extended versions (3.5) and (3.7). The key empirical estimates are remarkably similar across specifications.

The coefficients on lags pairs $\Delta ln\, p_{t-1}$ and $\Delta ln\, p_{t-2}$, and ε_{t-1} and ε_{t-2} are of different signs, suggesting that the gas price tends to "overreact" to shocks, and then move in the opposite direction two days later to compensate. Day of the week has a large impact on the change in natural gas price, which tends to increase on Monday and decrease towards Friday. Conditional on season, nation-wide storage inventory is negatively correlated with the natural gas price, although this effect is not statistically significant. Natural gas prices are the highest in the spring and the lowest in the fall, after the summer season of off-shore drilling which allows producers to replenish inventories.

Higher DD measures of temperature have a positive effect on the gas price in Texas, but a negative one in California, with both coefficients being significant but small. The absence of a large positive effect driven by derived demand for electricity could be due to temporal interactions of high temperatures with other events affecting the energy markets. For example, California's electricity generation fuel mix currently includes 20% hydro and 20% renewable sources, both of which are more productive in the spring-summer season (U.S. Department of Energy). This can temporarily reduce demand for natural gas for electricity generation, depressing its price.

The variance equation reveals several interesting features of natural gas price volatility. First, it is rather persistent, as evidenced by a large coefficients on h_{t-1} of $0.82 - 0.85$, and small coefficients on ε_{t-1}^2 of $0.15 - 0.17$. Secondly, there is a significant contagion effect from the stock market: a 1% increase in the Dow Jones index is associated with a 0.3% decrease in natural gas price volatility, and vice versa. Finally, volatility is the lowest on Tuesdays and the highest on Fridays. Table 3.2(C) reports descriptive and diagnostic statistics for standardized residuals $\hat{\varepsilon}_t = \frac{\varepsilon_t}{\sqrt{h_t}}$ (see section 2.4 for a description of the statistics included). The very large number of observations makes the data susceptible to failing the usual tests for residual normality, including skewness, kurtosis and Jarque-Bera tests. However, Figure 3.2, which includes a histogram of $\hat{\varepsilon}_t$ and a Gaussian Kernel Estimator plot, shows that the distribution of $\hat{\varepsilon}_t$ is actually close to normal. The McLeod-Li test for squared residual autocorrelation passes at various lag lengths; while the Ljung-Box test for residual correlation does not pass, a graphical analysis of residual correlation in Figures 3.3 and 3.4 shows that the small amount of autocorrelation present is not a cause for concern. The most complete model specification (3) performs slightly better in all diagnostic tests.

3.5 Forecasting

Figures 3.5 and 3.6 present 20- and 50-day forecasts for $\Delta \ln p_t$ obtained from the three GARCH models. Models (2) and (3) perform significantly better than model (1); incorporating additional information in the mean and variance equations allows these models to anticipate movements in the gas price with greater accuracy, and to narrow down the 95% forecast confidence bounds. Although no formal statistics exist to evaluate the accuracy of volatility forecasts, graphic analysis suggests that GARCH models employed to forecast natural gas volatility perform significantly better than those used for oil volatility in Chapter 2. Table 3.3 reports forecast performance statistics for , including Mean Error (ME), Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE) for 20-day forecasts, and Theil's U statistics for 1-day and 20-day forecasts (see section 2.5 for a description of these statistics). Models (2) and (3) have lower MAE and RMSE values than model (1). All three models significantly outperform the random walk model, especially for very short-term forecasts, with Theil's U statistics of 0.700-0.703. In fact, Theil's U values remain well below one even for a 20-day forecast horizon.

3.6 Conclusion

We presented a comprehensive study of natural gas price volatility in the United States during the period of 2000-2012, using daily data on wholesale natural gas prices, nation-wide storage inventories, weather conditions in key regions, and day of the week and seasonal dummies. GARCH models constructed yield valuable insights into the impact of these and other factors on the proportional change in the natural gas price, and into the characteristics and causes of natural gas price volatility. We find that natural gas price volatility is rather persistent, but also much more affected by actual events in the natural gas markets than oil price volatility. Unlike analagous models estimated as Box-Jenkins equations using Maximum Likelihood, GARCH models presented in this chapter adequately correct for heteroskedasticity, allowing us to obtain valid standard errors for hypothesis tests and meaningful conclusions regarding the impact of various factors on gas prices and volatility. In addition, our models produce forecasts for the direction and magnitude of natural gas price changes that are superior to those produced by the random walk model. Most importantly, they also generate forecasts for natural gas price volatility that are significantly more accurate and responsive to market information than those obtained from corresponding Box-Jenkins models.

Table 3.1: Wholesale Natural Gas Price: Summary Statistics

A. Summary Statistics

	p_t	$ln\, p_t$	$\Delta ln\, p_t$	$ln\, v_t$	$ln\, j_t$	$\Delta ln\, j_t$
Mean	5.476	1.612	-0.042	7.726	9.272	0.010
Standard Error	2.358	0.422	4.746	0.493	0.150	1.268
Variance	5.562	0.178	22.53	0.243	0.022	1.609
Skewness	1.134	-0.058	0.549	-7.039	-1.253	0.035
Excess kurtosis	1.852	-0.309	21.27	102.3	12.84	7.558
J-B normality	1034	13.14	54728	128860	20696	6808

B. Unit Root Tests

	p_t	$ln\, p_t$	$\Delta ln\, p_t$
ADF	-2.216	-1.942	-11.11
ADF - 5% c.v.	-2.863	-2.863	-2.863
Phillips-Perron	-3.399	-2.687	-53.94
PP - 5% c.v.	-2.863	-2.863	-2.863
KPSS - $\hat{\eta}_\mu$	8.264	9.504	0.051
$\hat{\eta}_\mu$ - 5% c.v.	0.463	0.463	0.463
KPSS - $\hat{\eta}_\tau$	7.932	9.032	0.054
$\hat{\eta}_\tau$ - 5% c.v.	0.146	0.146	0.146

Table 3.2: Empirical Estimates: univariate GARCH models with daily U.S. gas price data

A. Conditional Mean Equation

Coefficient on Variable	GARCH Model (1)	(2)	(3)
Constant	-0.148 (0.058)*	-5.456 (0.003)	-4.156 (0.027)
$\Delta ln\, p_{t-1}$	0.385 (0.000)	0.014 (0.546)	0.019 (0.385)
$\Delta ln\, p_{t-2}$	-0.687 (0.000)	-0.080 (0.000)	-0.081 (0.000)
ε_{t-1}	-0.369 (0.001)	-	-
ε_{t-2}	0.609 (0.000)	-	-
$\Delta ln\, j_t$	-	-0.077 (0.156)	-0.078 (0.134)
$ln\, v_t$	-	0.695 (0.007)	0.506 (0.048)
$ln\, v_t \times summer$	-	-0.928 (0.144)	-0.677 (0.287)
$ln\, v_t \times fall$	-	-1.415 (0.376)	-0.802 (0.602)
$ln\, v_t \times winter$	-	-0.824 (0.299)	-0.399 (0.602)
d_{ct}	-	-0.027 (0.001)	-0.023 (0.004)
d_{xt}	-	0.017 (0.050)	0.015 (0.078)
mon	-	0.896 (0.000)	0.928 (0.000)
tue	-	0.894 (0.000)	0.874 (0.000)
wed	-	0.723 (0.001)	0.724 (0.000)
thur	-	0.599 (0.004)	0.647 (0.000)
winter	-	-0.199 (0.390)	-0.170 (0.448)
summer	-	-0.173 (0.449)	-0.121 (0.602)
fall	-	-0.630 (0.026)	-0.404 (0.142)

B. Conditional Variance Equation

Coefficient on Variable	GARCH Model (1)	(2)	(3)
Constant	0.381 (0.000)	0.438 (0.000)	3.510 (0.000)
ε^2_{t-1}	0.151 (0.000)	0.171 (0.000)	0.167 (0.000)
h_{t-1}	0.847 (0.000)	0.827 (0.000)	0.823 (0.000)
$\Delta ln\, j_t$	-	-	-0.306 (0.034)
mon	-	-	-2.658 (0.011)
tue	-	-	-5.237 (0.000)
wed	-	-	-3.477 (0.000)
thur	-	-	-3.336 (0.001)

** P-values in parentheses*

(continued on next page)

Table 3.2 (continued)

C. Standardized Residual Diagnostics for Univariate GARCH Models

Statistic	GARCH Model		
	(1)	(2)	(3)
$\hat{\varepsilon}_t$ mean	0.017	0.018	0.014
$\hat{\varepsilon}_t$ variance	1.000	1.000	1.000
$\hat{\varepsilon}_t$ skewness	0.917	0.862	0.739
$\hat{\varepsilon}_t$ kurtosis	11.09	10.06	8.290
Jarque-Bera	15232	12565	8554
$Q(30)$	37.44 (0.17)*	50.84 (0.00)	45.67 (0.01)
$Q^2(30)$	11.92 (0.99)	12.15 (0.99)	12.70 (0.99)
$Q(100)$	139.3 (0.01)	163.9 (0.00)	153.2 (0.00)
$Q^2(100)$	27.91 (1.00)	30.73 (1.00)	35.52 (1.00)
Log-Likelihood	-8048	-8031	-8013

* *P-values in parentheses*

Table 3.3: Performance Statistics for 20-Day Forecasts

Forecast	GARCH Model		
Statistic	(1)	(2)	(3)
ME	0.156	-0.216	-0.231
MAE	2.191	1.695	1.686
RMSE	2.824	2.072	2.069
Theil's U (1-step)	0.703	0.701	0.700
Theil's U (20-step)	0.719	0.718	0.717

Figure 3.1: Wholesale Natural Gas Price at Henry Hub and Nationwide Storage Inventories

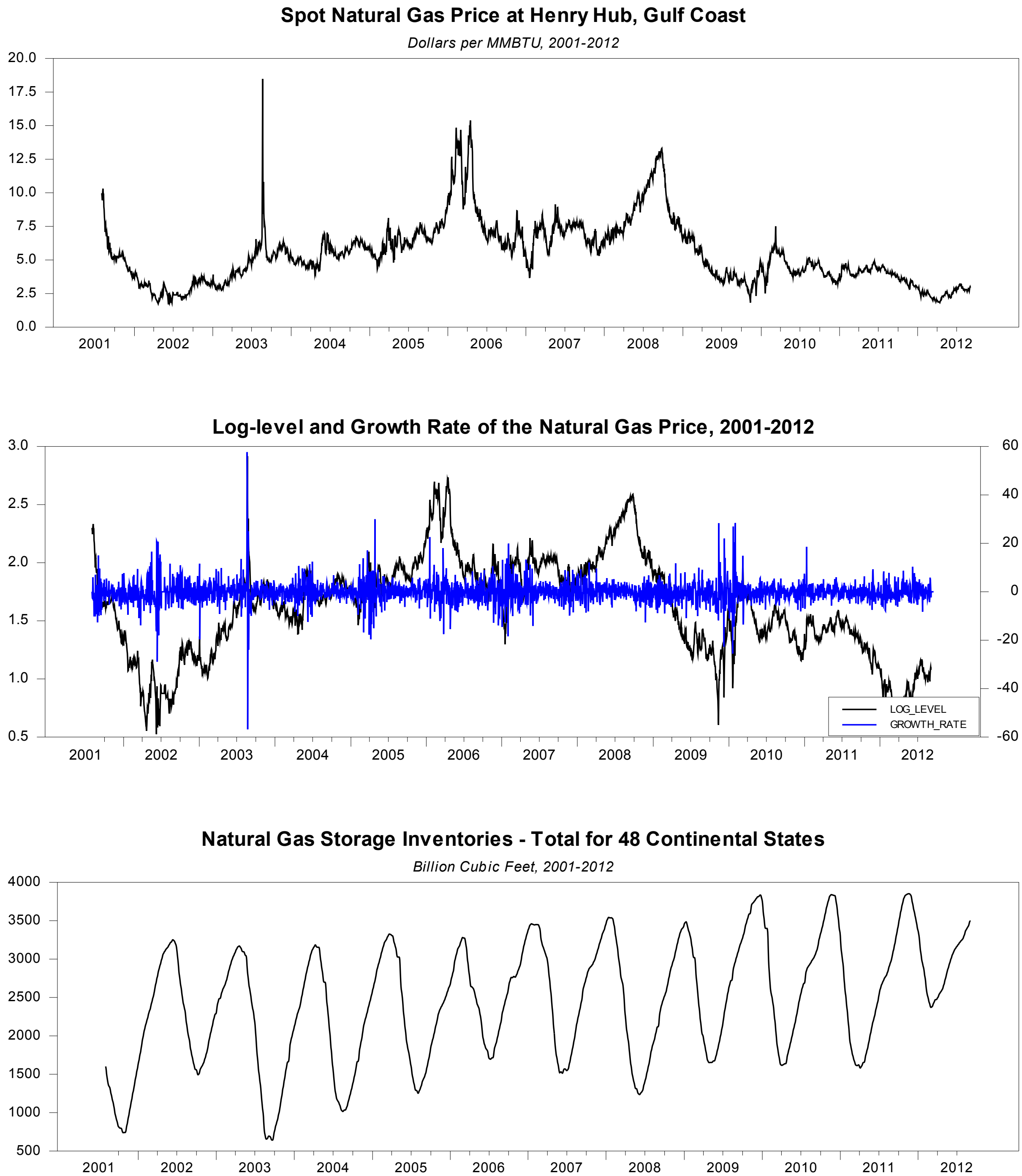


Figure 3.2: Residual Density Diagnosic Plots for Univariate GARCH Models of the Natural Gas Price

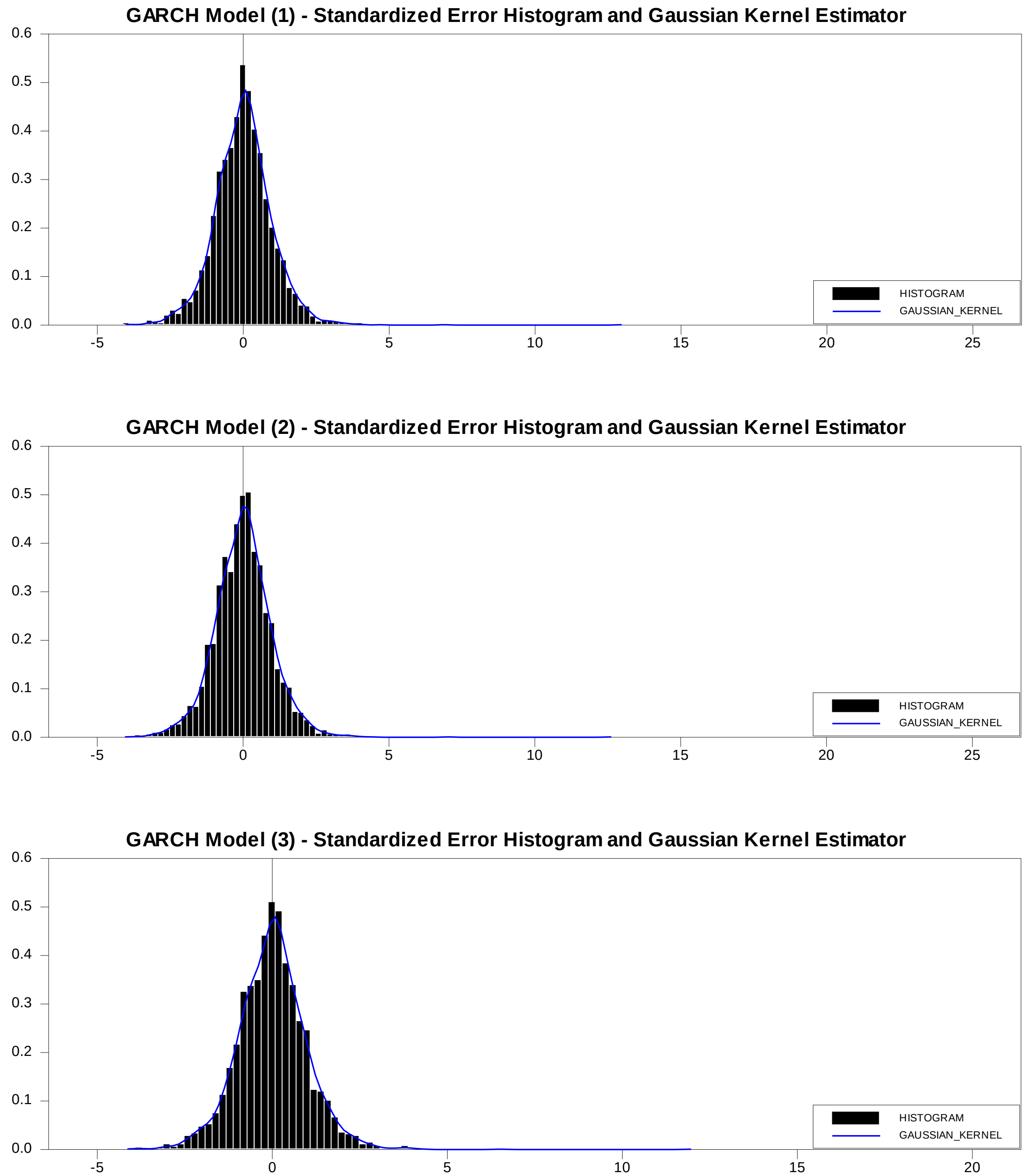


Figure 3.3: Residual Autocorrelation Tests for Univariate GARCH Models of the Natural Gas Price

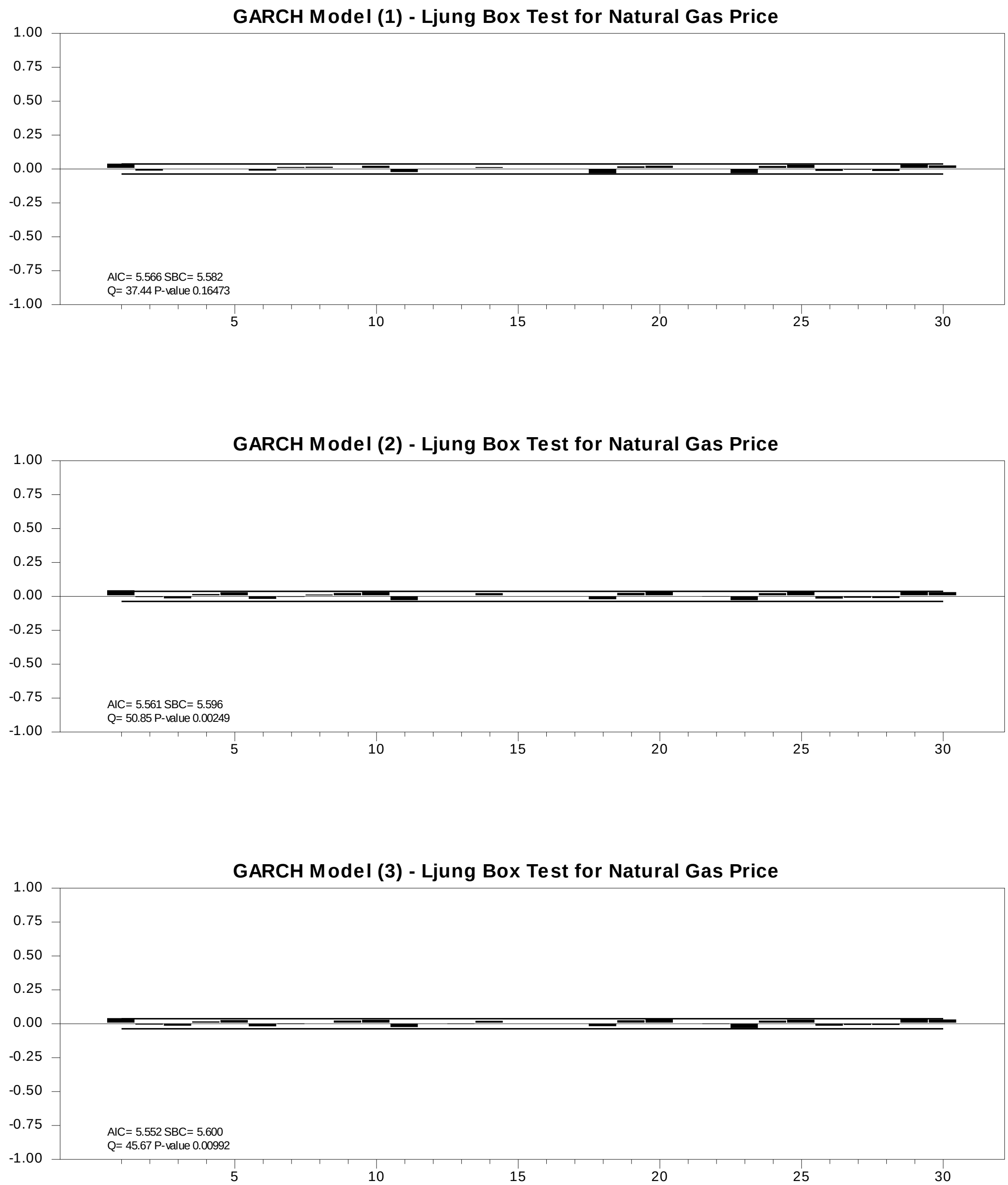
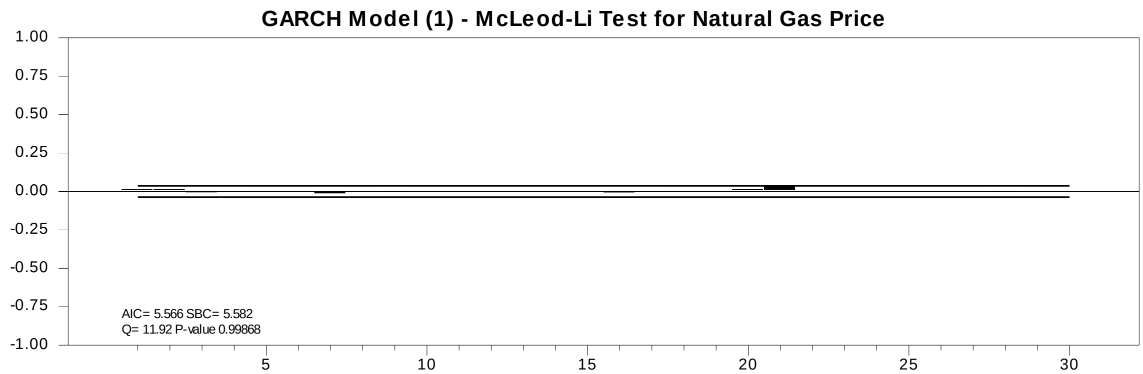
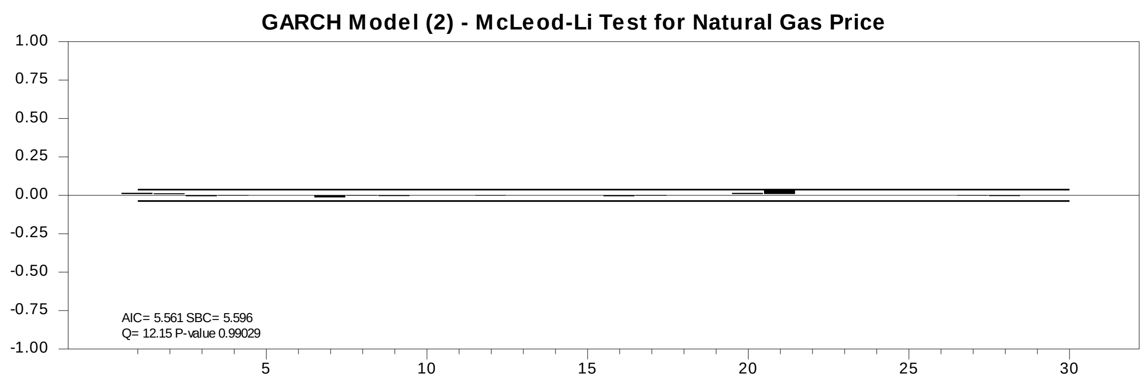


Figure 3.4: Squared Residual Autocorrelation Plots for Univariate GARCH Models of the Natural Gas Price





