IJEEP

INTERNATIONAL JOURNAL OF ENERGY ECONOMICS AND POLICY

EJ EconJournals

ISSN: 2146-4553

International Journal of Energy Economics and Policy

ISSN: 2146-4553

available at http: www.econjournals.com

International Journal of Energy Economics and Policy, 2016, 6(4), 814-821.

EJ EconJournals

Is the Best Generalized Autoregressive Conditional Heteroskedasticity(p,q) Value-at-risk Estimate also the Best in Reality? An Evidence from Australian Interconnected Power Markets

Rangga Handika[1,2]*, Sigit Triandaru[3]

[1]College of Business Administration, Abu Dhabi University, United Arab Emirates, [2]Faculty of Business and Economics Universitas, Indonesia, [3]Universitas Atmajaya Yogyakarta, Indonesia. *Email: rangga.handika@adu.ac.ae/rhandikapro@yahoo.com

ABSTRACT

This paper investigates whether the best value-at-risk (VaR) estimate will also perform the best in empirical performance. The study explores the linkage between statistical world and reality. This paper uses VaR generalized autoregressive conditional heteroskedasticity (GARCH)(p,q) estimates and performs the back testing from both generator (buyer) and retailer (seller) sides, at different confidence levels, and at different out-of-sample periods in the four regions of Australian interconnected power markets. Using VaR approach, we find that the best GARCH(p,q) model tends to generate best empirical performance. Our findings are consistent for both generator (buyer) and retailer (seller) sides, at different confidence levels and at different out-of-sample periods. However, our strong results are only in the daily series. Therefore, our study has two important practical implications in Australian power markets. First, generator and retailer can continue choosing the best GARCH(p,q) model based on statistical criteria. Second, the users of GARCH(p,q) model should be aware that the model tends to be appropriate for estimating the daily series only.

Keywords*:* Power Markets, Generalized Autoregressive Conditional Heteroskedasticity, Value-at-risk

JEL Classifications: G17, G32, Q40, Q47

1. INTRODUCTION

There are limited works investigating value-at-risk (VaR) in power market. One reason is that VaR normally applied in financial markets instead of power market. However, recent studies (Stoll and Whaley, 2010; Tang and Xiong, 2012; Basak and Pavlova, 2016) indicate that there is a sharp increase of investments in commodity markets. This is called financialization of commodity markets. As spot power market is considered a commodity (Joskow and Kahn, 2001), the VaR method should also be applied in the power market.

There are a number of recent studies on modelling volatility in commodity markets (Chkili et al., 2014; Efimova and Serletis, 2014; Youssef et al., 2015; Cabrera and Schulz, 2016). However, none of them investigate whether the best estimated model, based on a statistical criteria, is also the best in empirical performance (i.e., reality). Indeed, this is an important question exploring the relationship between statistical estimate and reality performance.

Using high frequency data in Australian interconnected power markets, we use generalized autoregressive conditional heteroskedasticity (GARCH) (Bollerslev, 1986) model to explore the linkage between statistical estimate and empirical performance. According to Higgs and Worthington (2008), Australian power markets are significantly more volatile than other comparable power markets. They use daily series, obtained from averaging different forty-eight half-hourly prices, to compromise difficulties in other intraday information. Therefore, our daily series in Australian interconnected power markets is appropriate modelled by GARCH (like other studies in power markets such as: Koopman et al., 2007; Efimova and Serletis, 2014).

We also realize that most papers investigating VaR in power market tend to take the position from generators', the sellers', perspective (like Chan and Gray, 2006; Walls and Zhang, 2006; Frauendorfer and Vinarski, 2007; Andriosopoulos and Nomikos, 2012). Indeed, analyzing VaR in power market from retailers', the buyers', perspective is also essential since the price spike generates market risks that must be managed properly. Recent paper by Handika and Triandaru (2016) addresses this importance since a dramatic increase of price change (i.e., huge positive "return") is unfavourable while a dramatic decrease of price change (i.e., huge negative "return") is favourable for the retailers in power market.

This paper contributes to the literature in twofold. First, it responds to the question whether the best VaR estimate will perform the best in empirical performance. Second, it covers the analysis from both buyers' and sellers' sides. The remainder of the paper is organized as follow. Section 2 discusses a brief overview about Australia interconnected power markets. Section 3 reviews relevant studies about VaR in power markets. Section 4 discusses the method. Section 5 describes the data and discusses the empirical results. Section 6 concludes.