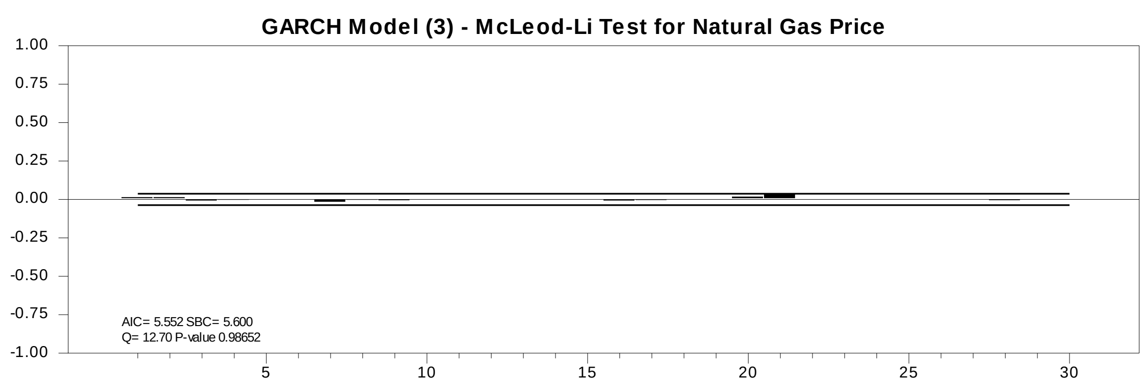
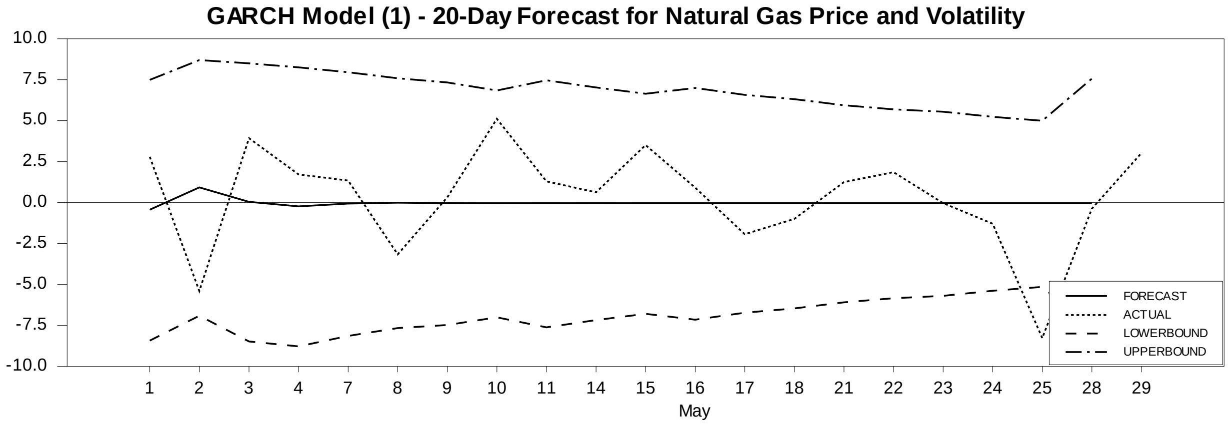
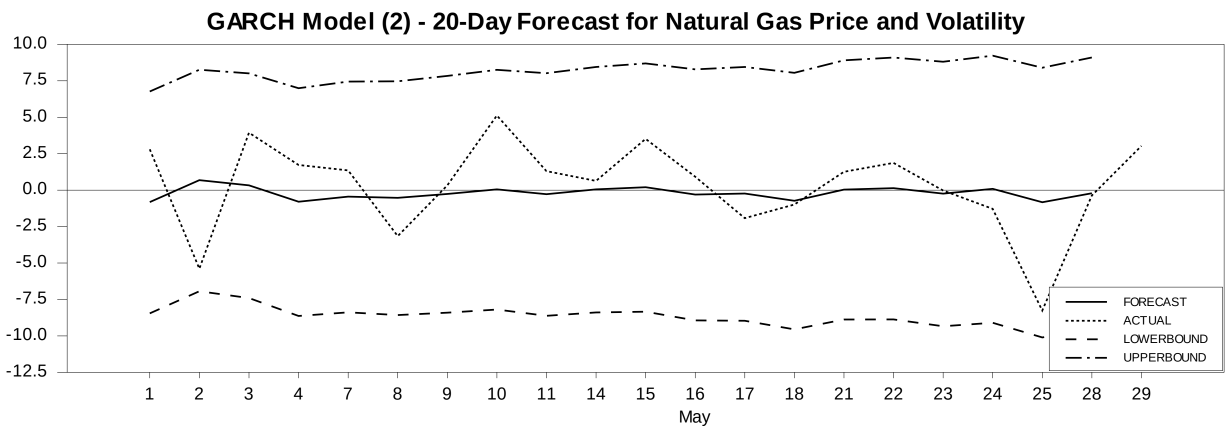


Figure 3.5: 20-Day Forecasts for Percentage Daily Change in Gas Prices by Univariate GARCH Models





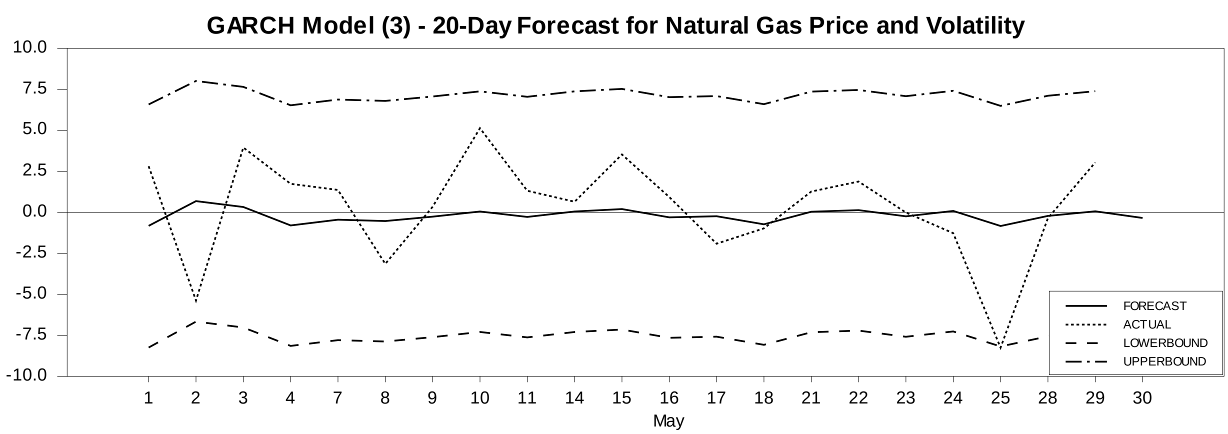
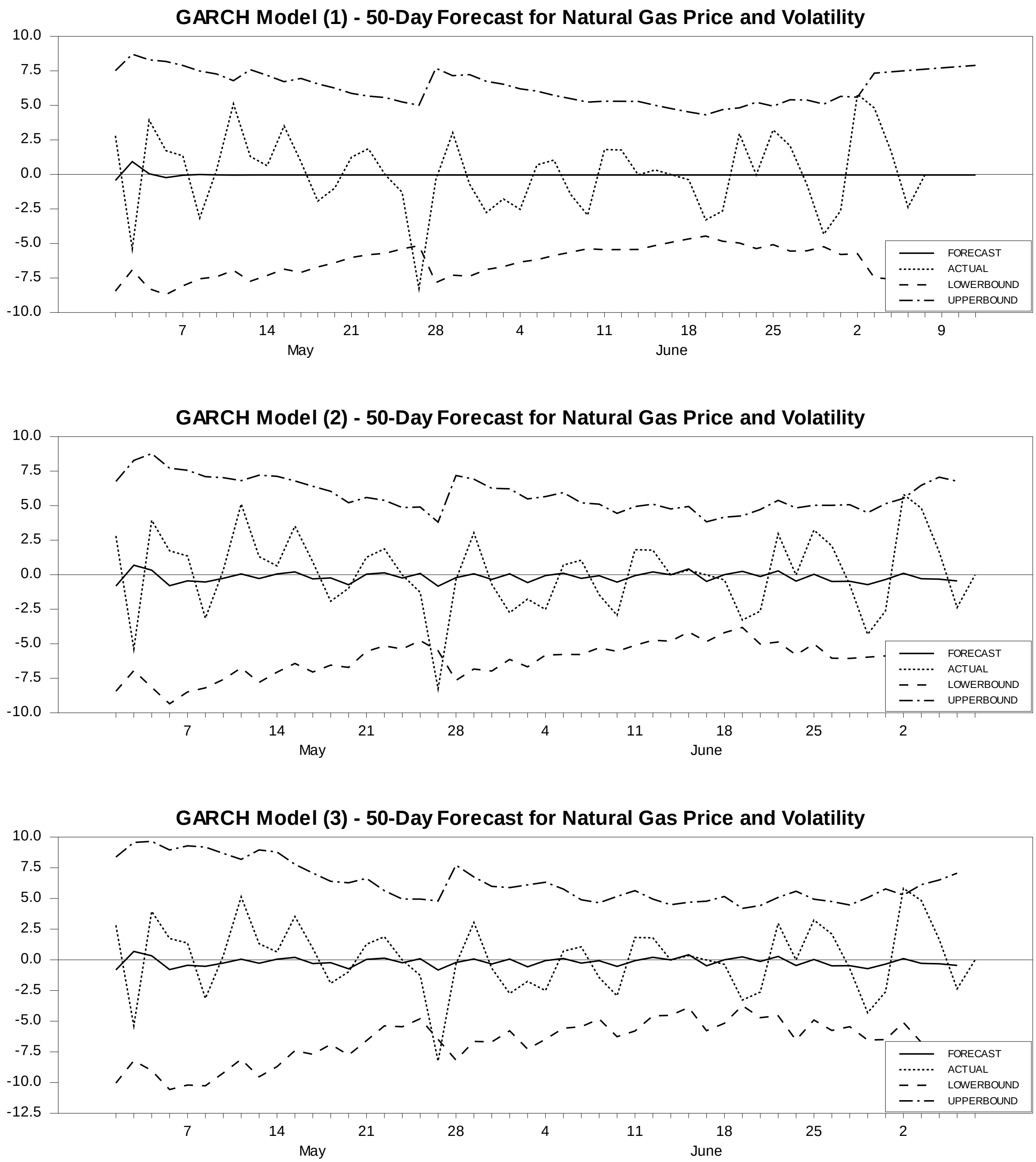


Figure 3.6: 50-Day Forecasts for Percentage Daily Change in Gas Prices by Univariate GARCH Models



4 Electricity Price Volatility – Univariate Models.

4.1 Introduction

In the United States, electricity has only been traded in competitive wholesale markets for just over a decade, following the initiation of energy sector deregulation by several states in 1996-2000. When vertically integrated utilities were restructured to separate the generation, transmission and distribution functions, markets were created where generators could sell electricity to retail distributors and directly to large industrial consumers. Despite several failures, including the California energy crisis in 2001 and the Northeast blackout in 2003, deregulation has been adopted to various degrees in 27 U.S. states, so that two-thirds of electricity consumed in the U.S. is currently bought and sold in wholesale markets (Electric Power Supply Association, 2012).

Electricity is much more vulnerable to extreme price events than other energy commodities, because of its nonstorability, high transportation costs caused by physical losses and transmission constraints, and highly inelastic demand. In addition, electricity supply is also inelastic at high output levels. Every region's generation capacity is composed of a unique mix of technologies, which differ by marginal cost and by their ability to quickly change the level of output. For example, coal plants have low marginal costs, but are unable to react to unexpected price changes, because shutting down and restarting a coal generator is extremely costly. On the other hand, "peaker" natural gas turbines have high marginal costs but can ramp up generation in as little as 15 minutes. Because they only produce at times of peak demand, these plants must charge a very high price to cover their fixed costs. At times of abnormally high demand or restricted supply, a region's generators may be pushed to full capacity, resulting in extreme price spikes. Electricity price volatility has become a political issue and resulted in the suspension of electricity deregulation in several states, and its reversal in Virginia (Energy Information Administration, 2010).

Being able to forecast electricity price volatility is of great importance to energy and financial market participants, as well as policymakers. As Sadorsky (2012) puts it, "Modelling and forecasting volatility lies at the heart of modern finance because good estimates of correlation and volatility are needed for derivative pricing, portfolio optimization, risk management, and hedging" (p. 240). At the same time, there is a relatively small body of recent research on volatility modelling of wholesale electricity prices. Worthington et. al. (2005) and Higgs (2009) use multivariate GARCH models to explore price and volatility spillovers among Australian wholesale electricity markets. Hickey et. al. (2012) compare the forecasting performance of several GARCH models over very short forecast horizons, using hourly U.S. data. Lindström and Regland (2012) estimate a range of Markov Regime Switching Models to measure the degree of contagion between six European electricity markets during extreme price events. Several studies using a univariate GARCH approach were published in late 1990s-early 2000s (for a survey, see Hadsell et al. (2004), Higgs and Worthington (2005) or Chan and Gray (2006)); however, no significant univariate GARCH study modelling the volatility of U.S. electricity markets has been published since 2006. This chapter addresses the issue by fitting a range of univariate GARCH models to wholesale U.S. spot electricity price data that is as recent as July 2012. It should be noted that although the volatility modelling literature has turned in favour of more complicated methods, univariate GARCH models continue to be useful as they often deliver more accurate estimates and superior forecasting performance to multivariate models (Wang & Wu, 2012), as demonstrated in section 4.5.

4.2 The Data

We use daily price data on wholesale electricity prices obtained from the U.S. Energy Information Administration (EIA). The time period covered is from January 1, 2001, to September 17, 2012. Unlike crude oil and natural gas, electricity is currently traded in over a dozen of wholesale markets across the United States. Moreover, new markets have started during the period of 2001-2012, and some of these have displaced trading activity from formerly major markets. We construct a nation-wide wholesale electricity price by taking a weighted average of daily prices at six wholesale markets, each of which has been among the largest at least at one point during the decade. The markets included are: Nepool Mass Hub in New England; PJM West in Pennsylvania; Entergy in Louisiana; Mid Columbia hub; SP15 EZ and SP15 hubs in California; and the ERCOT South hub in Texas. We found that wholesale electricity prices are very highly correlated across markets, which means that taking a weighted average hardly attentuates the degree of volatility in the markets. Figure 4.1 presents a graph of the wholesale electricity price in level, log and logged first difference forms for 2001-2012.

The only other external data we use is the mean daily temperature for Texas and California obtained from the National Oceanic and Atmospheric Administration's National Climatic Data Center, and converted into Degree Days (DD) (for a discussion of DD and their relevance to energy markets, see section 3.2, and Mu (2007)). DD is constructed from the sum of Cooling Degree Days (CDD) and Heating Degree Days (HDD) as follows:

$$DD_t = CDD_t + HDD_t \tag{4.1}$$

$$CDD_t = Max(0,\ T - 65^o F) \tag{4.2}$$

$$HDD_t = Max(0,\ 65^o F - T) \tag{4.3}$$

Table 4.1(A) reports summary statistics for levels (p_t), log-levels ($ln\, p_t$) and logged first-differences ($\Delta ln\, p_t$) of the wholesale electricity price. Only $\Delta ln\, p_t$ is scaled by a factor of 100 to keep its range in sync with that of other variables used. Heteroskedasticity is obvious in all three series, motivating our approach of using GARCH models in the following section.

However, we need to establish stationarity before estimating autoregressive models. Table 3.1(B) reports the results of ADF, PP and KPSS tests (see section 3.2 for a brief discussion on these). The ADF and PP tests assume a null hypothesis of a unit root and an alternative of stationarity, while the KPSS test does the opposite. KPSS statistic $\widehat{\eta}_\mu$ tests for stationarity around a constant, while $\widehat{\eta}_\tau$ tests stationarity around a trend. As reported in Table 3.1(B), unit root tests deliver contradictory results: the ADF and PP reject the null hypothesis of a unit root at a 5% level, while the KPSS rejects the null hypothesis of stationarity at a 5% level for p_t and $ln\, p_t$. Such erratic behaviour of unit root tests has been noted by economists for several macroecomic time series. For example, Caporale et. al. (2003) examine this problem in the study of purchasing power parity (PPP). The authors suggest constructing a recursive ADF test by first running the ADF regression using only a few initial observations in the dataset, then recursively adding observations and observing how the ADF t-statistic changes as a result. Adopting this approach, we create a plot of recursive ADF t-statistics for p_t and $ln\, p_t$ against 5% and 10% ADF critical values, displayed in Figure 4.2. The ADF test applied to electricity prices displays the same type of erratic behaviour that Caporale et. al. observed in regards to PPP. They suggest that sharp "spikes" in t-values point to breaks in the data-generating process, while drift over time may indicate non-stationarity of an unknown type (p. 284). Moreover, Figure 4.2 demonstrates that just dropping observations from 2010-2012 would cause the ADF test to consistently not reject the null for either p_t or $ln\, p_t$. A portion of this period, winter of 2011-2012, is associated with abnormal price fluctuations (see Figure 4.1). Therefore, we cannot rule out the presence of non-stationarity in the data, and use differenced series $\Delta ln\, p_t$ to construct autoregressive GARCH models in the following section.

4.3 Univariate Modelling of Electricity Prices

Following the same approach as for univariate oil and natural gas price volatility models, we build three GARCH models, starting with a baseline ARMA-GARCH(1,1) setup for model (1), then including additional variables in the mean equation for model (2), and finally including additional variables in both the mean and variance equations for model (3). We use the Schwarz Information Criterion to select the number of AR and MA lags in the mean equation, arriving at ARMA(2,1) as the optimal choice (equation 4.4). Models (2) and (3) use an extended version of ARMA(2,1) with variables representing day of the week w_t, season s_t, and temperature measured in DD for California and Texas, d_{ct} and d_{xt}. We originally estimated the model with DD measures for more states, as well as regional DD averages, but dropped those variables after obtaining small and statistically insignificant estimates. East-coast temperature fluctuations seem to not be extreme enough to cause swings in the electricity price, once season is controlled for, while averaging across states masks extreme temperature events that have the most impact on the price. Likewise, we do not include a GARCH-in-Mean term because preliminary estimations revealed it to be significant but very small in magnitude. The resulting extended mean model is represented in equation 4.5.

$$\Delta ln\, p_t = \alpha + \sum_{i=1}^{2} \beta_i \Delta ln\, p_{t-i} + \beta_3 \varepsilon_{t-1} + \varepsilon_t \tag{4.4}$$

$$\begin{aligned} \Delta ln\, p_t &= \alpha + \sum_{i=1}^{2} \beta_i \Delta ln\, p_{t-i} + \beta_3 \varepsilon_{t-1} + \beta_4 d_{ct} + \beta_5 d_{xt} + \\ &+ \sum_{j=1}^{4} \beta_{5+j} w_t + \sum_{k=1}^{3} \beta_{9+k} s_t + \varepsilon_t \end{aligned} \tag{4.5}$$

Models (1) and (2) use the standard GARCH(1,1) variance equation with the asymmetry coefficient of Glosten et. al. (1993), represented in equation 4.6. To construct its extended version for model (3), we add terms representing the lagged change in electricity price and day of the week (see equation 4.7). We use the normal error distribution, as it provides a better fit to the data than the alternatives, GED and Student's t. The two variance equations are reproduced below:

$$h_t = c_0 + a_1 u_{t-1}^2 + b_1 h_{t-1} + d_1 \varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}) \tag{4.6}$$

$$h_t = c_0 + a_1 u_{t-1}^2 + b_1 h_{t-1} + d_1 \varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1}) + d_2 \Delta ln\, p_{t-1} + \sum_{j=1}^{4} d_{1+j} w_t \tag{4.7}$$

4.4 Empirical Estimates

Empirical estimates for GARCH models (1), (2) and (3) defined in the previous section are reported in Table 4.2. Coefficients on AR and MA terms are large and highly significant in all three models, suggesting that the logged first difference of the wholesale electricity price is a true ARMA process. Estimates for models (2) and (3) suggest that price is most likely to decrease on tuesdays and thursdays, and increase on fridays. This is possibly due to the fact that most contracts for weekend delivery are finalized on friday. The seasonal variation is also very pronounced - on average, electricity prices are 25-36% higher in the winter and 21-26% higher in the summer, than in the spring and fall. Surprisingly, DD measures for California and Texas have the opposite effects - temperature extremes in Texas are associated with moderate electricity price increases, while those in California actually slightly depress the price, with both effects being statistically significant. This could be caused by time correlations between periods of hot weather in California and other factors affecting the electricity market. For example, the hot months of early summer in California are also associated with lower prices for natural gas (see chapter 3), and with the highest annual abundance of hydroelectricity, which are used to generate 52% and 17% of California's energy, respectively. Texas, on the other hand, relies more heavily on coal (37%), which has fixed marginal costs throughout the year (Get Energy Active, 2011).

In the variance equation, both the "ARCH" ε_{t-1}^2 and "GARCH" h_{t-1} terms are highly significant. The coefficient on h_{t-1} is estimated at 0.72 for all three models, which is significantly smaller than in natural gas and oil models, suggesting that electricity price volatility is less persistent. There is significant negative asymmetric effect - an unanticipated price increase is associated with 11-12% higher volatility than a decrease of the same magnitude. This effect disappears once we include $\Delta ln\, p_{t-1}$ in the variance equation, likely due to correlation between the two terms. It is also interesting to note that the direction of asymmetry is the opposite to that estimated in the oil price models of Chapter 2. The lagged first-difference in electricity price $\Delta ln\, p_{t-1}$ has a large and positive effect on variance: a 10%

change in electricity price in the previous period is associated with 11% higher variance, conditional on ε_{t-1}^2, $h_{t-1}$and other factors.

Table 4.2(C) reports descriptive statistics for standardized residuals $\hat{\varepsilon}_t = \frac{\varepsilon_t}{\sqrt{h_t}}$, as well as the Ljung-Box Q-statistic for residual autocorrelation, and the McLeod-Li Q^2 statistic for squared residual autocorrelation (see section 2.4 for a description). Although the latter two tests do not pass at conventional significance levels, Figures $4.4 - 4.5$ reveal that autocorrelations at each lag length are small and random. In addition, Figure 4.3 shows that the distribution of the standardized residuals is very close to normal. Based on all diagnostic tests, model (2) slightly outperforms models (1) and (3).

4.5 Forecasting

Figures 4.6 and 4.7 show sample 20- and 50-day forecasts for the logged first difference in electricity price $\Delta ln\ p_t$ and its volatility, represented as 95% forecast confidence bounds. All three GARCH models perform similarly well, accurately predicting periods of high and low price volatility. While no formal statistics exist to evaluate the performance of volatility forecasts, we report a range of statistics measuring the quality of forecasts for Δlnp_t produced by the mean equation in Table 4.3 (see section 2.5 for a discussion on these). RMSE estimates suggest that model (1) produces better forecasts than models (2) and (3). However, Theil's U statistic is close to 0.61 for all three models, indicating that all three ARMA specifications significantly outperform a random walk model.

4.6 Conclusion

This chapter presented three univariate ARMA-GARCH models of electricity price volatility in the U.S. All three effectively control for heteroskedasticity in the data and produce forecasts that are significantly superior to a benchmark "no-change" forecast by a random walk model. Surprisingly, model (1), which does not include any additional variables, produces slightly better forecasts for the change in the price than models augmented with variables representing weather, day of the week and season. However, model (3), augmented with additional variables, seems to produce better volatility forecasts. Estimates also yield insight into the nature of electricity price volatility (less persistent than oil and natural gas price volatility), and into the degree of influence of factors such as weather on electricity price fluctuations.

A significant advantage of univariate models such as the ones presented in this chapter is the relatively small number of parameters to be estimated. Such models not only yield more precise maximum likelihood estimates, but can also be augmented with additional variables in both the mean and variance equations, to reflect factors like weather conditions, supply shocks and new government regulation. Multivariate GARCH models, on the other hand, quickly become overparameterized with the addition of new variables because each one requires a vector of parameters.

Table 4.1: Wholesale Electricity Price: Summary Statistics

A. Summary Statistics

	p_t	$ln\, p_t$	$\Delta ln\, p_t$
Mean	52.18	3.877	-0.033
Standard Error	21.00	0.405	15.01
Variance	440.9	0.164	225.2
Skewness	1.380	-0.567	-0.074
Excess kurtosis	5.463	2.947	36.70
J-B normality	4533	1206.4	162933

B. Unit Root Tests

	p_t	$ln\, p_t$	$\Delta ln\, p_t$
ADF	-3.023	-2.946	-11.758
ADF - 5% c.v.	-2.863	-2.863	-2.863
Phillips-Perron	-11.72	-10.322	-95.75
PP - 5% c.v.	-2.863	-2.863	-2.863
KPSS - $\hat{\eta}_\mu$	7.129	7.695	0.008
$\hat{\eta}_\mu$ - 5% c.v.	0.463	0.463	0.463
KPSS - $\hat{\eta}_\tau$	7.120	7.708	0.007
$\hat{\eta}_\tau$ - 5% c.v.	0.146	0.146	0.146

Table 4.2: Empirical Estimates: univariate GARCH models with daily U.S. electricity price data

A. Conditional Mean Equation

Coefficient on Variable	GARCH Model (1)	(2)	(3)
Constant	-0.002 (0.962)*	0.435 (0.443)	0.253 (0.667)
$\Delta ln\, p_{t-1}$	0.535 (0.000)	0.556 (0.000)	0.547 (0.000)
$\Delta ln\, p_{t-2}$	-0.131 (0.000)	-0.126 (0.000)	-0.118 (0.000)
ε_{t-1}	-0.691 (0.000)	-0.716 (0.000)	-0.717 (0.000)
d_{ct}	-	-0.022 (0.010)	-0.021 (0.016)
d_{xt}	-	0.022 (0.015)	0.025 (0.008)
mon	-	-0.175 (0.783)	-0.201 (0.764)
tue	-	-1.250 (0.027)	-1.271 (0.025)
wed	-	-0.309 (0.605)	-0.446 (0.446)
thur	-	-1.120 (0.089)	-1.093 (0.110)
winter	-	0.252 (0.066)	0.361 (0.015)
summer	-	0.259 (0.132)	0.208 (0.257)
fall	-	-0.016 (0.898)	-0.063 (0.630)

B. Conditional Variance Equation

Coefficient on Variable	GARCH Model (1)	(2)	(3)
Constant	6.965 (0.000)	6.740	13.20 (0.009)
ε_{t-1}^2	0.339 (0.000)	0.333	0.278 (0.000)
h_{t-1}	0.716 (0.000)	0.720	0.722 (0.000)
$\varepsilon_{t-1}^2 I_{\varepsilon<0}(\varepsilon_{t-1})$	-0.120 (0.002)	-0.112	-0.009 (0.863)
$\Delta ln\, p_{t-1}$	-	-	1.111 (0.013)
mon	-	-	-3.313 (0.678)
tue	-	-	-15.00 (0.029)
wed	-	-	-10.89 (0.120)
thur	-	-	-2.213 (0.776)

** P-values in parentheses*

Table 4.2 (continued)
C. Standardized Residual Diagnostics for Univariate GARCH Models

Statistic	GARCH Model		
	(1)	(2)	(3)
$\hat{\varepsilon}_t$ mean	0.006	0.000	-0.008
$\hat{\varepsilon}_t$ standard error	1.000	0.999	0.999
$\hat{\varepsilon}_t$ variance	1.000	0.999	0.999
$\hat{\varepsilon}_t$ skewness	0.283	0.262	0.270
$\hat{\varepsilon}_t$ kurtosis	5.482	5.344	5.166
Jarque-Bera	3674	3487	3263
$Q(30)$	89.48	90.64	92.13
$Q^2(30)$	84.35	83.99	84.79
$Q(100)$	201.9	197.9	200.1
$Q^2(100)$	176.3	175.6	174.3
Log-Likelihood	-10934	-10894	-10888

** P-values in parentheses*

Table 4.3: Performance Statistics for 20-Day Forecasts

Forecast Statistic	GARCH Model (1)	(2)	(3)
ME	0.135	-0.743	0.200
MAE	11.64	14.92	15.18
RMSE	14.50	21.06	21.06
Theil's U (1-step)	0.611	0.609	0.608
Theil's U (20-step)	0.688	0.681	0.685

Figure 4.1: Wholesale Electricity Price - Weighted Average at Major U.S. Trading Hubs

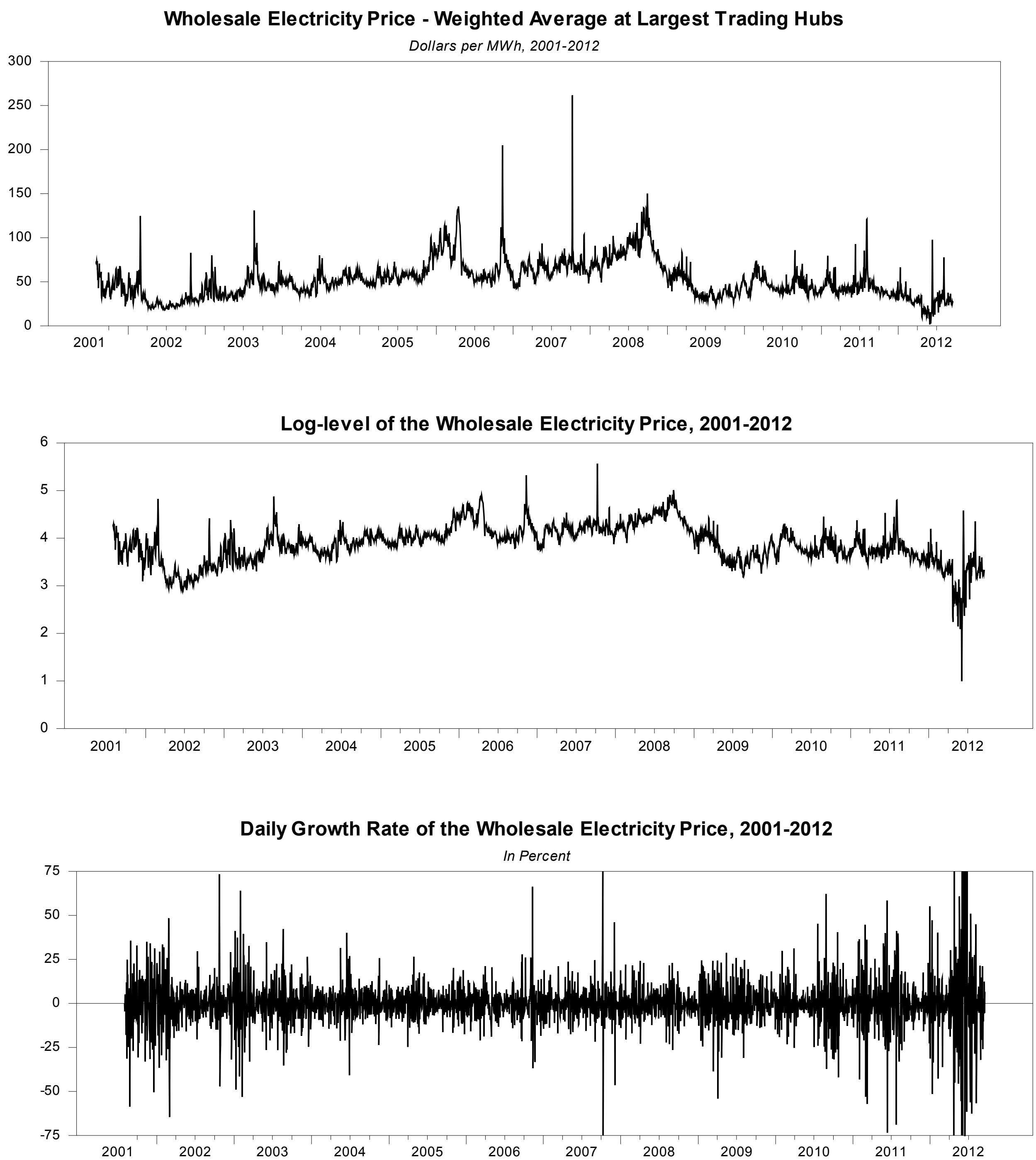
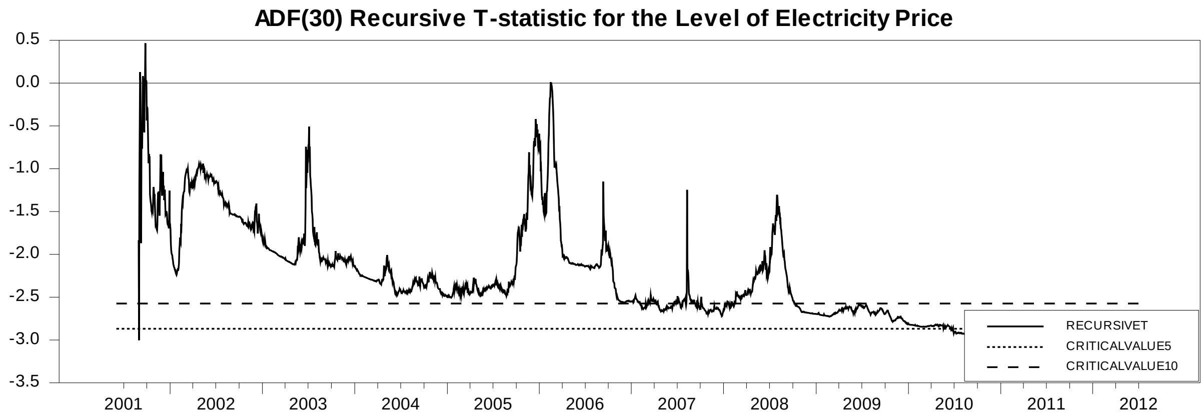


Figure 4.2: Recursive Augmented Dickey-Fuller Test Statistic – Electricity Price



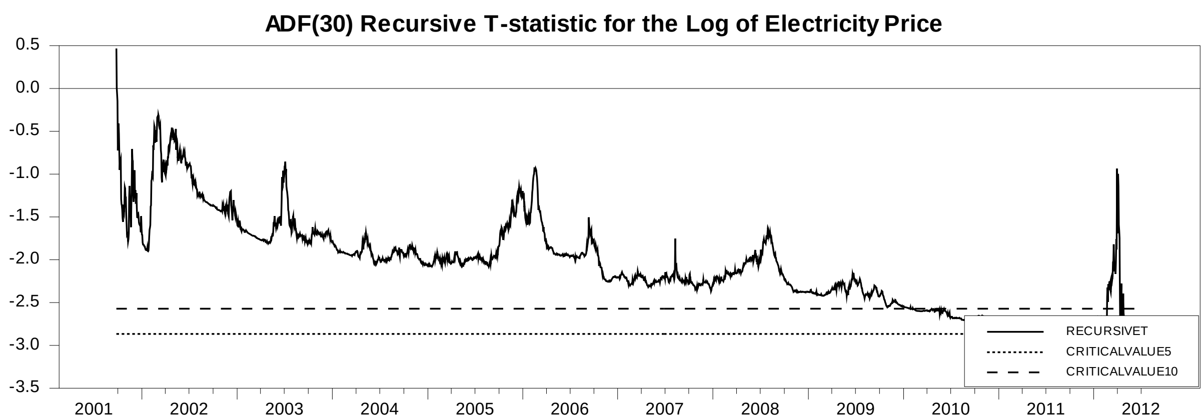


Figure 4.3: Residual Density Diagnosic Plots for Univariate GARCH Models of the Electricity Price

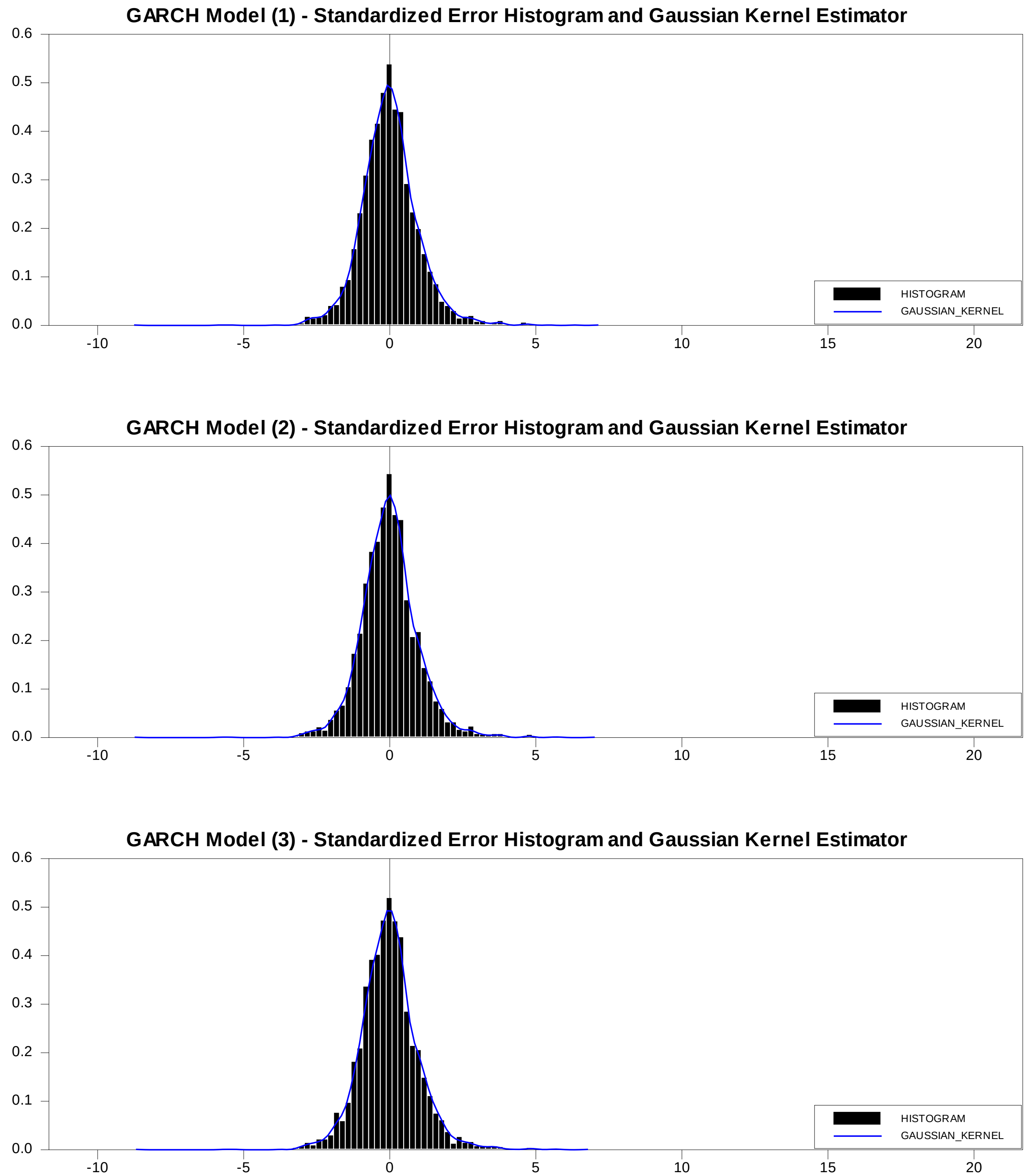
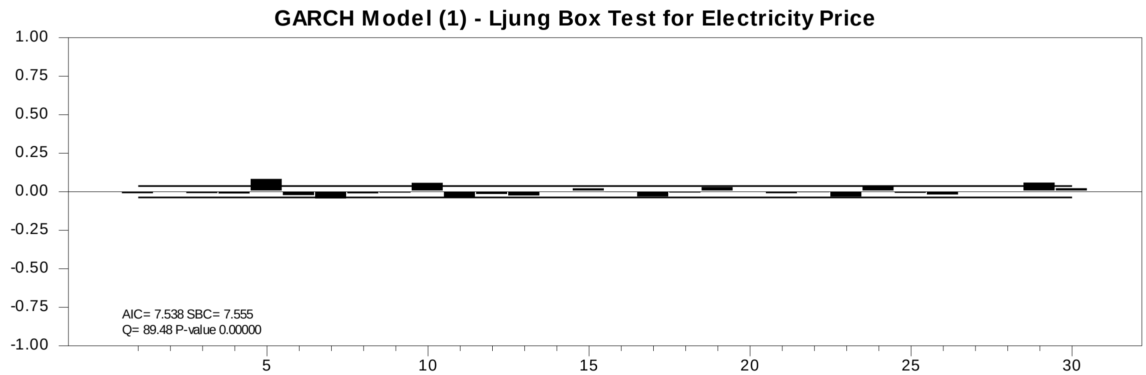
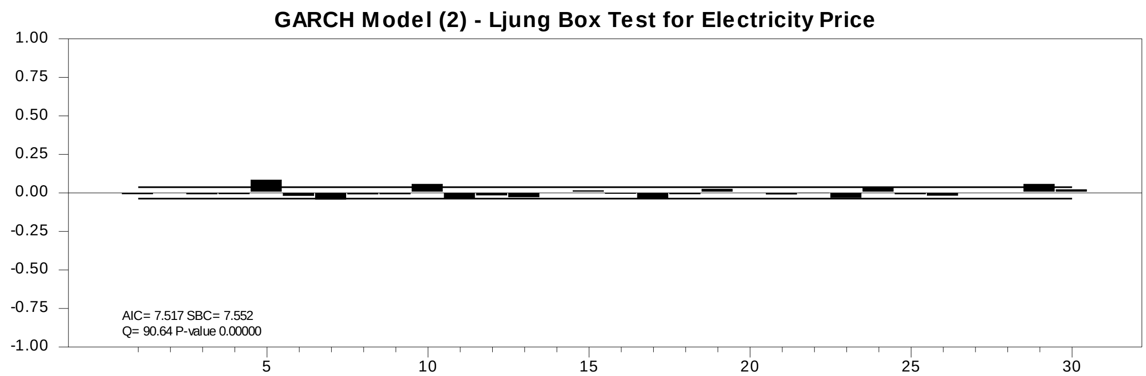


Figure 4.4: Residual Autocorrelation Tests for Univariate GARCH Models of the Electricity Price





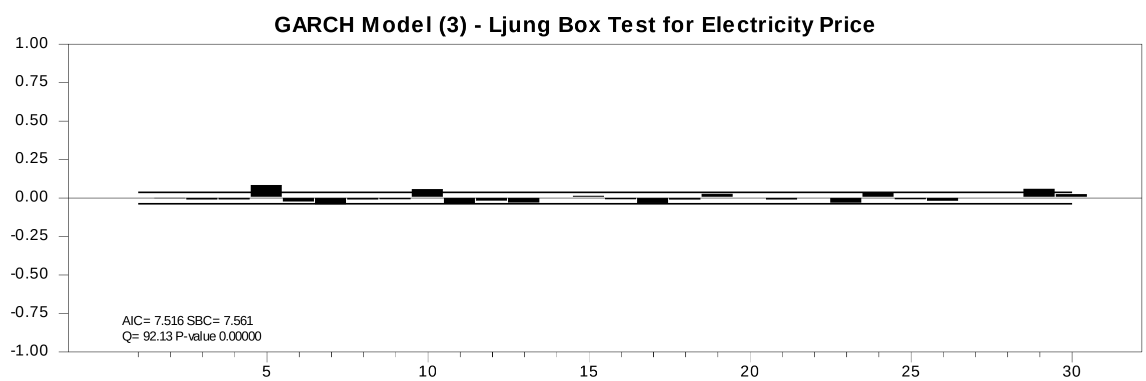
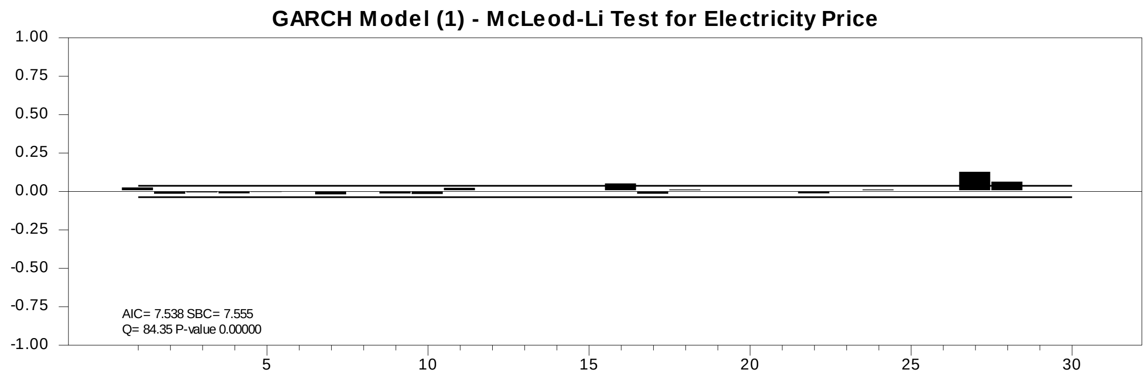
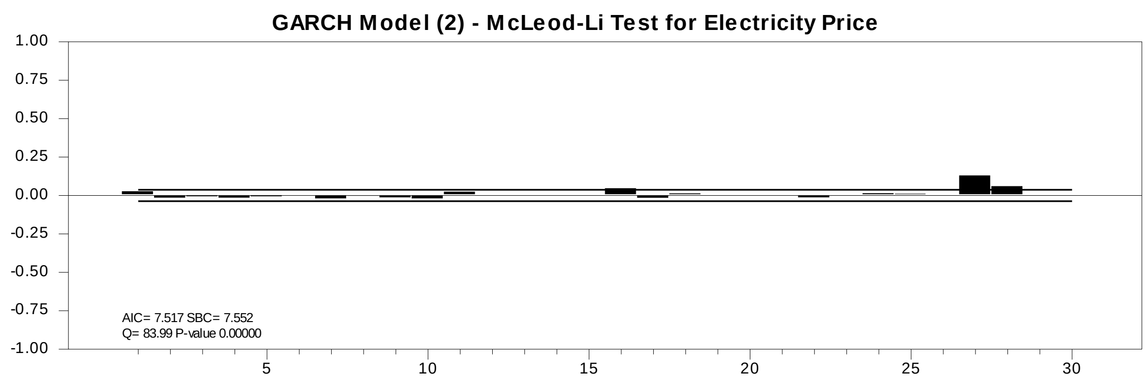


Figure 4.5: Squared Residual Autocorrelation Plots for Univariate GARCH Models of the Electricity Price





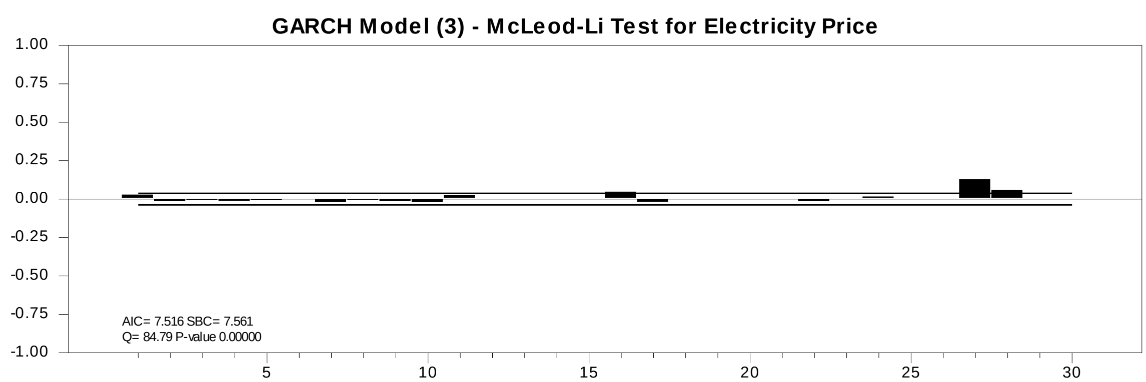


Figure 4.6: 20-Day Forecasts for Daily Change in Electricity Price by Univariate GARCH Models