2. AUSTRALIAN INTERCONNECTED POWER MARKETS

According to Higgs and Worthington (2008), Australian interconnected power markets are considered more volatile and spike-prone than many comparable power markets. Australian power market is on one of the world's longest interconnected power systems comprising several regional networks supplying electricity to retailers and end-users.

The Australian power market has transformed over the last two decades. Before 1997, the market participants of the power market were owned by government and monopolies. Then, there was a significant structural reform in the late 1990s. The reform creates market competition in the retail power market. Nowadays, the Australian power market is an interconnected power market among several regional networks between power suppliers and retailers. The market is managed by the Australian Energy Market Operator (AEMO) which operates under the National Electricity Law and National Electricity Rules. The price in the power market is formed by following steps: (i) The generators submit offers every five minutes, (ii) the submitted offers then become the basis in determining the number of generators that are required to produce electricity, (iii) the final price is constructed every half-hour for each of region by averaging the five minutes spot prices.

3. LITERATURE REVIEW

Both academics and practitioners have investigated volatility modelling in commodity markets in the recent decade. For instance, Chan and Gray (2006) develop and apply extreme value theory (EVT) to model the tails of the return distribution in the power markets. Another paper by Walls and Zhang (2006) also uses EVT in their extended VaR model and reports that the extended VaR is more accurate in the Alberta power market. Frauendorfer and Vinarski (2007) conduct sensitivity analysis of the VaR with respect to the risk factors price and volatility in the power markets. Andriosopoulos and Nomikos (2012) extend a number of VaR models to capture the dynamics of energy prices. A recent work by Chkili et al. (2014) models the conditional volatility of four widely traded commodities. Efimova and Serletis (2014) use various univariate and multivariate GARCH models to investigate the empirical properties of price volatilities in three different commodities. Youssef et al. (2015) forecast volatility in oil and natural gas markets by using long-memory-models. Another paper by Cabrera and Schulz (2016) examines volatilities using an asymmetric dynamic conditional correlation GARCH model and a multivariate multiplicative volatility model in various commodity markets. However, none of them has addressed whether the best VaR estimate will also generate the best empirical performance in power markets. Furthermore, the papers discussing VaR analysis are limited from buyers' side. Therefore, this paper explores the linkage between VaR estimate and reality in both buyers' and sellers' sides.

4. METHODS

We use GARCH model for estimating VaR in the interconnected power markets. According to Bollerslev (1986), GARCH(p,q) model can be written as follow:

$$\sigma_t^2=\omega+\sum_{i=1}^{q}\alpha_i y_{t-1}^2+\sum_{j=1}^{q}\beta_j\sigma_{t-1}^2 \tag{1}$$

Where, σ_t denotes the volatility forecast for time t and y_t is the realized return at time t.

We compare various p and q values up to GARCH(2,2), like Chinn and Coibion (2014).

We do the reality performance analysis by conducting the back-testing VaR (like Chkili et al., 2014). Our back-testing VaR procedure is intended to find out whether the best VaR estimate also generates the best performance. The best performance is the VaR estimate with minimum number of VaR violation. For robustness tests, we perform the back-testing VaR at various confidence levels (90%, 95% and 99%) for all regions at different data frequencies (daily, weekly, and monthly series).

We perform the procedures for both buyer (generator) and seller (retailer) sides. Indeed, an increase price change (i.e., a positive price change) is favourable while a decrease price change (i.e., a negative price change) is unfavourable for the generators in power market. However, a decrease price change (i.e., a negative price change) is favourable while an increase price (i.e., a positive price change) is favourable for the retailers in power market. Therefore, we cover the VaR back-testing analysis from both left-tailed and right-tailed sides.

5. EMPIRICAL ANALYSIS

We obtained the daily series of Australian power market prices in four interconnected markets (NSW, QLD, SA and VIC) from the AEMO website[1]. Then, we calculated the daily price for each region by averaging the different forty-eight half-hour power prices. After that, we obtained daily price change (i.e., "daily return") by calculating the natural log of the today price divided by the previous day price.