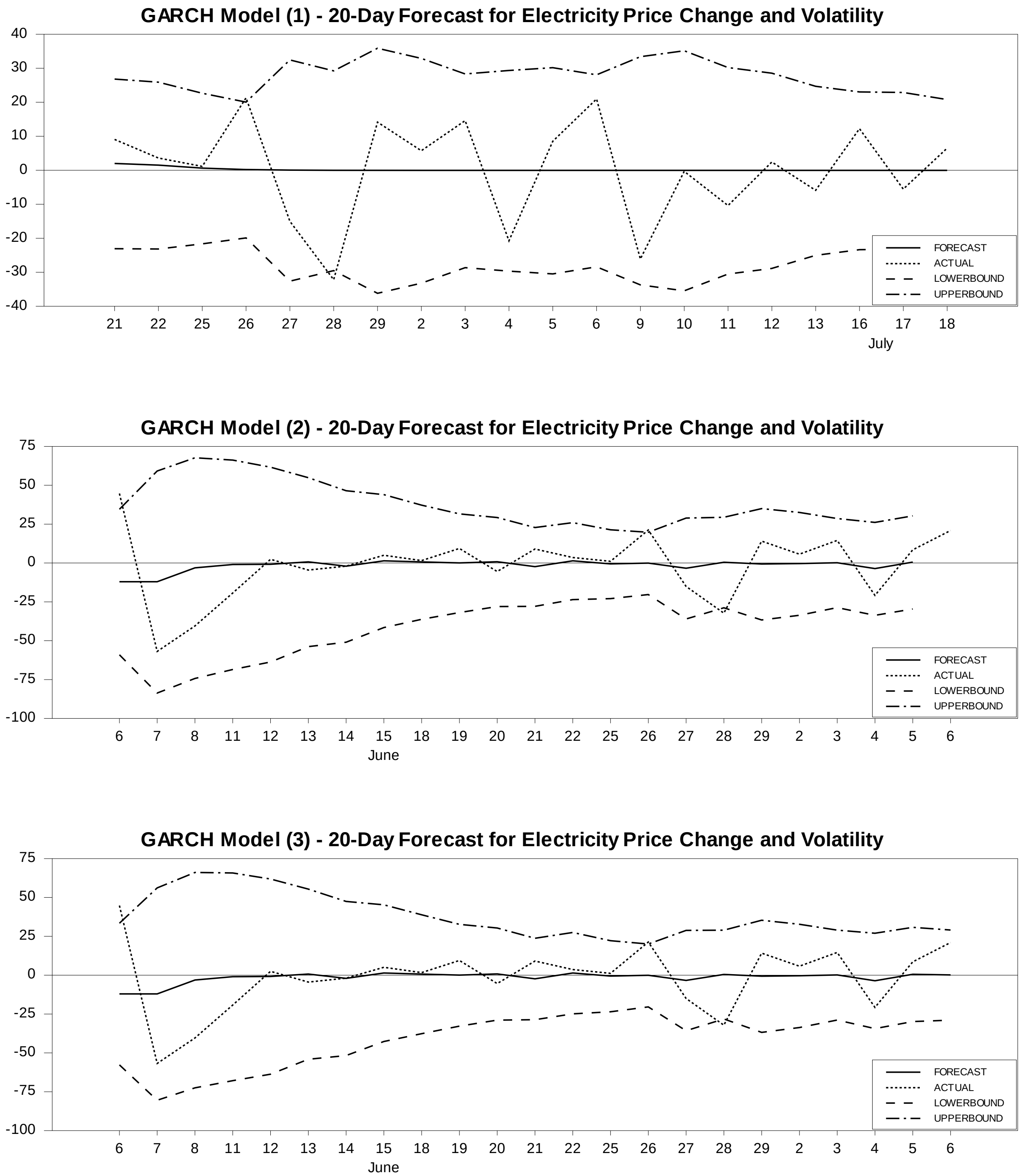
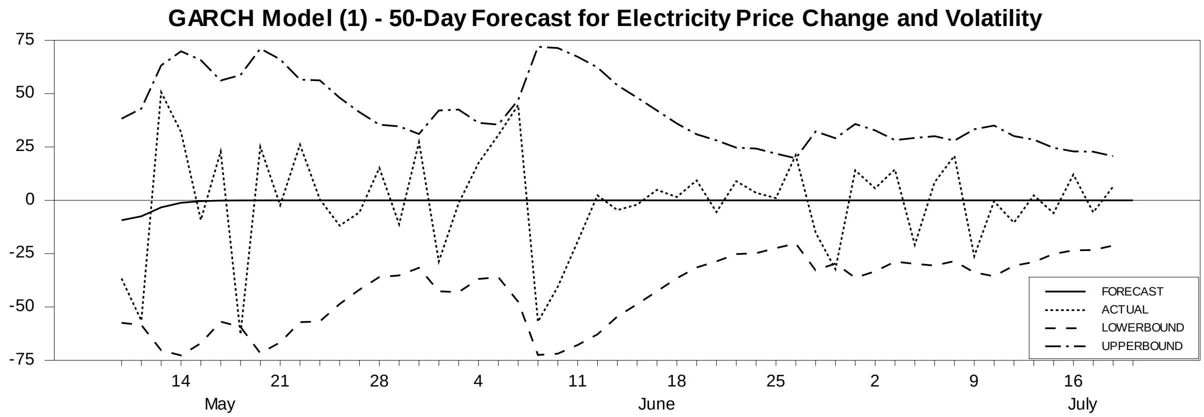
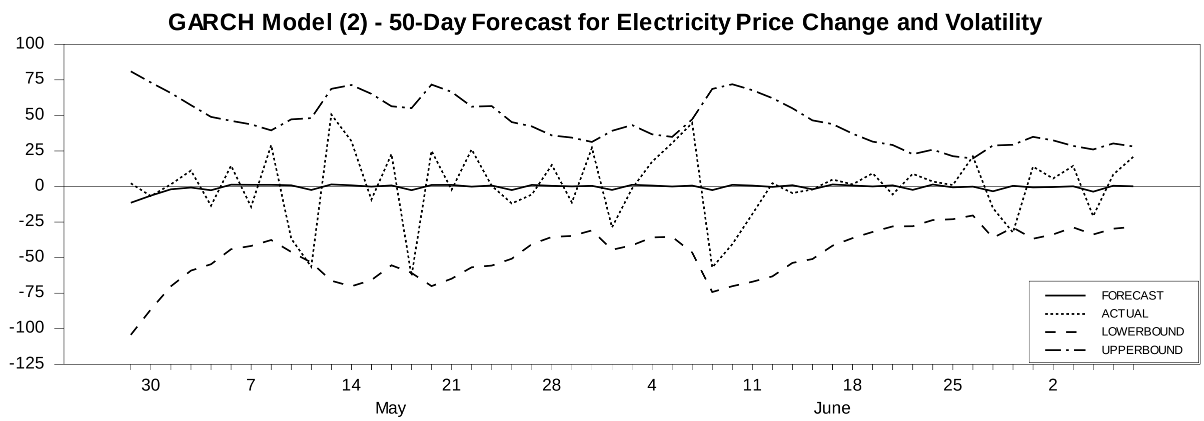
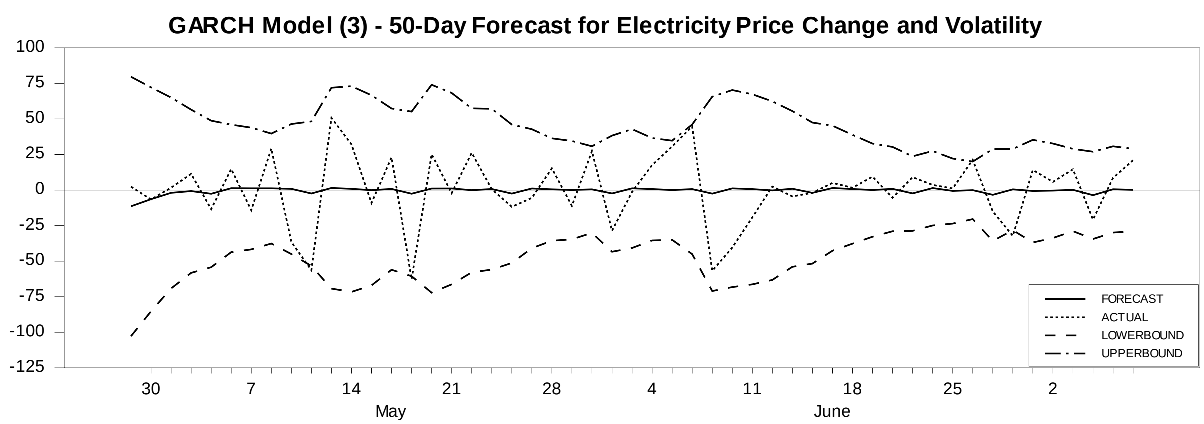


Figure 4.7: 50-Day Forecasts for Daily Change in Electricity Price by Univariate GARCH Models







5 Multivariate GARCH Models.

5.1 Introduction

Multivariate GARCH (MGARCH) models are becoming standard in finance and energy economics. Combined with a vector autoregressive (VAR) or vector error-correction (VEC) model for the mean equation, they allow for rich dynamics in the variance-covariance structure of series, making it possible to model spillovers in both the values and the conditional variances of series under study. MGARCH models also treat all variables as potentially endogenous, and combining them with a VEC mean equation allows one to explicitly model cointegrating relationships among series (Lütkepohl in Lütkepohl & Krätzig, 2004). Using data on oil-derived energy products, Wang and Wu (2012) find that MGARCH models produce better forecasts of individual asset volatility than univariate GARCH models. The original MGARCH model, the VECH, was proposed by Bollerslev et. al. (1988). The VECH is defined as

$$\begin{aligned} z_t &= \phi_0 + \sum_{j=1}^{s} \phi_j z_{t-j} + \varepsilon_t \qquad (5.1) \\ \varepsilon_t &\mid \Omega_{t-1} \sim (0,\ H_t) \\ vech(H_t) &= C + \sum_{j=1}^{q} B_j vech(H_{t-j}) + \sum_{j=1}^{p} A_j vech(\varepsilon_{t-j}\varepsilon'_{t-j}) \qquad (5.2) \end{aligned}$$

where 0 is the null vector; Ω_{t-1} represents the information set available in period $t-1$; $vech(.)$ denotes the column-stacking operator applied to the upper-trianguar portion of a matrix; H_t is an $n \times n$ variance/covariance matrix; B_j and A_j are square $n(n+1)/2$ matrices; and C is an $n(n+1)/2$ vector (notation from Serletis, 2012). While allowing for great flexibility, VECH has two disadvantages – the fact that H_t is positive definite only under restrictive assumptions, and the large number of parameters, which equals $n(n+1)/2 + (p+q)(n(n+1)/2)^2$. In reality, this sets the practical limit of VECH to two variables. To reduce the number of parameters, Bollerslev et. al. (1988) also suggested the diagonal VECH formulation, where all off-diagonal elements in matrices A_j and B_j are set to zero; however, this restricts each element h_{ijt} to depend only on its own past values and on the past values of ε_{ijt}. As a result, the most popular model today is BEKK (attributed to Baba, Engle, Kraft and Kroner), proposed by Engle and Kroner (1995). While the mean equation is the same as (1), the variance equation in BEKK is defined as:

$$H_t = C'C + \sum_{j=1}^{q}\sum_{k=1}^{K} B'_{kj} H_{t-j} B_{kj} + \sum_{j=1}^{p}\sum_{k=1}^{K} A'_{kj}\varepsilon_{t-j}\varepsilon'_{t-j}A_{kj} \tag{5.3}$$

where A_{kj}, B_{kj} and C are $n \times n$ matrices, with C being upper-triangular. The BEKK formulation makes H_t positive definite by construction, and has fewer parameters than the full VECH, while still allowing for volatility spillovers.

Two other popular formulations are the Constant Conditional Correlation (CCC) model of Bollerslev (1990), and the Dynamic Conditional Correlation (DCC) model of Engle (2002), which derives from the CCC (more on these models in section 5.4). GARCH volatility modelling is a rapidly evolving field, and at least two new methodologies were developed very recently. Hou and Suardi (2011) proposed a nonparametric GARCH approach to forecasting volatility. Engle and Sokalska (2012) introduced a multiplicative component GARCH model which can be used to forecast the volatilities of hundreds of series at the same time, which is especially useful in finance.

Several recent studies have used the multivariate GARCH approach to study the volatility of energy commodities; however, most of these have focused on modelling volatility spillovers among different types of oil products (Wang & Wu, 2012); spillovers between oil and the financial markets (Sadorsky, 2011, 2012); and those between oil markets and the macroeconomy (Elder & Serletis, 2010; Rahman & Serletis, 2012). There have been a few studies applying MGARCH to electricity markets (Goto & Karolyi, 2004; Worthington et. al, 2005) and to gas markets. Serletis and Shahmoradi (2006) use a bivariate VARMA-BEKK model to investigate volatility spillovers between the oil and natural gas markets in Canada. However, to the best of our knowledge, no paper has yet attempted to apply trivariate GARCH to jointly model the volatility of oil, gas and electricity prices. This chapter fills this void by constructing an asymmetric VARMA(1,1) BEKK-in-Mean model of oil, gas and electricity price volatilities in U.S. wholesale markets. MGARCH is a valuable approach in our case because volatility spillovers are expected among oil, natural gas and electricity markets: not only are the three substitutes in consumption, but also, natural gas and oil are both used as inputs in electricity generation (mid-peak and on-peak, respectively); and natural gas and oil are complements in production. The chosen formulation allows us to model the transmission of price volatility from one energy commodity to another, and estimate the effects of volatility in any of the three markets on the price of each commodity.

5.2 The Data

We use daily data from the U.S. Energy Information Administration (EIA); specifically, the spot oil West Texas Intermediate (WTI) price at Cushing, Oklahoma; the spot natural gas price at Henry Hub, Gulf Coast; and the day-ahead electricity price, which is a weighted average of prices at the five largest trading hubs – NEPOOL, Ercot, PJM, SP15 and Entergy. The time period is from January 1, 2001, to April 27, 2012. Table 5.1 presents descriptive statistics for log-levels of the series, as well as for log-differences scaled by 100. We also report the results of unit root, stationarity and cointegration tests. As discussed in sections 2.2, 3.2 and 4.2, ADF, PP and KPSS test statistics indicate taht all three series contain unit roots.

Normally, the presence of unit roots (also known as integration of order one or I(1)) would suggest logged first differences as the correct data representation in our model; however, we find evidence of cointegration among the three energy commodity prices. A system of I(1) variables is cointegrated if there exists a linear combination of them that is stationary or I(0) (Lütkepohl in Lütkepohl & Krätzig, ed., 2004, p. 88). In a VAR or VARMA framework, cointegration encourages both maximum-likelihood and OLS estimation methods to select parameters that correspond to this stationary combination, since parameters that eliminate the trends are always associated with the smallest deviations of actual observations from their predicted values. This result was formally proven by Davidson & MacKinnon (1993), who also showed that a VAR with cointegrated series produces estimates that are not only consistent, but superconsistent (converge to their true values faster than normal).

VEC models are often used with cointegrated series because they allow for an explicit analysis of cointegrating relations. However, a VAR in levels is sufficient if the latter are not the focus of study, as in our case. In fact, VAR and VEC are equivalent, as demonstrated in Lütkepohl's Applied Time Series Econometrics (Lütkepohl & Krätzig, ed., 2004, p. 88). A VAR system of order p (VAR(p)) in general form can be expressed as:

$$y_t = A_1 y_{t-1} + ... + A_p y_{t-p} + e_t \tag{5.4}$$

where y_t and e_t are n-dimensioned vectors, and A_k $(1 \leq k \leq p)$ are parameter matrices. A VEC can be obtained from the above equation by subtracting y_{t-1} from both sides and rearranging the terms by adding and subtracting $A_k y_{t-k+1}$ for each time period $2 \leq k \leq p$ from the right side of the equation:

$$\begin{aligned}
y_t - y_{t-1} &= A_1 y_{t-1} - y_{t-1} + A_2 y_{t-2} .... + A_p y_{t-p} + e_t \\
\Delta y_t &= -(I - A_1) y_{t-1} + A_2 y_{t-1} - A_2 y_{t-1} + A_2 y_{t-2} + ... + A_p y_{t-p+1} - A_p y_{t-p+1} \\
&+ A_p y_{t-p} + e_t \\
\Delta y_t &= -(I - A_1 - A_2) y_{t-1} - A_2 \qquad (5.5) \\
\Delta y_t &= -(I_n - A_1 - A_2 - ... - A_p) y_{t-1} - (A_2 + ... + A_p) \Delta y_{t-1} - ... - (A_p) y_{t-p+1} + e_t \\
\Delta y_t &= \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + ... + \Gamma_{p-1} \Delta y_{t-p+1} + e_t \qquad (5.6)
\end{aligned}$$

where $\Pi = -(I_n - A_1 - ... - A_p)$, $\Gamma_i = -(A_{i+1} + ... + A_p)$ for $i = 1, ...p-1$, and I_n is an n-dimensioned identity matrix (Lütkepohl in Lütkepohl & Krätzig, ed., 2004, p. 89).

We test for cointegration using the Engle & Granger (1987) and the Johansen (1988,

1991) methods. The first involves applying the ADF test to the residuals from the regression involving the three series. The null hypothesis of a unit root in residuals is rejected at' the 5% level when o_t or e_t is taken as the dependent variable; it is also rejected with g_t as the dependent variable if an allowance is made for a trend (g_t exhibits strong seasonality; the null hypothesis was also rejected with the introduction of seasonal dummy controls). ADF test results are reported in Table 5.2(C).

The second method, developed independently by Johansen and Stock & Watson (1988), attemps to detect the implied restrictions on an otherwise unrestricted VAR involving the series in question. An implied restriction suggests that there exists a VEC model that is equivalent to the VAR, something that can only happen in an integrated system. The Johansen method also allows to test for higher orders of integration. The *trace* version of the test evaluates the null hypothesis of r or fewer linearly independent cointegrating vectors. The *eigenvalue* version tests the null of exactly r cointegrating vectors. In our case, the trace test suggests a cointegration order of no more than two, and the eigenvalue test suggests an order of one (see Table 5.2(D)). We conclude that the series are cointegrated, which motivates us to use the VARMA in log-levels formulation for multivariate GARCH models in the following section.

5.3 The VARMA GARCH-M, asymmetric BEKK Model

5.3.1 Specification

The goal of this chapter is to build and estimate the most effective multivariate GARCH model possible for modelling the volatility of oil, gas and electricity prices. With this objective in mind, we choose a trivariate VARMA (vector autoregressive moving average) asymmetric GARCH-in-Mean specification, with logarithms of oil, gas and electricity prices forming the dependent variables in the mean equation:

$$
\begin{aligned}
y_t &= \phi_t + \sum_{i=1}^{p} \Gamma_i y_{t-i} + \Psi \sqrt{h_t} + \sum_{i=1}^{q} \Theta_i \varepsilon_{t-i} + \Lambda_t z_t + \varepsilon_t \\
\varepsilon_t &\mid \Omega_{t-1} \sim (0,\ H_t)
\end{aligned}
\tag{5.7}
$$

where Ω_{t-1} is the information set available in period $t-1$, and

$$
\begin{aligned}
y_t &= \begin{bmatrix} o_t \\ g_t \\ e_t \end{bmatrix}; \quad h_t = \begin{bmatrix} h_{oo,t} \\ h_{gg,t} \\ h_{ee,t} \end{bmatrix}; \quad z_t = \begin{bmatrix} j_t \\ j_{t-1} \end{bmatrix}; \\
\Gamma_i &= \begin{bmatrix} \gamma_{11}^i & \gamma_{12}^i & \gamma_{13}^i \\ \gamma_{21}^i & \gamma_{22}^i & \gamma_{23}^i \\ \gamma_{31}^i & \gamma_{32}^i & \gamma_{33}^i \end{bmatrix}; \quad \Psi = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} \\ \psi_{21} & \psi_{22} & \psi_{23} \\ \psi_{31} & \psi_{32} & \psi_{33} \end{bmatrix}; \\
\Theta_i &= \begin{bmatrix} \theta_{11}^i & \theta_{12}^i & \theta_{13}^i \\ \theta_{21}^i & \theta_{22}^i & \theta_{23}^i \\ \theta_{31}^i & \theta_{32}^i & \theta_{33}^i \end{bmatrix}; \quad \Lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \\ \lambda_{21} & \lambda_{22} \\ \lambda_{31} & \lambda_{32} \end{bmatrix};
\end{aligned}
$$

Vector z_t contains additional explanatory variables, which include the log-difference of the Dow Jones index j_t with one lag. The oil, gas and electricity prices appear in log-level form due to cointegration between them (see discussion in section 1.2); however, variables representing the trending Dow Jones index must be differenced because they are "unmodelled" components of the VAR that appear only as explanatory variables.

Introducing each additional explanatory variable into a VAR-type model adds n parameters. For this reason, we are not able to add as many additional explanatory variables as in the univariate models of preceding chapters, but we still include the Dow Jones index as we found it to have the most predictive power in modelling energy prices, compared to variables representing days of the week, season and weather. The Schwarz information criterion selects

$p = 1$ and $q = 1$ as the optimal numbers of autoregressive and moving-average lags. For the variance equation, we use the asymmetric version of the BEKK model introduced by Grier et. al. (2004). We choose the BEKK(1,1,1) specification, which is a multivariate extension of GARCH(1,1). The resulting variance equation is:

$$H_t = C'C + B'H_{t-1}B + A'\varepsilon_{t-1}\varepsilon'_{t-1}A + D'u_{t-1}u'_{t-1}D \tag{5.8}$$

where C is an upper-triangular matrix to ensure the positive definiteness of H; C, A, B and D are 3×3 matrices representing the intercepts, ARCH and GARCH parameters and asymmetry coefficients, respectively. The asymmetry vector is defined as $u_{t-1} = \varepsilon_{t-1} \circ I_{\varepsilon<0}\varepsilon_{t-1}$, where $\circ$ denotes elementwise product of vectors (Grier et. al., 2004). Assuming matrix H is symmetric, the model produces six unique equations modeling the dynamic variances of oil, gas and electricity prices, as well the covariances between them.

5.3.2 Empirical Estimates

Equations (4) and (5) are estimated by using quasi-Maximum Likelihood. We used BFGS (Broyden, Fletcher, Goldfarb & Shanno, described in Press et. al. (2007) estimation algorithm, which is recommended for GARCH models (Estima RATS (2012), p. UG-115). To obtain faster convergence, we combine BFGS with the derivative-free Simplex pre-estimation method in Estima RATS. Table 5.3 reports the coefficients obtained (significance levels in parentheses), as well as key diagnostics for standardized residuals $\hat{z}_{jt} = \frac{e_{jt}}{\sqrt{\hat{h}_{jt}}}$ for $j = o,\ g,\ e$.

The AR(1) coefficients in matrix Γ_1 are very close to one along the main diagonal, suggesting that yesterday's price of one of the three assets is a very good predictor of its price today. However, this effect is weaker for electricity, which is also the only asset that experiences significant price spillovers from the other two markets. 83.2% of the current electricity price is explained by its own price in the previous period, 11.6% is explained by the natural gas price, and 3.2% by the oil price in the previous period, amounting to significant spillover effects. On the other hand, the MA(1) coefficients along the diagonal of matrix Θ_1, are largely insignificant, suggesting that the dynamics of energy commodity prices do not correspond to a typical ARMA (although the Schwarz Information Criterion still chooses ARMA(1,1) over various AR specifications). One interesting result from the Θ_1 matrix is the significant impact of a surprise change in the oil price on gas and electricity prices in the next period. For example, a 10% unexplained increase in the oil price is expected to raise the gas price by 3.0% and the electricity price by 2.4% in the next period. However, fluctuations in the gas and electricity markets have very little effect on the oil market.

The same direction of influence (from oil to gas and electricity markets but not vice versa) is also evident in the GARCH-in-Mean coefficient matrix Ψ. Volatility in the oil market has a large impact on prices in all three markets. A 10% increase in the oil price variance is associated with an 84% increase in the oil price, 8.6% increase in the gas price and a 33% decrease in the electricity price. At the same time, the oil market is only mildly affected by volatility in the gas market (a 10% increase in the gas price variance leads to a 1% decrease in the oil price), and not at all by that in the electricity market. On an intuitive level, it is obvious that dynamics in the oil market significantly affect other markets and indeed the entire economy; however, the degree of impact resulting from daily wholesale oil price fluctuations is staggering. One possible explanation could be that speculative activity by financial traders contributes to periods of volatility in both the oil market, and the other two energy commodity markets. The impact of the former is undeniable - according to the deputy editor of the Council of Foreign Relations Toni Johnson, “Increasingly speculative behavior by a more diverse set of investors outside the oil industry - including hedge funds, pension funds, and investment banks - has made oil-market trends harder to predict” (Johnson, 2011). Substitution among the three commodities, in response to daily price and volatility changes, is unlikely over the very short time horizons of one to several days represented in the BEKK model, due to technological constraints.

The estimates for the variance equation show high GARCH coefficients along the main diagonal of matrix B, suggesting that volatility is persistent in all three markets. ARCH coefficients in matrix A indicate that gas and electricity price volatilities are both somewhat "spiky", but the oil price volatility is not. Moderate volatility contagion happens across all three markets, with no obvious uni-directional mechanisms like the one described above. Matrix D shows that asymmetric ARCH effects arise from the electricity and gas markets. A 10% unexpected positive shock in the electricity price increases the volatility in the electricity market by an additional 1%, in the gas market by 2.5% and in the oil market by 2.2%. A 10% unexpected increase in the gas price raises gas market volatility by 1.8%, but has little effect on the other markets. These asymmetric effects add on to the regular ARCH coefficients only for positive and not for negative price shocks.

Overall, the VARMA-BEKK model shows significant interactions between the three wholesale energy commodity markets, including spillovers from volatility of one asset to a price change in another asset. Thanks to the large dataset, we are able to not only detect these spillover effects, but also estimate their magnitude. BEKK is arguably the most powerful GARCH formulation for a trivariate case such as ours; however, we also build and estimate CCC and DCC models in section 5.4, primarily to compare forecasting performance and assess whether the above results are robust to changes in model specification.

5.4 Additional Models

5.4.1 The VARMA GARCH-M, Constant Conditional Correlation Model

The constant conditional correlation (CCC) model was developed by Bollerslev (1990) as another response to the excessive number of parameters in the original VECH model. The CCC model restricts conditional correlations ρ among the elements of ε_t to be constant; however, the standard deviations of the series forming the system still change, which allows their variances and covariances to change over time. The variance-covariance matrix H_t is defined as $H_t = S_t \rho S_t$, where S_t is an $n \times n$ diagonal matrix containing the conditional standard deviations $\sqrt{h_{ii,t}}$ of the elements in ε_t. The covariance between series i and j can be expressed as:

$$h_{ij,t} = \rho_{ij,t}\sqrt{h_{ii,t}}\sqrt{h_{jj,t}} \tag{5.9}$$

The CCC model is less computationally demanding than BEKK. If both models are estimated correctly, we should get similar results; the only information lost in the CCC specification is the dynamic change in the relationships among oil, gas and electricity price volatilites. For example, if the feedback from one market to another has weakened between 2001 and 2012, we wouldn't be able to tell. We use the same mean equation as for the BEKK model, but define the CCC variance equation as follows:

$$\begin{aligned} H_t &= S_t \rho S_t && (5.10) \\ diag\,(H_t) &= C + B\,diag\,(H_{t-1}) + A\,diag(e_{t-1}e'_{t-1}) + D'u_{t-1}u'_{t-1}D && (5.11) \end{aligned}$$

where *diag* is an operator that extracts the diagonal from a square matrix, C and D are 3×1 vectors, and A and B are a 3×3 matrices. The CCC model produces three equations modelling the conditional variances of oil, gas and electricity prices. We can easily obtain the covariances between any two of these assets with equation (5.6), which uses conditional correlations estimated by RATS and reported in matrix R.

We use the same VARMA(1,1) mean equation (5.6) as for the BEKK specification. Like before, we combine the BFGS estimation method with a set of Simplex pre-iterations. The resulting estimates are reported in Table 5.4. The coefficients in the mean equation are predictably similar to those obtained with BEKK. The main difference is that the GARCH-in-mean coefficients are small and insignificant in CCC whereas they are large and significant in BEKK. Also, the Dow Jones coefficients are insignificant in CCC, although they were not large in BEKK. The fact that Dow Jones had a powerful predictive effect in univariate models but not in a VARMA model is possibly due to the fact that it was capturing the interdependence of the three energy commodities, each of which is also related to the general level of economic activity, rather than an actual relationship between a particular energy commodity and the macroeconomy.

In the variance equation, the CCC specification yields considerably greater long-term volatility spillover effects, evidenced by large and significant off-diagonal elements in matrix B. For example, a 100% increase in the variance of oil price is predicted to cause a 39% increase in the variance of gas price, and a 100% increase in the variance of gas price yields a 37% increase in that of the electricity price. Another difference is that the CCC produces a large asymmetric effect in the oil market - an unexpected 10% dip in the oil price is predicted to increase oil market volatility by an extra 1.3%, in addition to the increase predicted by other variance equation coefficients. By contrasting large coefficients in matrix B with smaller ones in matrix A, we once again conclude that energy volatility is very persistent, especially in the oil market.

The estimated conditional correlations are realistic; the largest of these is a 29.7% correlation between gas and electricity volatilities. Overall, the CCC model yields sensible results while requiring fewer parameters to be estimated than with BEKK. One might still question the realism of the assumption of constant conditional correlations during the 2000-2012 period, which has included important events that affected each energy market individually, such as the shale gas boom, the conflict in the Middle East, and the deregulation of U.S. electricity markets. The next sub-section estimates a dynamic conditional correlation (DCC) model of Engle (2002), which may provide a better fit to the data while still using a limited number of parameters.

5.4.2 The VARMA GARCH-M, Dynamic Conditional Correlation Model

The Dynamic Conditional Correlation (DCC) model was suggested by Robert Engle in 2002. The DCC creates an ingenious way to make conditional correlations among series time-varying using only two additional parameters, regardless of the number of series in the model (Engle, 2002). The first step in estimating a DCC model is to obtain conditional correlations from the covariance matrix Q_t, which is typically estimated with a "GARCH

(1,1)" equation governed by two scalar parameters a and b:

$$Q_t = (1 - a - b)Q_0 + a\varepsilon_{t-1}\varepsilon'_{t-1} + bQ_{t-1} \tag{5.12}$$

where Q_0 is the unconditional covariance matrix (Engle, 2002). The matrix Q_t does not replace H_t; its sole purpose is to provide conditional correlations $\sqrt{Q_{ij,t}}$, $i \neq j$. The H_t matrix is generated by fitting univariate GARCH models to estimate the variances, and combining these variances with $\sqrt{Q_{ij,t}}$ to estimate the covariances. The process is summarized as:

$$H_{ij,t} = \frac{Q_{ij,t}\sqrt{H_{ii,t}}\sqrt{H_{jj,t}}}{\sqrt{Q_{ii,t}Q_{jj,t}}} \tag{5.13}$$

Because the elements of Q_t are required to equal unity along the diagonal, it is clear that variances in H_t do not depend on Q_t (Engle, 2002).

The DCC model is very useful in our study of energy commodity volatility, because given the significant shocks that exogenously affected oil, gas and electricity markets during the past decade, it is implausible to assume that the conditional correlations among the three assets remained constant. When we estimate the DCC, we use a "VARMA" specification for the variances in H_t:

$$H_{ii,t} = c_{ii} + \sum_{j=1}^{3} a_{ij}\varepsilon^2_{j,t-1} + \sum_{j=1}^{3} b_{ij}H_{jj,t-1} + d_{ii}u^2_{t-1} \tag{5.14}$$

where the last term represents the asymmetry coefficient. This specification allows for spillovers among the variances of the three series, and also makes the form almost identical to that used for the CCC and BEKK models, allowing for direct comparisons of model performance. The main disadvantage of the DCC model is that there is no simple extrapolative technique for making forecasts. From equation (5.9), we can see that the current conditional correlations depend on the last period's error vector from the mean equation; because the expected value of each error term is zero, the best we can do is use today's errors to make one-day-ahead forecasts. This is the reason why we do not report DCC forecasts in this chapter.

Table 5.5 presents the empirical estimates obtained from the DCC model. These are similar to those obtained with BEKK and especially CCC. Like the latter, DCC finds insignificant GARCH-in-mean and Dow Jones coefficients in the mean equation, as well as a strong asymmetric effect in the oil market variance equation. However, the pattern of long-term variance spillover effects is more similar to that obtained with the BEKK model. The DCC model produces significant short-term spillovers in volatility, evidenced by large and significant off-diagonal elements in matrix A. Once again, DCC reinforces the finding that all three markets experience persistent volatility, especially the oil market. In comparison, electricity market volatility is the most "spiky".

5.5 Forecasting

Volatility forecasting is arguably the most important application of GARCH models. A complex VARMA model such as ours incorporates information about past prices and variances of all three commodities in every forecast. Furthermore, the forecasts are dynamic in the sense that a price and volatility forecast for a particular period are used as "current" values in forecasting for the next period. We produce 20- and 40-day forecasts for illustrative purposes (see Figures 5.4 and 5.5), as well as forecast performance statistics based on rolling 20-day forecasts for the entire period of January 2001 to May 2012. Based on a visual analysis of the graphs, BEKK produces slightly better volatility forecasts than the CCC, although the two are similar. Forecasting is not possible using the DCC model, which lacks an extrapolative formula. The volatility of the electricity price is the most amenable to forecasting - for example, Figure 5.5(c) shows both BEKK and CCC models correctly identifying periods of high and low volatility. On the other hand, the highly persistent nature of oil price volatility yields 95% confidence bounds that change only slightly over time (see Figure 5.5(a)).

While no formal statistics exist to compare the quality of volatility forecasts, we can objectively evaluate the performance of the VARMA GARCH-in-Mean model specification in forecasting future prices. Table 5.6 reports forecast Mean Error (ME), Mean Absolute Error (MAE), Root Mean Square Error (RMSE) and Theil's U statistics (Theil, 1971) for both BEKK and CCC specifications, as well as for a benchmark random walk with drift model. Theil's U statistic is simply the ratio of RMSE from the model under question to the RMSE of a "no-change" forecast. If $U > 1$, then our model produces forecasts that are inferior to this "naive" benchmark, while if $U < 1$, our model produces superior forecasts. We find that over a very short forecast horizon (e.g. one day), neither BEKK nor CCC outperform a simple random walk with drift model; however, as the forecast horizon increases (e.g. to 20 days), our models yield better forecasts for gas and electricity prices (see Table 5.6). BEKK produces the best electricity price forecasts, while the random walk model yields the best oil price forecasts even with a 20-day horizon.

5.6 Conclusion

This chapter presents three multivariate GARCH models that jointly model the price volatilities in wholesale oil, natural gas and electricity markets, as well as covariances between the markets. BEKK is our model of choice because, being the least restrictive, it allows us to identify many powerful and unexpected spillover effects among the three energy markets. For example, we find asymmetric price-to-volatility spillover effects between natural gas and electricity markets, and strong impacts from the oil price, its unexpected change, and its variance, to natural gas and electricity prices and their variances. Also, several differences emerge between multivariate and univariate models. The MA coefficients in the multivariate mean equation are estimated as small and sometimes insignificant by both BEKK and CCC, although they were consistently large and significant in univariate models. On the other hand, the GARCH-in-Mean effects are much more pronounced in multivariate models, especially BEKK. These include significant volatility-to-price spillovers: in fact, BEKK and DCC estimates suggest that oil price volatility has a larger imact on gas and electricity prices than their own volatilities.

Although univariate models produce more accurate forecasts, especially for the mean equation, multivariate formulations presented in this chapter allow us to uncover fascinating dynamics and interactions among energy markets.

Table 5.1: Summary Statistics

A. Summary Statistics in log levels

Variable	Mean	Variance	Skewness	Excess kurtosis	J-B normality
o_t	3.988	0.250	-0.266	-1.071	166.9
g_t	1.634	0.170	-0.089	-0.152	6.37
e_t	3.906	0.136	0.032	0.071	1.10

B. Summary Statistics in logged first differences

Variable	Mean	Variance	Skewness	Excess kurtosis	J-B normality
Δo_t	0.048	6.694	-0.090	5.043	2969
Δg_t	-0.057	22.967	0.554	21.149	52290
Δe_t	-0.026	143.6	-0.065	8.430	8286

Table 5.2: Unit Root and Cointegration Tests

A. Unit root and stationarity tests in log levels

	Unit root tests		KPSS stationarity tests	
Variable	ADF	PP	$\widehat{\eta}_\mu$	$\widehat{\eta}_\tau$
o_t	-1.221	-1.282	44.48	4.838
g_t	-1.569	-2.407	8.360	8.363
e_t	-2.819	-8.217	8.066	7.149
5% cv	-2.863	-2.863	0.463	0.146
1% cv	-3.436	-3.436	0.739	0.216

B. Unit root and stationarity tests in logged first differences

	Unit root tests		KPSS stationarity tests	
Variable	ADF	PP	$\widehat{\eta}_\mu$	$\widehat{\eta}_\tau$
Δo_t	-9.17	-54.36	0.034	0.034
Δg_t	-10.92	-53.14	0.073	0.064
Δe_t	-11.97	-81.72	0.010	0.01
5% cv	-2.863	-2.863	0.463	0.146

C. Engle-Granger Cointegration Test

Dependent Variable	ADF - with trend	5% cv with trend	ADF - no trend	5% cv no trend
o_t	-4.64	-4.12	-5.37	-3.75
g_t	-4.00	-4.12	-1.72	-3.75
e_t	-7.61	-4.12	-7.39	-3.75

D. Johansen Cointegration Tests

Cointegration Order	Eigenvalue	Trace statistic	5% critical value (trace)	Max-Eigenv. statistic
At most 1	0.0189	56.46	29.80	52.70*
At most 2	0.0012	3.767*	15.41	3.318
At most 3	0.0002	0.449	3.84	3.84

Table 5.3: The trivariate VARMA(1,1) asymmetric BEKK Model with daily U.S. oil, natural gas and electricity price data

A. Conditional mean equation

$$\phi = \begin{bmatrix} -0.227 \\ (0.000) \\ -0.009 \\ (0.513) \\ 0.417 \\ (0.000) \end{bmatrix}; \quad \Gamma_1 = \begin{bmatrix} 0.994 & -0.001 & 0.008 \\ (0.000) & (0.770) & (0.055) \\ 0.009 & 0.988 & -0.006 \\ (0.001) & (0.000) & (0.295) \\ 0.032 & 0.116 & 0.832 \\ (0.000) & (0.000) & (0.000) \end{bmatrix}; \quad \Lambda = \begin{bmatrix} -0.054 & 0.087 \\ (0.258) & (0.068) \\ -0.099 & 0.012 \\ (0.094) & (0.846) \\ -0.039 & 0.038 \\ (0.790) & (0.811) \end{bmatrix};$$

$$\Psi = \begin{bmatrix} 8.401 & -0.110 & 0.002 \\ (0.000) & (0.002) & (0.917) \\ 0.862 & -0.129 & -0.015 \\ (0.000) & (0.036) & (0.552) \\ -3.327 & -0.064 & 0.120 \\ (0.000) & (0.438) & (0.071) \end{bmatrix}; \quad \Theta_1 = \begin{bmatrix} 0.008 & 0.013 & -0.013 \\ (0.701) & (0.313) & (0.031) \\ 0.301 & -0.036 & 0.047 \\ (0.000) & (0.131) & (0.000) \\ 0.238 & -0.013 & -0.028 \\ (0.000) & (0.741) & (0.241) \end{bmatrix};$$

B. Conditional variance-covariance structure

$$C = \begin{bmatrix} -0.013 & 0.000 & 0.004 \\ (0.000) & (0.361) & (0.000) \\ & 0.004 & -0.001 \\ & (0.000) & (0.102) \\ & & 0.013 \\ & & (0.000) \end{bmatrix}; \quad B = \begin{bmatrix} 0.864 & 0.058 & 0.148 \\ (0.000) & (0.000) & (0.000) \\ 0.003 & 0.918 & -0.005 \\ (0.054) & (0.000) & (0.604) \\ 0.000 & -0.004 & 0.923 \\ (0.707) & (0.096) & (0.000) \end{bmatrix};$$

$$A = \begin{bmatrix} 0.005 & -0.073 & -0.066 \\ (0.544) & (0.010) & (0.020) \\ -0.012 & 0.351 & -0.035 \\ (0.029) & (0.000) & (0.172) \\ -0.003 & 0.023 & 0.373 \\ (0.094) & (0.000) & (0.000) \end{bmatrix}; \quad D = \begin{bmatrix} -0.073 & 0.037 & 0.216 \\ (0.000) & (0.370) & (0.000) \\ 0.028 & 0.180 & 0.251 \\ (0.000) & (0.000) & (0.000) \\ -0.007 & 0.044 & -0.097 \\ (0.063) & (0.000) & (0.000) \end{bmatrix};$$

C. Residual Diagnostics

	Mean	Variance	$Q(40)$	$Q^2(40)$	$Q(100)$	$Q^2(100)$
z_{ot}	−0.002	0.996	48.11 (0.177)	57.60 (0.035)	98.72 (0.517)	132.9 (0.016)
z_{gt}	0.023	0.967	85.90 0.000	21.13 0.993	330.3 (0.000)	38.56 (1.00)
z_{et}	0.040	0.978	210.7 0.000	68.20 (0.004)	330.4 (0.000)	127.3 (0.034)

Table 5.4: The trivariate VARMA(1,1) asymmetric CCC Model with daily U.S. oil, natural gas and electricity price data

A. Conditional mean equation

$$\phi = \begin{bmatrix} 0.0031 \\ (0.692) \\ 0.0185 \\ (0.143) \\ 0.3484 \\ (0.000) \end{bmatrix}; \quad \Gamma_1 = \begin{bmatrix} 0.999 & -0.001 & 0.001 \\ (0.000) & (0.822) & (0.814) \\ 0.001 & 1.004 & -0.007 \\ (0.746) & (0.000) & (0.166) \\ 0.027 & 0.121 & 0.831 \\ (0.000) & (0.000) & (0.000) \end{bmatrix}; \quad \Lambda = \begin{bmatrix} -0.001 & -0.001 \\ (0.080) & (0.752) \\ -0.001 & 0.000 \\ (0.306) & (0.756) \\ -0.001 & -0.000 \\ (0.651) & (0.961) \end{bmatrix};$$

$$\Psi = \begin{bmatrix} 0.084 & -0.050 & 0.005 \\ (0.330) & (0.006) & (0.560) \\ 0.032 & 0.025 & -0.014 \\ (0.710) & (0.643) & (0.314) \\ -0.073 & 0.157 & -0.011 \\ (0.729) & (0.032) & (0.827) \end{bmatrix}; \quad \Theta_1 = \begin{bmatrix} -0.039 & 0.015 & -0.003 \\ (0.043) & (0.102) & (0.564) \\ 0.215 & -0.036 & 0.035 \\ (0.000) & (0.110) & (0.000) \\ 0.169 & -0.008 & -0.078 \\ (0.007) & (0.846) & (0.017) \end{bmatrix};$$

B. Conditional variance-covariance structure

$$C = \begin{bmatrix} 0.000 \\ (0.000) \\ 0.000 \\ (0.010) \\ 0.001 \\ (0.000) \end{bmatrix}; \quad B = \begin{bmatrix} 0.858 & 0.078 & -0.089 \\ (0.000) & (0.067) & (0.059) \\ 0.387 & 0.771 & 0.014 \\ (0.070) & (0.000) & (0.324) \\ -0.248 & 0.365 & 0.705 \\ (0.853) & (0.008) & (0.000) \end{bmatrix}; \quad D = \begin{bmatrix} 0.115 \\ (0.000) \\ -0.001 \\ (0.973) \\ -0.043 \\ (0.158) \end{bmatrix};$$

$$A = \begin{bmatrix} 0.031 & -0.021 & -0.005 \\ (0.009) & (0.003) & (0.044) \\ -0.094 & 0.185 & 0.020 \\ (0.000) & (0.000) & (0.000) \\ -0.095 & -0.162 & 0.276 \\ (0.121) & (0.000) & (0.000) \end{bmatrix}; \quad R = \begin{bmatrix} 1.000 & 0.119 & 0.059 \\ & 1.000 & 0.297 \\ & & 1.000 \end{bmatrix};$$

C. Residual Diagnostics

	Mean	Variance	$Q(40)$	$Q^2(40)$	$Q(100)$	$Q^2(100)$
z_{ot}	−0.008	1.000	50.64 (0.121)	58.05 (0.033)	102.9 (0.402)	138.2 (0.007)
z_{gt}	0.013	1.001	83.27 0.000	20.58 0.995	165.8 (0.000)	42.95 (1.00)
z_{et}	0.022	1.005	211.1 0.000	69.61 (0.003)	327.6 (0.000)	129.0 (0.027)

Table 5.5: The trivariate VARMA(1,1) asymmetric DCC Model with daily U.S. oil, natural gas and electricity price data

A. Conditional mean equation

$$\phi = \begin{bmatrix} 0.003 \\ (0.718) \\ 0.022 \\ (0.080) \\ 0.361 \\ (0.000) \end{bmatrix}; \quad \Gamma_1 = \begin{bmatrix} 0.999 & -0.001 & 0.001 \\ (0.000) & (0.796) & (0.847) \\ 0.001 & 1.005 & -0.009 \\ (0.723) & (0.000) & (0.100) \\ 0.028 & 0.126 & 0.824 \\ (0.000) & (0.000) & (0.000) \end{bmatrix}; \quad \Lambda = \begin{bmatrix} -0.001 & -0.000 \\ (0.122) & (0.766) \\ -0.001 & 0.000 \\ (0.312) & (0.952) \\ -0.001 & -0.000 \\ (0.469) & (0.741) \end{bmatrix};$$

$$\Psi = \begin{bmatrix} 0.107 & -0.034 & 0.002 \\ (0.209) & (0.035) & (0.793) \\ 0.029 & 0.033 & -0.014 \\ (0.723) & (0.520) & (0.318) \\ -0.054 & 0.149 & -0.013 \\ (0.769) & (0.033) & (0.821) \end{bmatrix}; \quad \Theta_1 = \begin{bmatrix} -0.032 & 0.012 & -0.001 \\ (0.115) & (0.157) & (0.747) \\ 0.214 & -0.029 & 0.035 \\ (0.000) & (0.192) & (0.000) \\ 0.155 & -0.018 & -0.073 \\ (0.009) & (0.652) & (0.023) \end{bmatrix};$$

B. Conditional variance-covariance structure

$$C = \begin{bmatrix} 0.000 \\ (0.000) \\ 0.000 \\ (0.000) \\ 0.001 \\ (0.000) \end{bmatrix}; \quad B = \begin{bmatrix} 0.843 & -0.064 & -0.027 \\ (0.000) & (0.113) & (0.123) \\ 0.413 & 0.783 & -0.002 \\ (0.000) & (0.000) & (0.900) \\ 1.293 & -0.058 & 0.686 \\ (0.029) & (0.468) & (0.000) \end{bmatrix}; \quad D = \begin{bmatrix} 0.130 \\ (0.000) \\ 0.001 \\ (0.960) \\ -0.047 \\ (0.127) \end{bmatrix};$$

$$A = \begin{bmatrix} 0.033 & -0.016 & -0.005 \\ (0.023) & (0.047) & (0.076) \\ -0.105 & 0.171 & 0.024 \\ (0.000) & (0.000) & (0.001) \\ -0.181 & -0.112 & 0.295 \\ (0.006) & (0.000) & (0.000) \end{bmatrix}; \quad a = 0.006; \quad b = 0.994$$