We use in-sample period is from January 1, 2000 to December 31, 2009 and the out-of-sample period is from January 1, 2010 to December 31, 2015 (10 years of in-sample period and 6 years out of sample period). Our choice about in-sample and out-of-sample periods is motivated by the studies indicating that financialized of commodity markets has started since 2000s (Rossi, 2012; Tang and Xiong, 2012). Therefore, we start our in-sample period at 1-Jan-2000. Ledoit and Wolf (2008) document that many finance empirical works use 10 years. Therefore, we end the in-sample period at 31-Dec-2009.

Table 1 reports the descriptive statistics of daily price change in the NSW, QLD, SA and VIC regions. The descriptive statistics include the mean, standard deviation, minimum, maximum and the number of observation of each region. Panel A reports the descriptive statistics from January 1, 2000 to December 31, 2015 (all periods), Panel B reports the descriptive statistics from January 1, 2000 to December 31, 2009 (in-sample period) and Panel C reports the descriptive statistics from January 1, 2010 to December 31, 2015 (out-of- sample period). Note that we found few missing values in our observations, specifically 2 daily values in QLD region, 20 daily values in SA region and 4 daily values in VIC region. We replaced those missing values by previous values.

We find that average daily power price changes range from negative 0.11% in SA to 0.09% in VIC during all sample period, range from negative 0.01% in QLD to 0.13% in SA during in-sample period and range from negative 0.53% in SA to 0.05% in NSW and QLD during out-of-sample period. Overall, SA region tends to report the smallest price change while NSW, QLD and VIC regions tend to report the highest price change.

We also find that volatility of power price change range from 34.69% in VIC to 47.43% in SA during all sample period, range from 38.01% in VIC to 48.15% in SA during in-sample period and range 24.42% in NSW to 46.22% in SA during out-of-sample period. Overall, SA region tends to report the most volatile of price change while NSW and VIC regions tend to report the least volatile of price change.

In the next part, we perform the calibration of fitting the series into various p and q values of the GARCH(p,q) model up to GARCH(2,2). For robustness checks, we perform the calibration (and later the VaR modelling and backtesting) for different data frequencies including daily, weekly and monthly. We perform analysis at different data frequencies because there is an issue of data availability dictating the time span for commodity markets (Narayan et al., 2013).

1 The data can be downloaded here http://www.aemo.com.au/Electricity/Data/Price-and-Demand/Aggregated-Price-and-Demand-Data-Files.

Table 1: Descriptive statistics of daily price change in the four power market regions

Statistics descriptive	NSW (%)	QLD (%)	SA (%)	VIC (%)
Panel A: All sample period (from 1, Jan 2000 to 31, Dec 2015)				
Mean	0.03	0.01	−0.11	0.09
Standard deviation	38.66	45.10	47.43	34.69
Minimum	−403.28	−415.47	−398.58	−399.29
Maximum	406.45	426.88	439.62	381.15
Number of observation	5844	5844	5844	5844
Panel B: In sample period (from 1, Jan 2000 to 31, Dec 2009)				
Mean	0.02	−0.01	0.13	0.11
Standard deviation	45.06	47.95	48.15	38.01
Minimum	−403.28	−414.45	−395.46	−368.41
Maximum	406.45	426.88	439.62	381.15
Number of observation	3653	3653	3653	3653
Panel C: Out of sample period (from 1, Jan 2010 to 31, Dec, 2015)				
Mean	0.05	0.05	−0.53	0.04
Standard deviation	24.51	39.91	46.22	28.33
Minimum	−248.50	−415.47	−398.58	−399.29
Maximum	239.78	415.77	352.82	361.77
Number of observation	2191	2191	2191	2191

Table 2 reports the log-likelihood values for various p and q values up to GARCH(2,2) models for different data frequencies (daily, weekly and monthly series) in the four regions. According to Danielsson (2011), one of statistical evaluation methods of GARCH(p,q) model calibration is evaluating based on the log-likelihood value. The best model is the model with highest log-likelihood value (i.e. the least negative log-likelihood value). For daily series, the best model(s) is (are) GARCH(1,2) for all regions and GARCH(2,2) for NSW and SA regions. For weekly series, the best model(s) is(are) GARCH(2,2) for NSW, SA and VIC regions and GARCH(2,1) for NSW region, GARCH(1,1) for QLD region and GARCH(1,2) for VIC region. For monthly series, the best model(s) is (are) GARCH(1,1) for NSW, SA and VIC regions and GARCH(2,1), GARCH(1,2), GACRH(2,2) for QLD region. Overall, we conclude that GARCH(2,2) tends to be the best model based on statistical criteria.