C. Residual Diagnostics

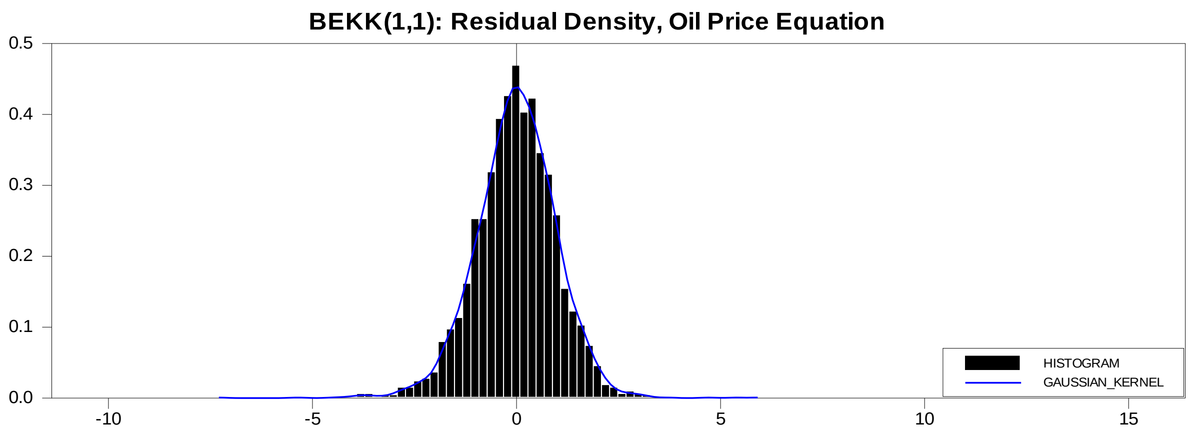
	Mean	Variance	$Q(40)$	$Q^2(40)$	$Q(100)$	$Q^2(100)$
z_{ot}	-0.016	0.998	48.66 (0.164)	53.16 (0.080)	102.6 (0.410)	136.5 (0.009)
z_{gt}	0.009	0.994	83.62 0.000	19.04 0.998	165.7 (0.000)	40.79 (1.00)
z_{et}	0.024	0.996	205.6 0.000	70.90 (0.002)	322.4 (0.000)	132.7 (0.016)

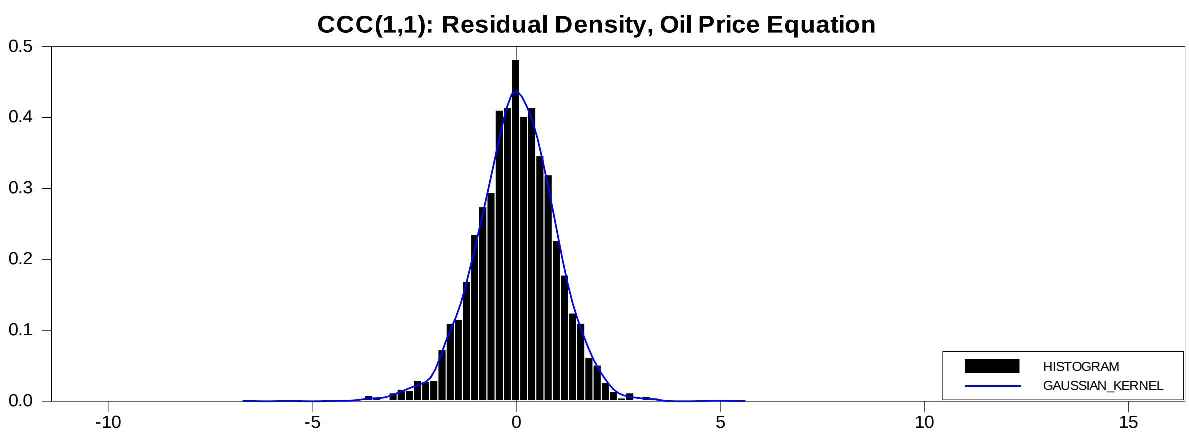
Table 5.6: Performance Statistics for 20-Day Forecasts

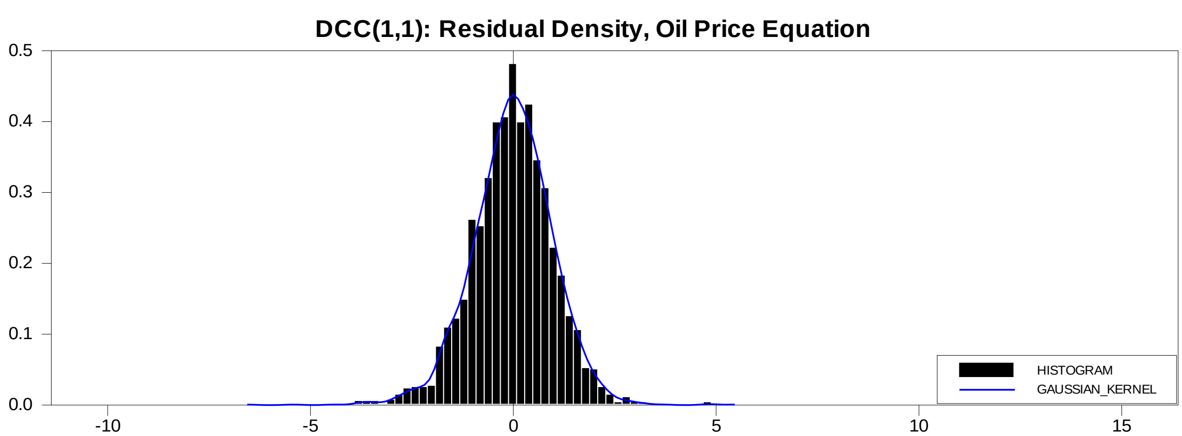
	ME (in dollars)	MAE (in dollars)	RMSE (in dollars)	Theil's U (1-step)	Theil's U (20-step)
VARMA-BEKK					
o_t	0.285	3.489	5.203	1.001	1.115
g_t	0.053	0.477	0.774	0.987	0.976
e_t	1.802	7.205	11.77	0.948	0.853
VARMA-CCC					
o_t	-0.228	3.430	5.038	1.000	1.071
g_t	0.036	0.479	0.775	0.987	0.963
e_t	-9.289	10.67	13.60	1.792	5.419
Random Walk					
o_t	0.040	3.249	4.670	1.000	0.992
g_t	0.053	0.497	0.793	0.999	0.985
e_t	1.064	7.223	11.56	0.987	1.124

Figure 5.1: Residual Density Plots for the VARMA(1,1) asymmetric GARCH-in-Mean Models

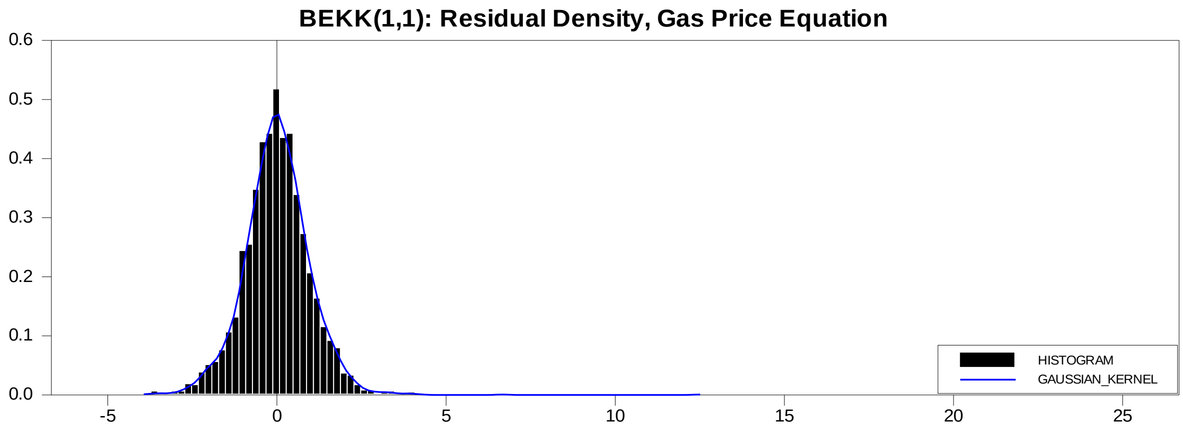
a) Oil Price Equation

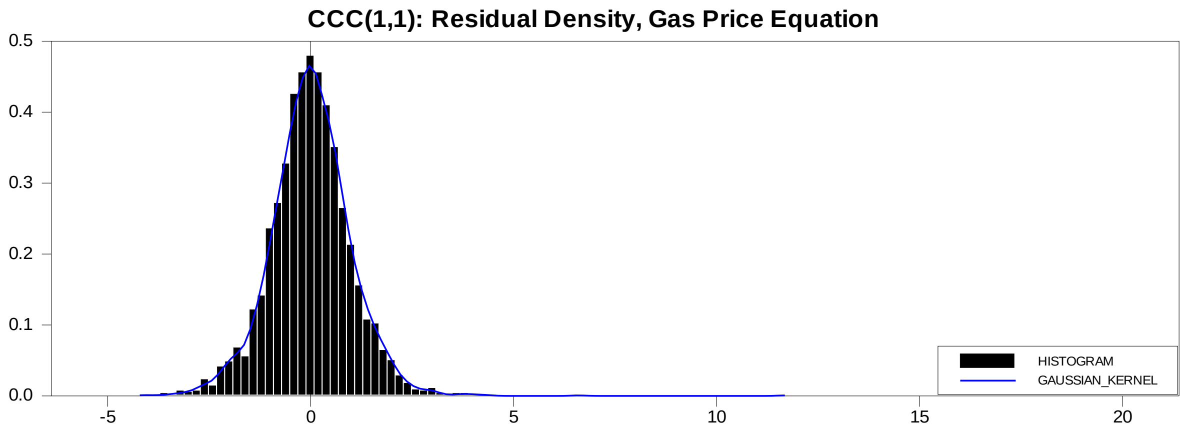


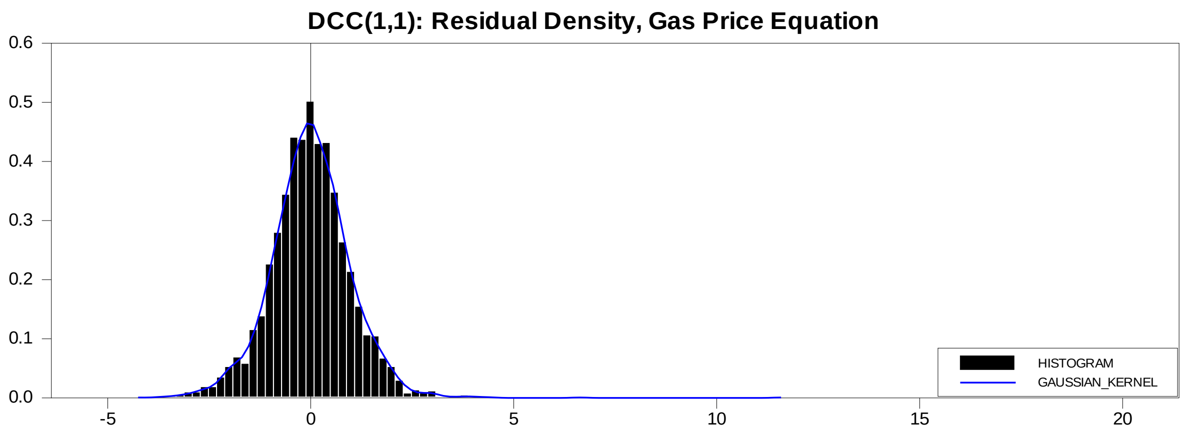




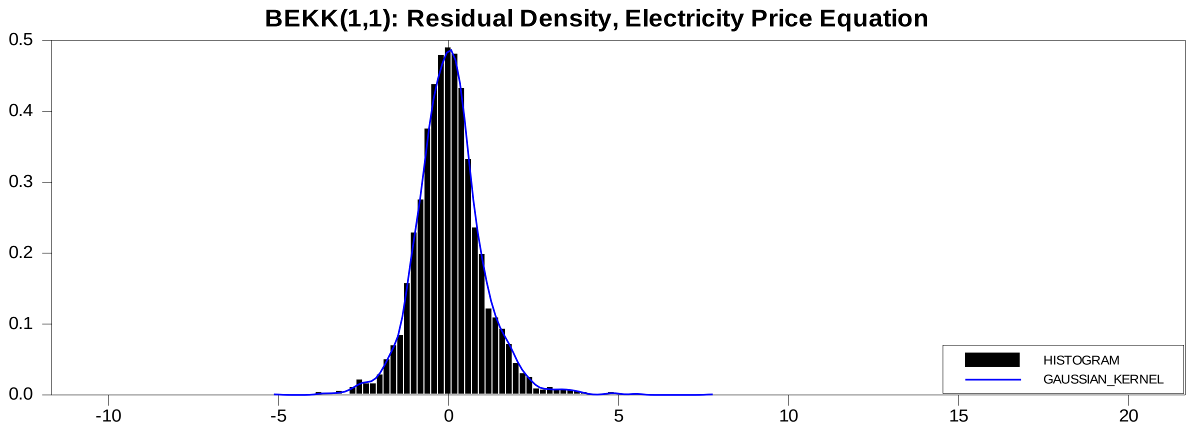
b) Gas Price Equation

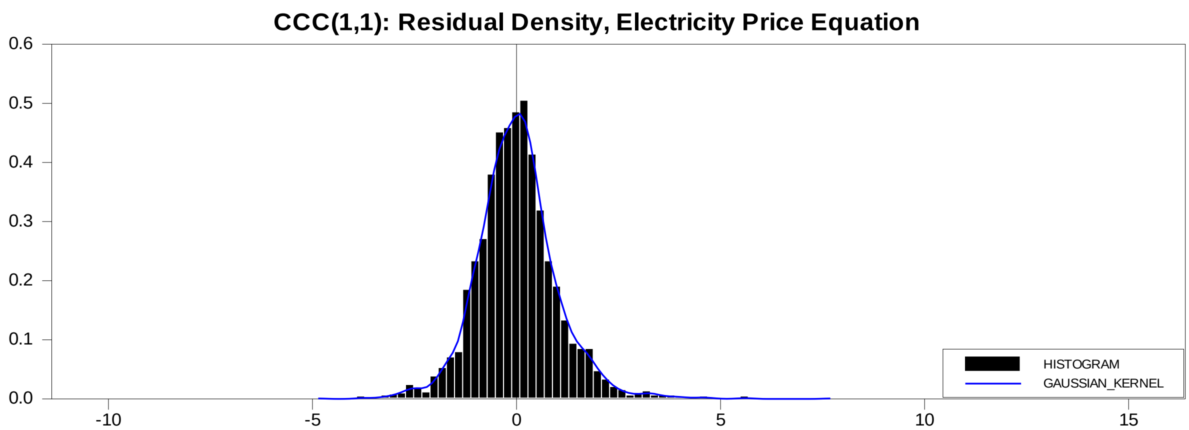






c) Electricity Price Equation





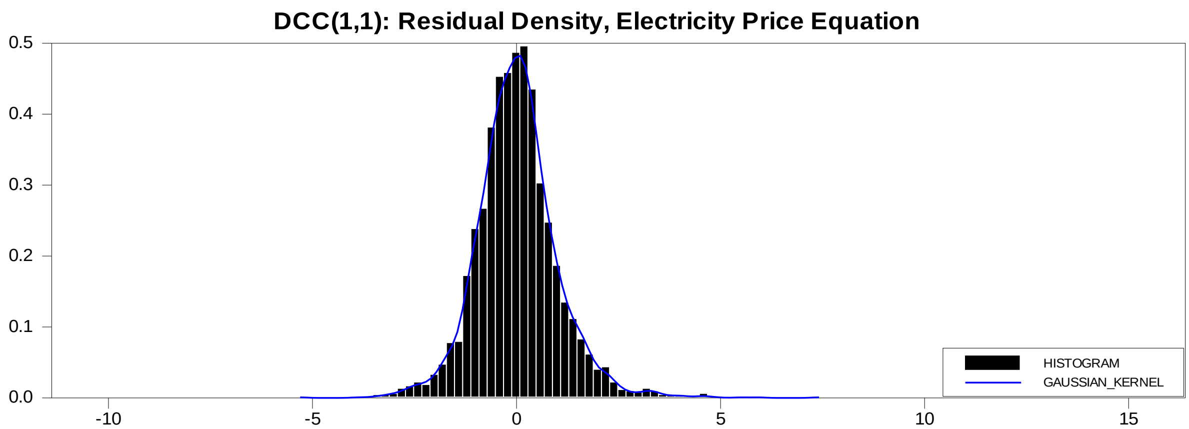
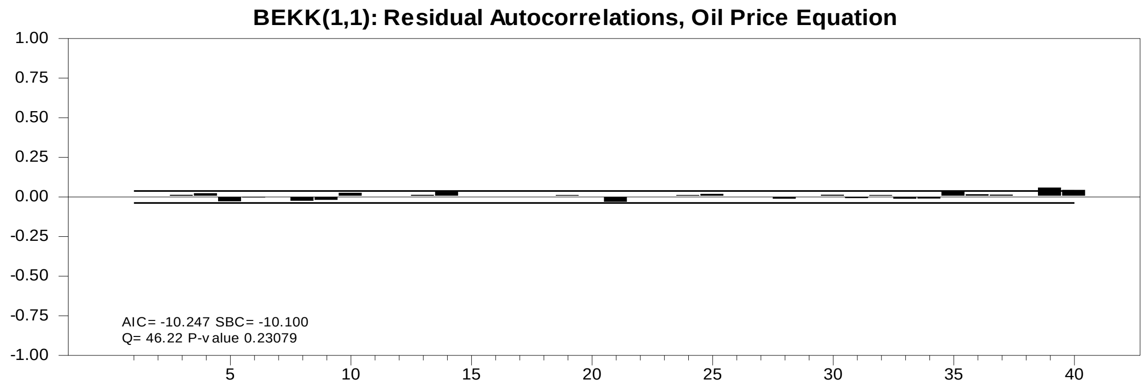
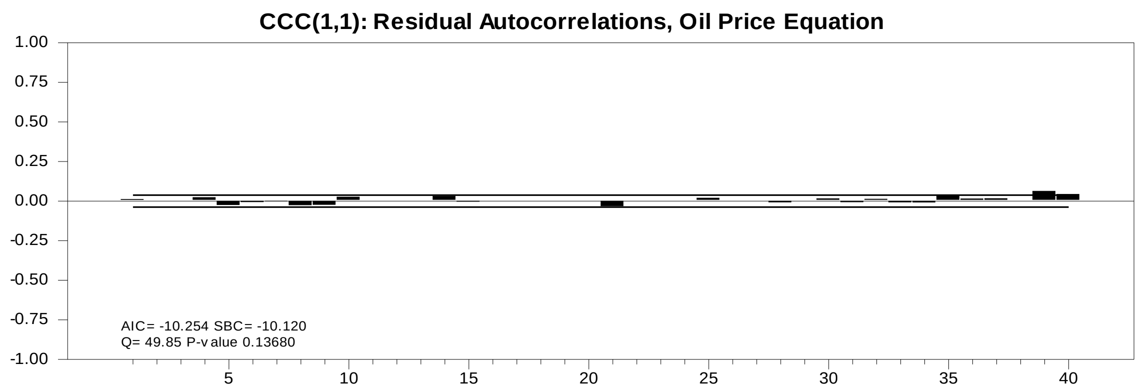
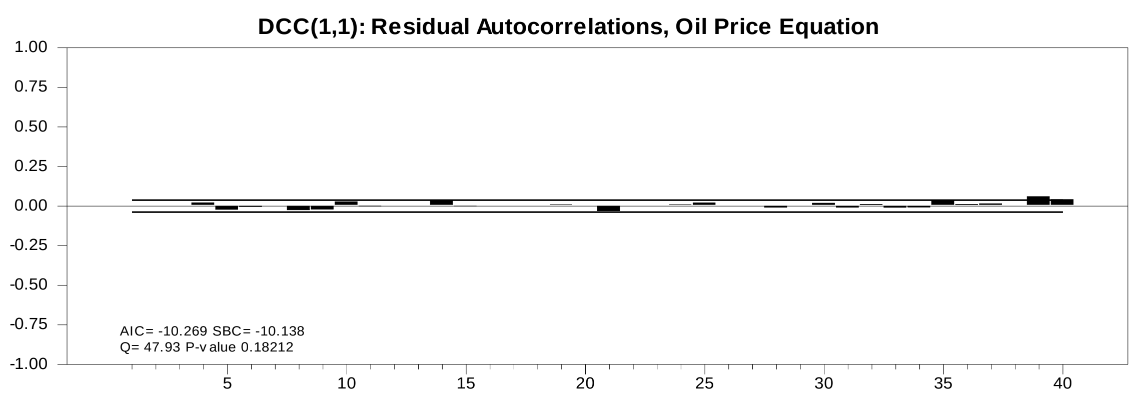


Figure 5.2: Residual Autocorrelation Plots for the VARMA(1,1) Asymmetric GARCH-in-Mean Models

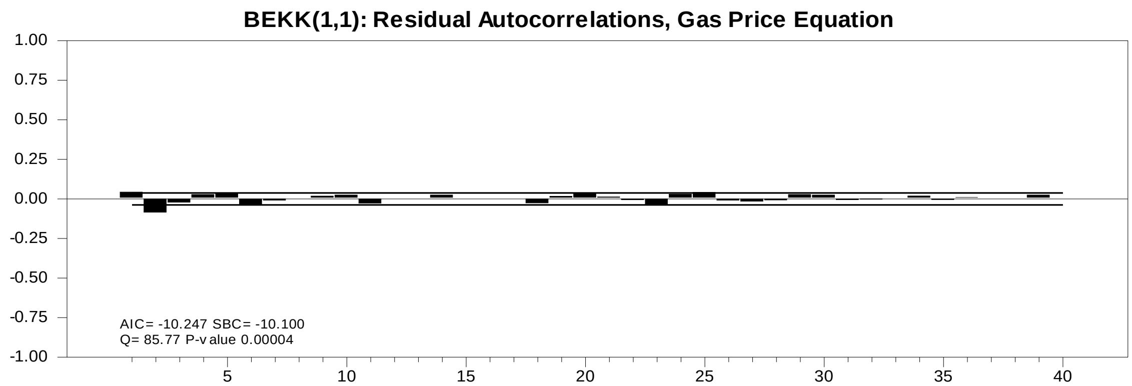
a) Oil Price Equation

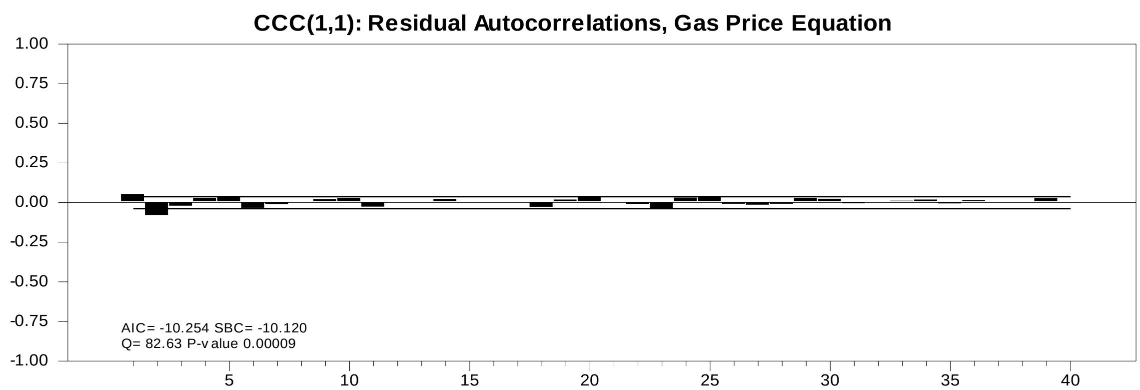


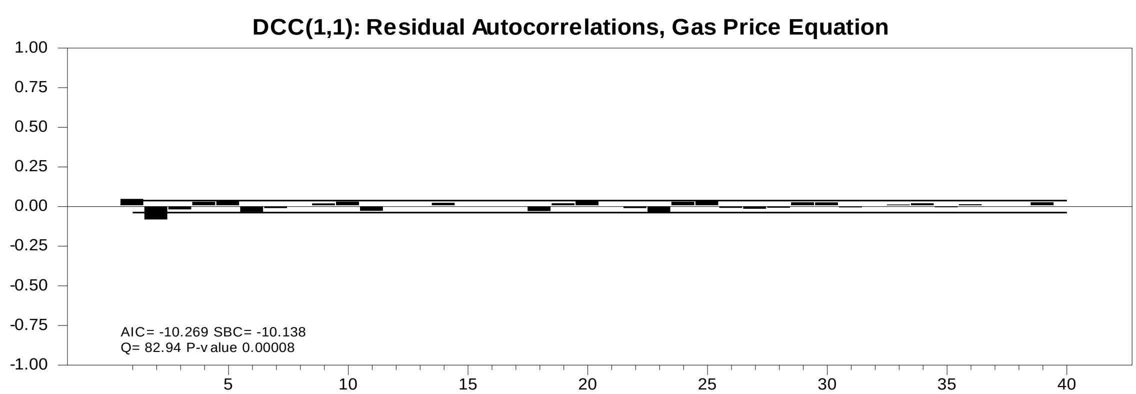




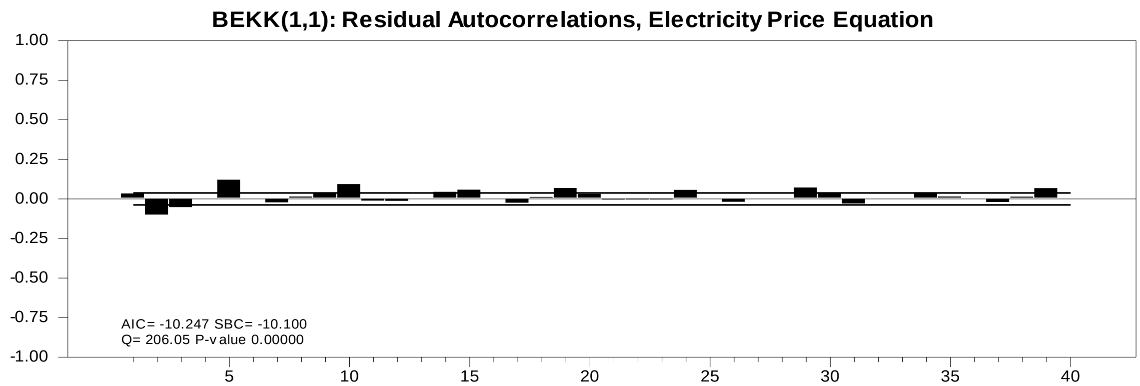
b) Gas Price Equation

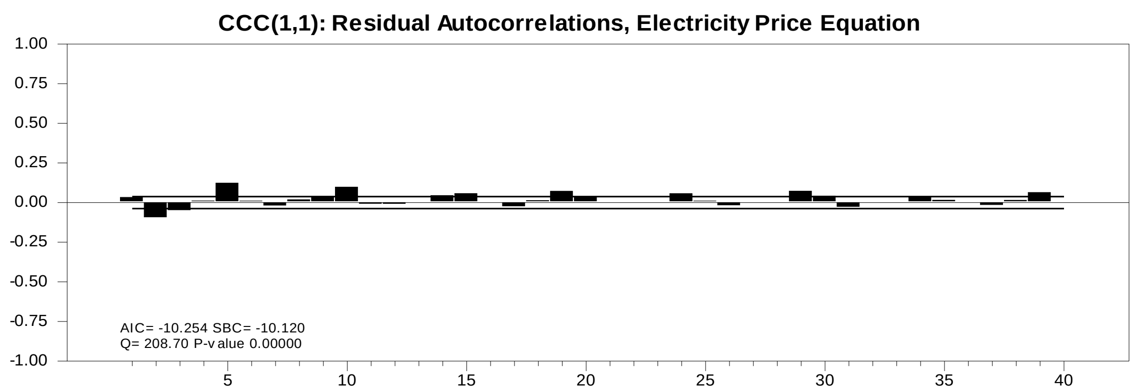






c) Electricity Price Equation





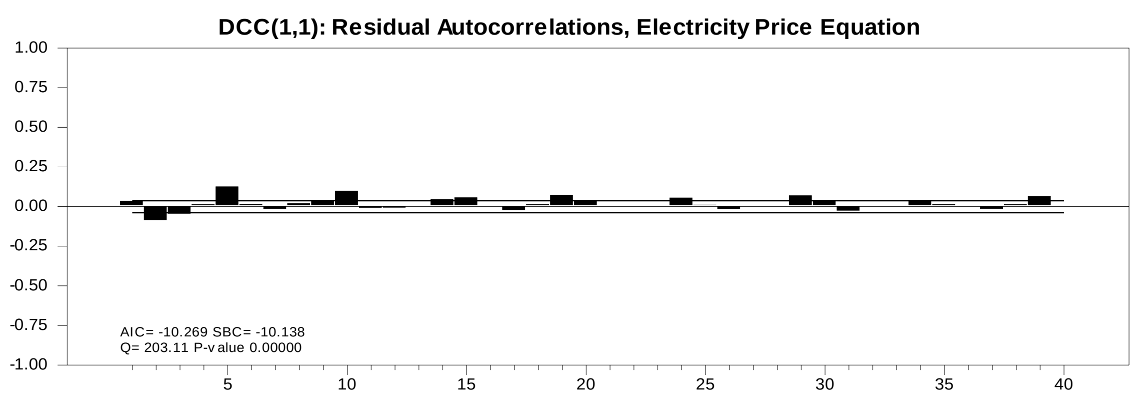
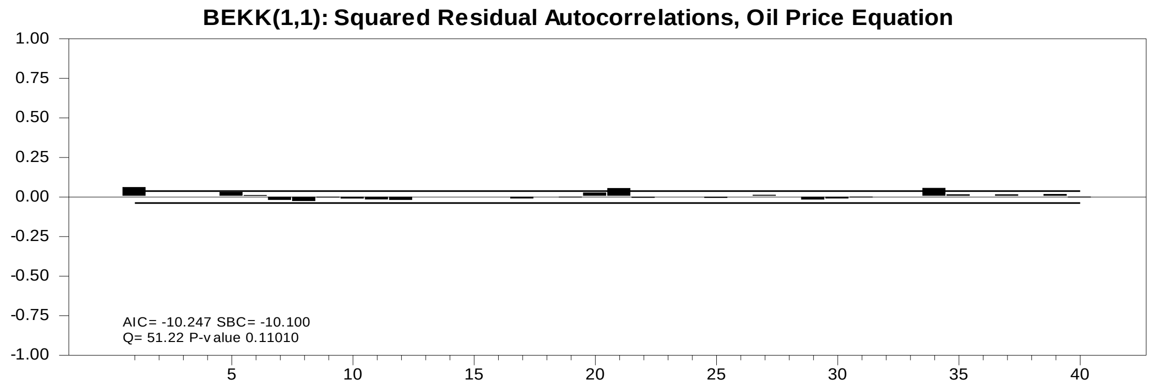
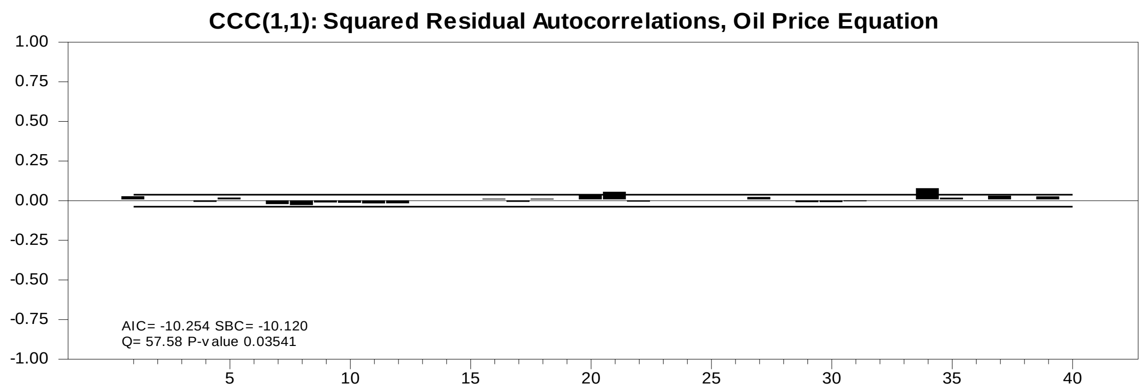
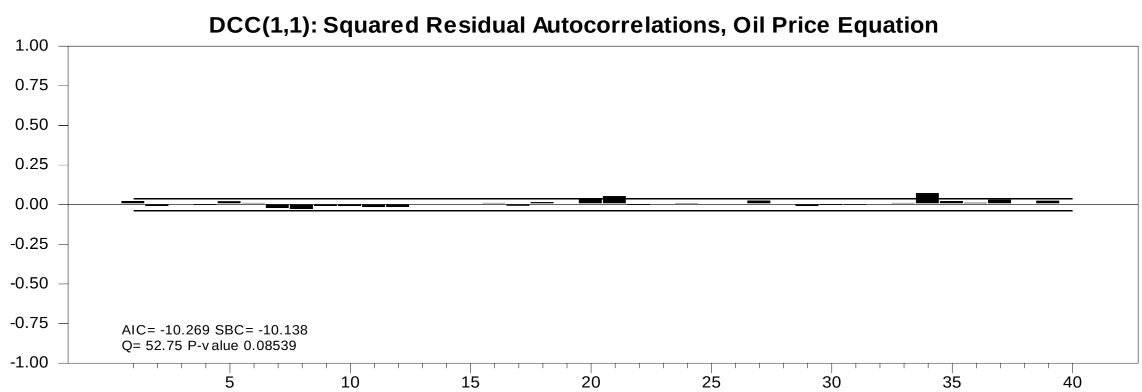


Figure 5.3: Squared Residual Autocorrelation Plots for the VARMA(1,1) Asymmetric GARCH- in-Mean Models

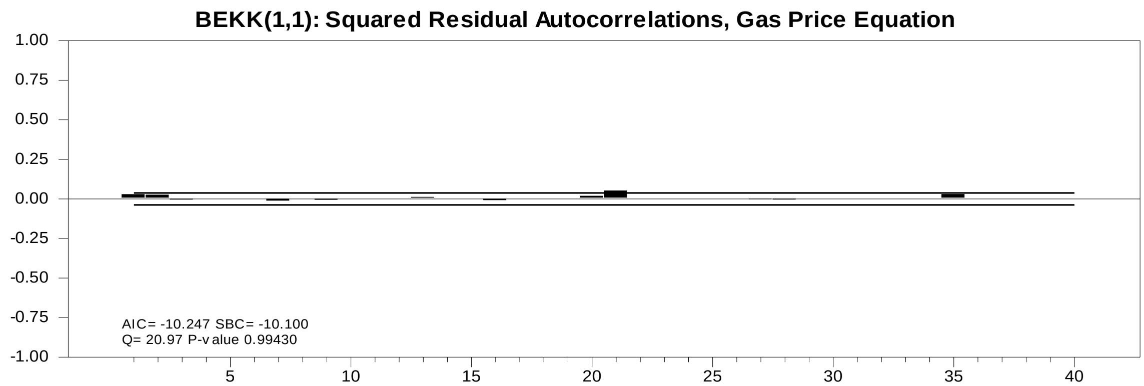
a) Oil Price Equation

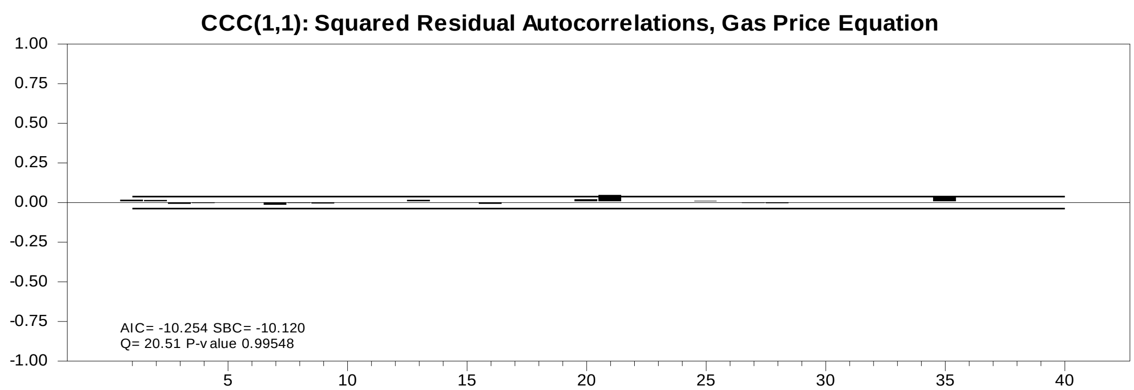


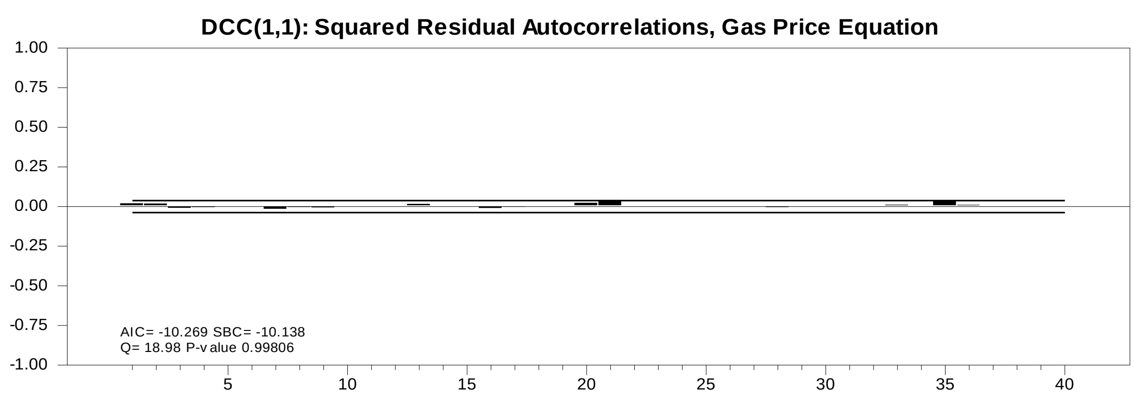




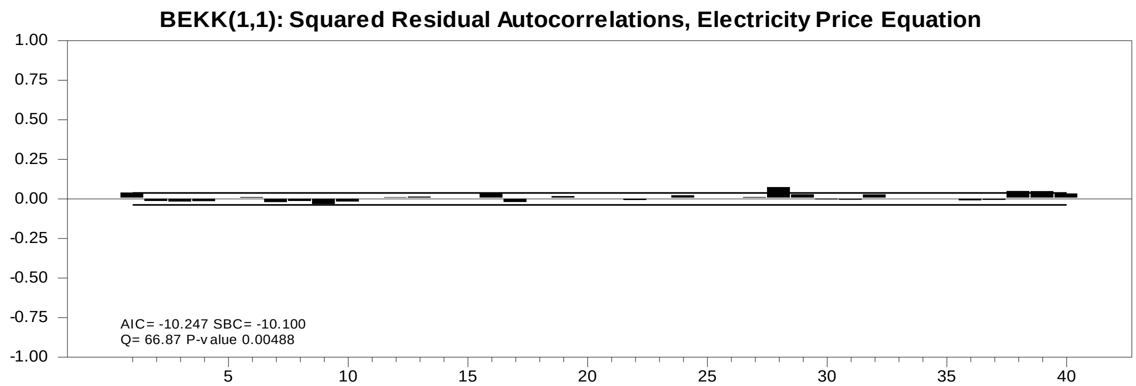
b) Gas Price Equation

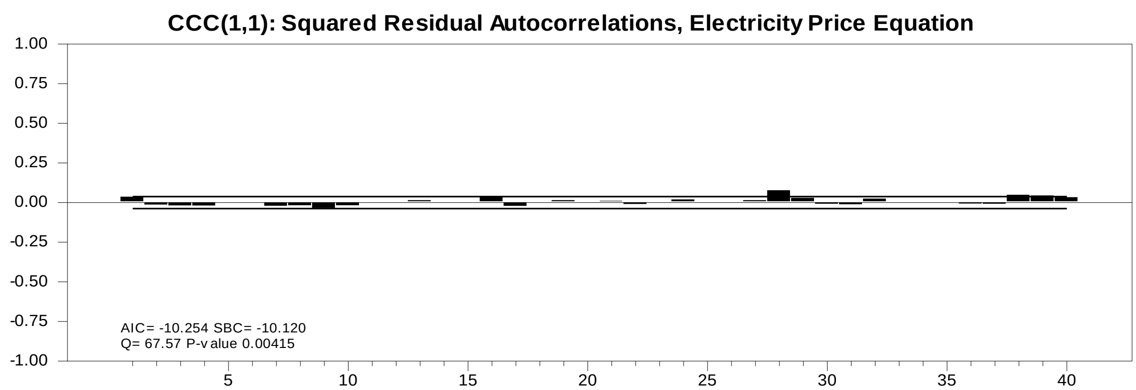






c) Electricity Price Equation





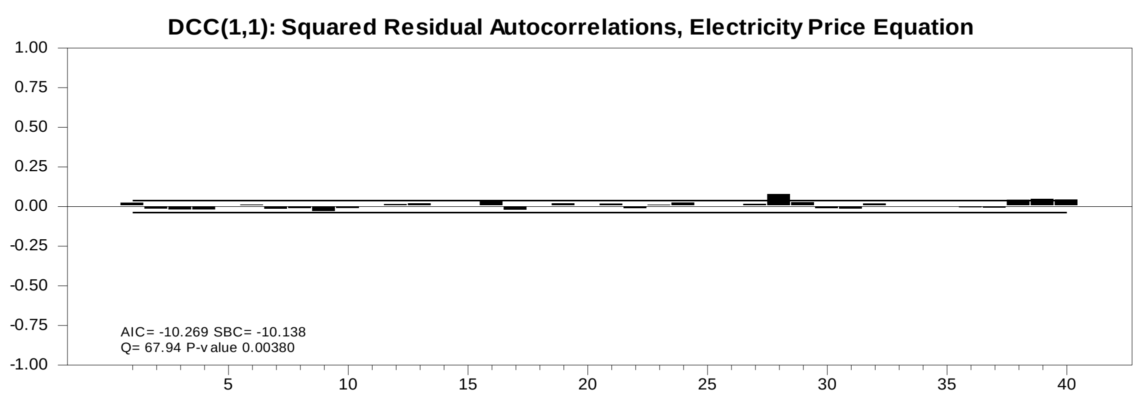
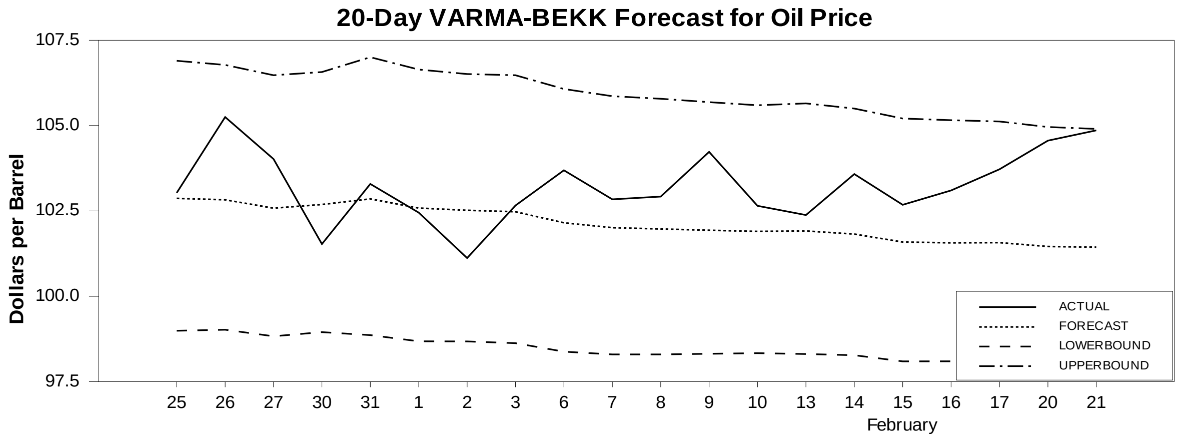
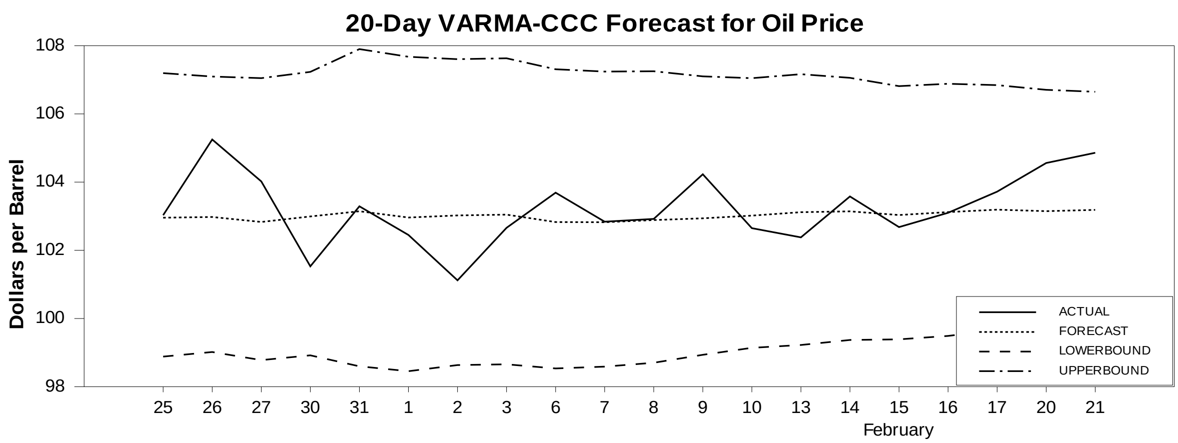


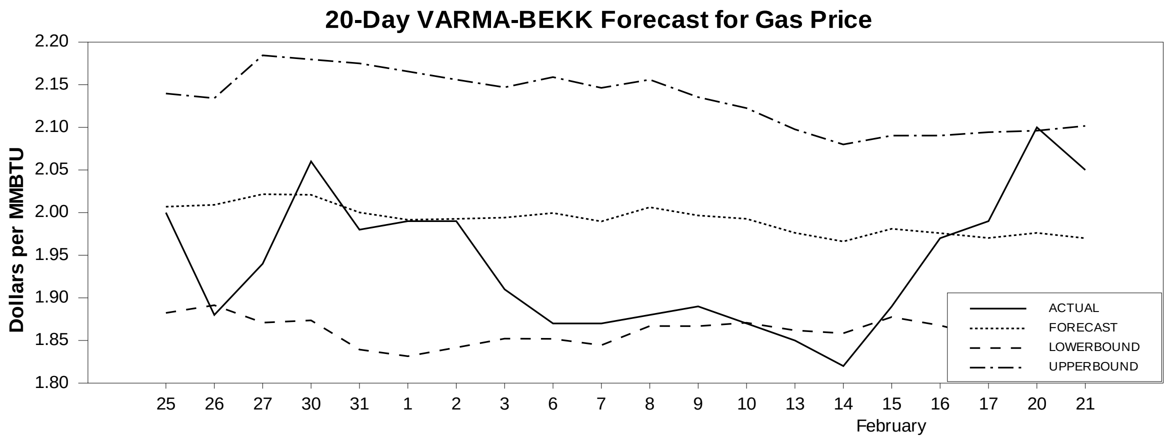
Figure 5.4: 20-Day Forecasts by VARMA(1,1) Asymmetric GARCH-in-Mean Models

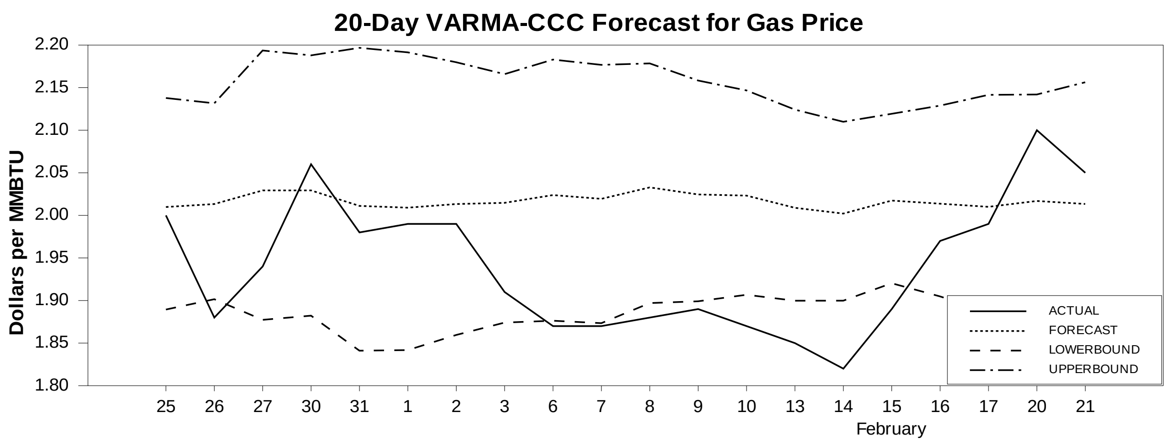
a) Oil Price Volatility Forecasts



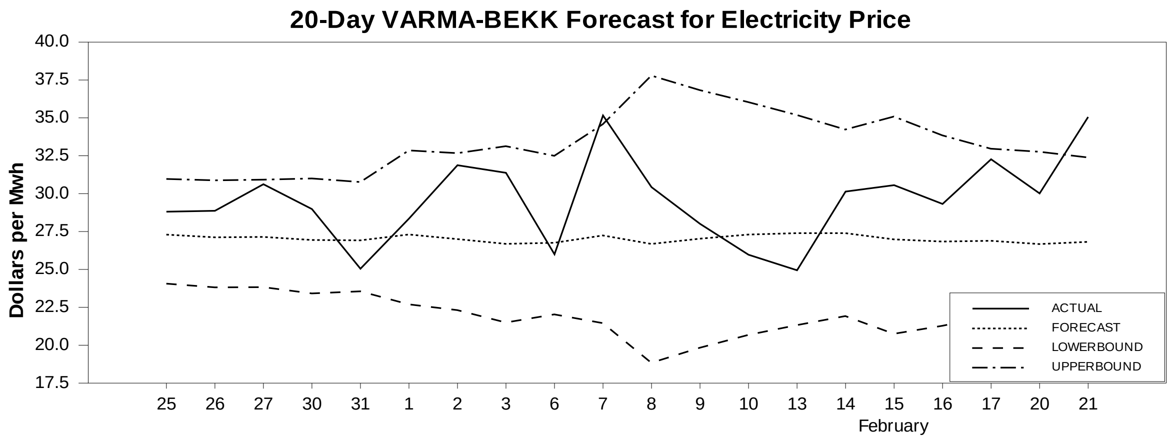


b) Gas Price Volatility Forecasts





c) Electricity Price Volatility Forecasts



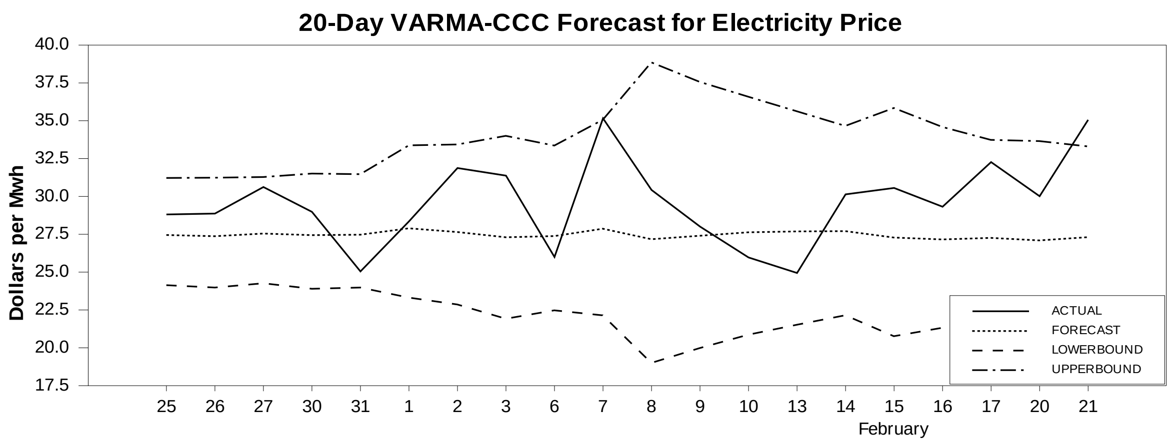
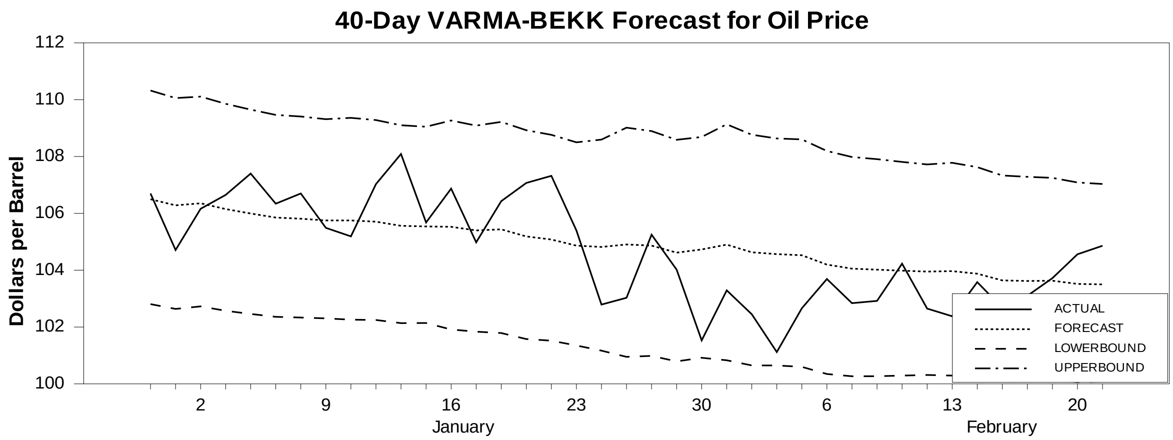
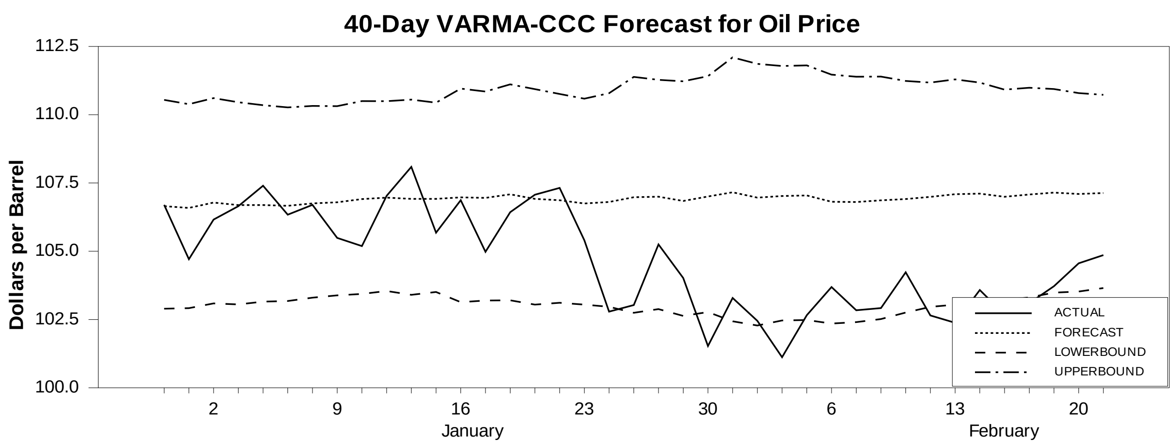


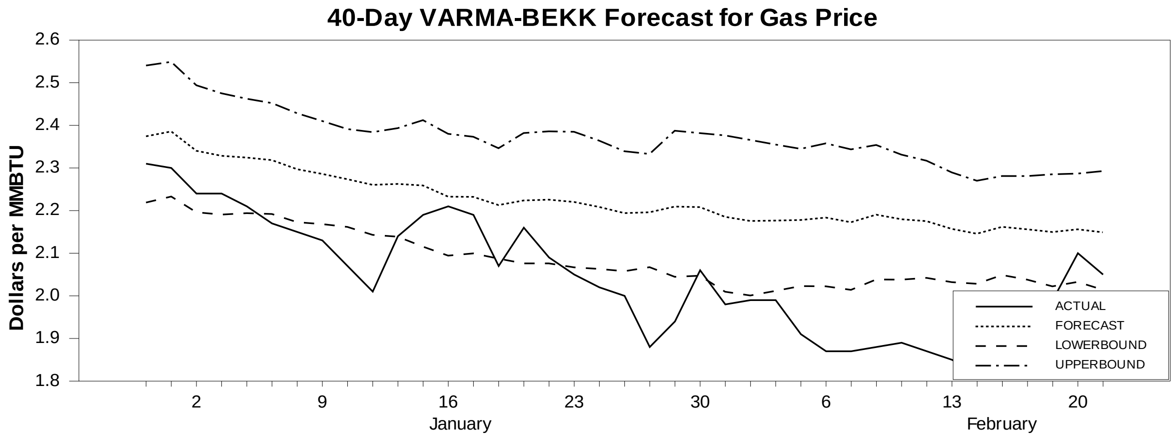
Figure 5.5: 40-Day Forecasts by VARMA(1,1) Asymmetric GARCH-in-Mean Models

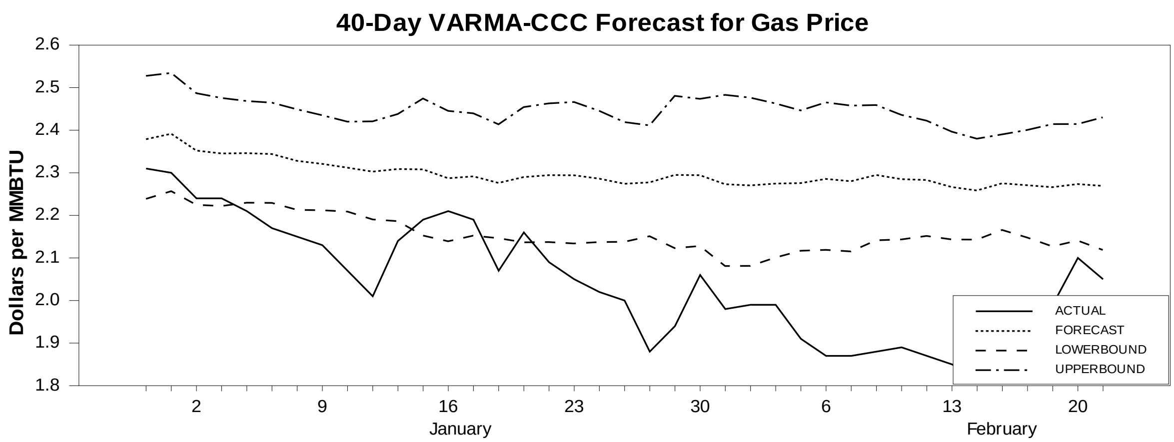
a) Oil Price Volatility Forecasts



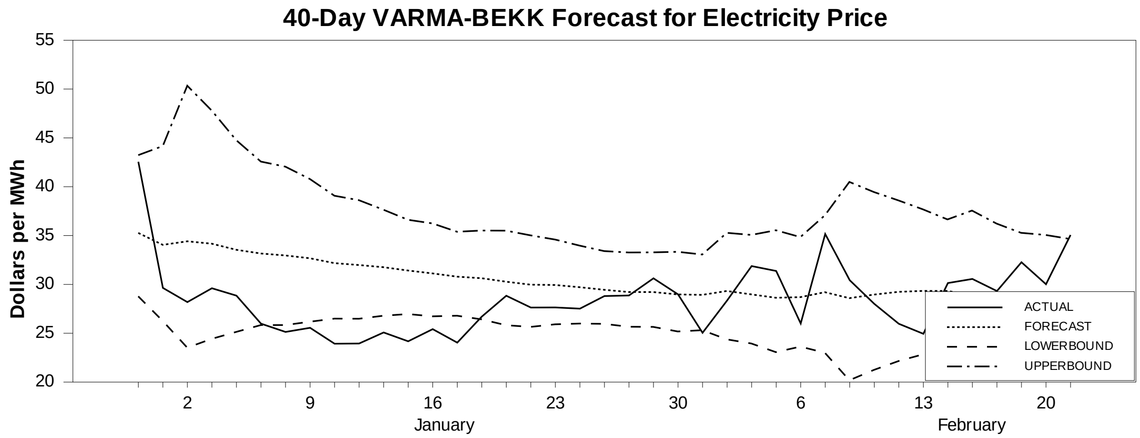


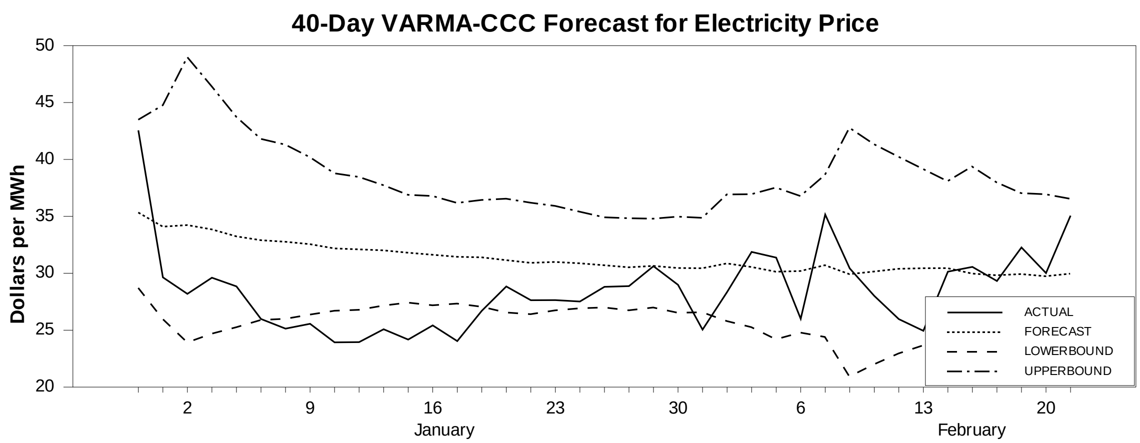
b) Gas Price Volatility Forecasts





c) Electricity Price Volatility Forecasts





6 Conclusion

Globalization, growing energy demand, increasing intensity of extreme weather events and geopolitical tensions, as well as the deregulation of electricity markets, all mean that price volatility will remain a central feature of oil, natural gas and electricity markets for decades to come. Originally developed in finance, GARCH models have become indispensable in short-term volatility modelling of energy commodity prices, largely because they are very efficient at accommodating irregular periods of price volatility and tranquility that are characteristic of energy markets. This thesis presents a comprehensive empirical study that applies a range of univariate and multivariate GARCH models to daily oil, natural gas and electricity price data from U.S. wholesale markets, for the period from 2001 to 2012.

We find that univariate models produce superior mean equation forecasts, as evidenced by impressive Theil's U statistics of 0.6-0.7 for all three commodities, compared to multivariate models' Theil's U of 0.95-1.00. The latter indicates forecasting performance that is not much better than that of a random walk. For variance forecasts, informal graphic analysis suggests that both univariate and multivariate models yield good estimates, correctly predicting periods of price volatility and tranquility within 95% confidence bounds. An excellent direction for future research would be to develop objective measures for the performance of volatility forecasts, such as those that exist for point forecasts.

Given our large dataset of over 2,800 daily observations for each price series, univariate GARCH models also allow a wide selection of additional variables to be included in both the mean and variance equations. The models presented in Chapters 2, 3 and 4 use dummy variables for days of the week and seasons, as well as daily data on weather, the Dow Jones index, and natural gas storage inventories. Many other variables can potentially be included to reflect relevant political events, extreme weather events like hurricanes, facility breakdowns causing supply shocks, and oil storage inventories. Also, with wind energy generation growing at a rapid rate, future studies on electricity volatility could benefit from incorporating wind speed data for key producing regions. With multivariate GARCH models, on the other hand, our choice of additional regressors is extremely limited, since each one adds an entire vector of parameters.

However, one advantage of using multivariate GARCH in our case is the opportunity to forego first-differencing in the vector autoregressive mean equation due to the presence of cointegration among the three series, leading to information preservation and more robust estimates. The very fact that oil, natural gas and electricity prices cointegrate shows a considerable degree of interdependence among the three respective markets. Another suggestion for future research is to explore in more depth the nature and reasons behind this cointegration.

The greatest strength of multivariate GARCH is the potential to investigate the interactions among all three commodity prices and their volatilities, which makes it possible for us to discover surprising and significant spillover effects. We find that both price and volatility spillovers are rather unidirectional, suggesting the existence of a hierarchy of influence from oil to gas and electricity markets. Specifically, electricity price and its volatility have little impact on either of the other two markets; natural gas price and volatility significantly affect electricity price and volatility, respectively. This is expected since natural gas is an input in electricity generation, but not vice versa. However, oil price and volatility generate large and significant spillovers into both natural gas and electricity markets, including price-to-volatility and volatility-to-price spillovers. The degree of influence of the oil market is quite impressive. In fact, these findings strongly underline the importance of oil in the U.S. economy today, and the far-reaching implications of events in wholesale oil markets.

Recent advances in volatility modelling have resulted in an explosion of new types of GARCH models including asymmetric, multivariate, nonparametric, structural break and multiplicative models. This thesis shows that model type and specification must be carefully considered to reflect features of the data, the size of the dataset, the availability of useful data for additional regressors, and, most importantly, the research question. Many of the new multivariate models are rich and flexible but computationally demanding, making their estimation as much of an art as science. By applying several types of GARCH models to oil, natural gas and energy price data, we are able to not only compare their perfomance, but also draw out common trends and effects that persist across formulations. Thus we contribute to the understanding of price volatility in wholesale energy markets, and suggest several effective models that would be of use to energy market participants, derivatives market participants, large energy consumers interested in hedging strategies, and policymakers.

Bibliography

Andersen, T. G., and Bollerslev, T. (1998). Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts. *International Economic Review* 39(4), 885-905.

Baillie R.T., Bollerslev, T., and Mikkelsen, H.O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3-30.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. J*ournal of Econometrics* 31, 307-327.

Bühlmann, P. and McNeil, A.J. (2002). An algorithm for nonparametric GARCH modelling. *Journal of Computational Statistics & Data Analysis* 40, 665-683.

Caporale, G.M., Pittis, N., Sakellis, P. (2003). Testing for PPP: the erratic behaviour of unit root tests. *Economics Letters* 80(2), 277-284.

Chan, K.F., Gray, P. (2006). Using extreme value theory to measure value at risk for daily electricity spot prices. *International Journal of Forecasting* 22 (2), 283–300.

Chang, C.L., McAleer, M., and Tansuchat, R. (2010). Analyzing and forecasting volatility spillovers, asymmetries and hedging in major oil markets. *Energy Economics* 32, 1445-1455.

Davidson, R., and MacKinnon, J. (1993). *Estimation and Inference in Econometrics.* London: Oxford University Press.

Dickey, D.A., Fuller,W.A., (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74, 427–431.

Elder, J., and Serletis, A., (2010). Oil price uncertainty. *Journal of Money, Credit and Banking* 42(6), 1137-1159.

Electric Power Supply Association (2012). *What Is a Wholesale Electricity Market?* Retrieved from http://www.epsa.org/industry/primer/?fa=wholesaleMarket

Energy Information Administration (2012). *Annual Energy Outlook 2012 with projections to 2035.* Retrieved from http://www.eia.gov/forecasts/aeo/pdf/0383(2012).pdf

Energy Information Administration (2012). *Sources and Uses - Data.* Retrieved from http://www.eia.gov/

Energy Information Administration (2010). *Status of Electricity Restructuring by State.* Retrieved from http://www.eia.gov/cneaf/electricity/page/restructuring/restructure_elect.html

Engle, R.F. (2002). Dynamic conditional correlation. *Journal of Business & Economic Statistics* 20(3), 339-350.

Engle, R.F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50(4), 987-1007

Engle, F.F., and Granger, C.W. (1987). Co-integration and Error Correction: Representation, Estimation, and Testing. *Econometrica*, 55(2), 251-276.

Engle, R.F., and Kroner, K.F. (1995). Multivariate Simultaneous Generalized Arch. *Econometric Theory* 11(1) Vol. 11, 122-150.

Engle, R. & Sokalska, M.E. (2012). Forecasting intraday volatility in the US equity market. Multiplicative Component GARCH. *Journal of Financial Econometrics* 10(1), 54–83.

Estima (2012). *RATS Version 8 User's Guide.* Evanston, IL: Estima.

Ewing, B.T., Malikb, F., and Ozfidanc, O. (2002). Volatility transmission in the oil and natural gas markets. *Energy Economics* 24, 525–538.

Get Energy Active (2012). *Keep Our Fuel Mix Diverse.* Retrieved from http://www.getenergyactive.org/fuel/state.htm

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), 1779–1801.

Goto, M., and Karolyi, G. A. (2004). Understanding Electricity Price Volatility within and across Markets. Ohio State University, Charles A. Dice Center for Research in Financial Economics. Working Paper Series 2004-12.

Grier, K.B., Henry, O.T., Olekalns, N., and Shields, K. (2004). The asymmetric effects of uncertainty on inflation and output growth. *Journal of Applied Econometrics* 19(5), 551-565.

Hadsell, L., Marathe, A., Shawky, H.A. (2004). Estimating the volatility of wholesale electricity spot prices in the US. *Energy Journal* 25 (4), 23–40.

Hansen, P.R., and Lunde, A. (2001). A Comparison of Volatility Models: Does Anything

Beat a GARCH(1,1)? University of Aarhus Centre for Analytical Finance. Working Paper Series 84.

Hickey, E., Loomis, D.G., and Mohammadi, H. (2012). Forecasting hourly electricity prices using ARMAX–GARCH models: An application to MISO hubs. *Energy Economics*. doi:http://dx.doi.org/10.1016/j.eneco.2011.11.011

Higgs, H. (2009). Modelling price and volatility inter-relationships in the Australian wholesale spot electricity markets. *Energy Economics* 31, 748–756.

Higgs, H., Worthington, A. (2008). Stochastic price modelling of high volatility, mean-reverting, spike-prone commodities: the Australian wholesale electricity market. *Energy Economics* 30 (6), 3172–3195.

Hou, A., and Suardi, S. (2011). A nonparametric GARCH model of crude oil price return volatility. *Energy Economics* 34, 618-626.

Johansen, S. (1988). Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2-3), 231-254.

Johnson, T. (2011). *Oil Market Volatility*. Council on Foreign Relations. Retrieved from http://www.cfr.org/energy/oil-market-volatility/p15017

Kwiatowski, Phillips, Schmidt & Shin(1992). Testing the null hypothesis of stationarity against the alternative of a unit root: how sure are we that economic time series have a unit root? *Journal of Econometrics* 54, 159-178.

Lee, K., Ni, S., and Ratti, R. A. (1995). Oil Shocks and the Macroeconomy: The Role of Price Variability. *The Energy Journal* 16(4), 39-56.

Lin, S.X., and Tamvakis, M.V. (2001). Spillover effects in energy futures markets. *Energy Economics* 23, 43-56.

Lindström, E., and Regland, F. (2012). Modeling extreme dependence between European electricity markets. *Energy Economics* 34(4), 899-904.

Lütkepohl, H. (2004). Vector Autoregressive and Vector Error Correction Models. In Lütkepohl, H., and Krätzig, M. (Eds.), *Applied Time Series Econometrics* (pp. 86-158). Cambridge, UK: Cambridge University Press.

Morana, C. (2001). A semiparametric approach to short-term oil price forecasting. *Energy Economics* 23, 325-338.

Namit, S. (1998). *Forecasting Oil Price Volatility.* Faculty of the Virginia Polytechnic Institute and State University

National Climatic Data Centre. (2012). *Online Climate Data Directory.* Retrieved from http://www.ncdc.noaa.gov/oa/climate/climatedata.html

Mu, X. (2007). Weather, storage, and natural gas price dynamics: Fundamentals and volatility. *Energy Economics* 29, 46-63.

Rahman, S., and Serletis, A. (2012). Oil price uncertainty and the Canadian economy: evidence from a VARMA, GARCH-in-Mean, asymmetric BEKK model. *Energy Economics*

34(2), 603-610.

Sadorsky, P. (2011). Correlations and volatility spillovers between oil prices and the stock prices of clean energy and technology companies. *Energy Economics* 34, 248-255.

Serletis, A. (2013). *Oil Price Uncertainty.* Hackensack, NJ: World Scientific Publishing Co.

Serletis, A., and Shahmoradi, A. (2006). Measuring and testing natural gas and electricity markets volatility: evidence from Alberta's deregulated markets. *Studies in Nonlinear Dynamics & Econometrics* 10(3), article 10.

Wang, Y., and Wu, C. (2012) Forecasting energy market volatility using GARCH models: Can multivariate models beat univariate models? *Energy Economics* 34(2), doi:10.1016/j.eneco. 2012.03.010.

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, 337–350.