Then, we attempt to answer question whether the best model based on statistical criteria will generate the best performance. We also do another robustness test by performing yearly sub-sample analysis during out-of-sample period. This yearly robustness checks follow methods from Gorton and Rouwenhorst (2006), Wong (2010), and Perignon et al. (2008).

Table 3 reports the results of back-testing VaR model for p and q values of various GARCH(p,q) models up to GARCH(2,2) for daily, weekly and monthly data series, at 99%, 95% and 90% confidence levels and for one year, two years, three years, four years, five years and six years out-of-sample periods from generator (buyer) perspective. The reported numbers states the number of VaR violations. A VaR violation from generators (buyer) perspective occurs when a negative return is worse than left-tailed VaR limit in the designated confidence level. The best VaR estimate implies the least VaR violation. The least VaR violation

Table 2: The log-likelihood values of various GARCH (p, q) models up to GARCH (2,2) for different data frequencies in the four regions

Series	GARCH (p, q)	Log-likelihood			
		Region			
		NSW	QLD	SA	VIC
Daily	GARCH (1,1)	−656.29	−2043.21	−2715.48	−1001.69
	GARCH (2,1)	−656.62	−2038.21	−2715.73	−1002.02
	GARCH (1,2)	−653.88	−2042.84	−2694.16	−994.64
	GARCH (2,2)	−653.88	−2036.08	−2694.16	−994.19
Weekly	GARCH (1,1)	−183.45	−512.57	−497.26	−277.08
	GARCH (2,1)	−182.69	−513.05	−496.94	−276.87
	GARCH (1,2)	−183.26	−512.96	−496.71	−274.37
	GARCH (2,2)	−182.69	−512.96	−496.61	−274.37
Monthly	GARCH (1,1)	−176.81	−161.22	−148.25	−137.28
	GARCH (2,1)	−177.06	−160.27	−148.55	−138.07
	GARCH (1,2)	−177.06	−160.27	−148.55	−138.81
	GARCH (2,2)	−177.06	−160.27	−148.55	−138.05

The best model is the highest log-likelihood value. VaR: Value-at-risk, GARCH: Generalized autoregressive conditional heteroskedasticity

was typed bold to clearly show the best empirical performance of the respective GARCH(p,q) model.

For daily series, overall we find that GARCH(1,2) tends to perform best. However, we should note the results are contradict between NSW and QLD regions and SA and VIC regions. GARCH(1,2) and GARCH(2,2) perform best for all out-of-sample periods for 95% confidence level in the NSW region. For 90% confidence level, however, GARCH(1,1) and GARCH(2,1) perform best for long years (5 and 6 years) out-of-sample period while GARCH(1,2) and GARCH(2,2) perform best for short years (1-4 years) out-of-sample periods in the NSW region. For 99% confidence level, none of GARCH is superior to others in the NSW region. GARCH(2,1) and GARCH(2,2) perform best for almost all of out-of-sample period for 95% confidence level while GARCH(2,1) perform best for 90% confidence level in the QLD region. For 99% confidence level, GARCH(1,1) and GARCH(1,2) perform best for long years (4-6 years) out-of-sample periods and none of GARCH is superior to others for shot years (1-3 years) out-of-sample periods in the QLD region. GARCH(1,2) and GARCH(2,2) consistently perform best at 99%, 95% and 90% confidence levels in both SA and VIC regions.

For weekly series, overall we find that GARCH(1,1) tends to perform slightly better compare to other GARCH(p,q). The overall results from weekly series indicate that the best statistical GARCH(p,q) model does not tend to generate best empirical performance. GARCH(1,1) and GARCH(1,2) perform best at 95% confidence level in the NSW region for almost all out-of-sample periods (except for 1 year) in the NSW region. However, most results in the NSW region report that none of GARCH is superior to others for 99% and 90% confidence levels at all out-of-sample periods. Furthermore, almost all results in the QLD also show that none of GARCH is superior to others at all confidence levels. GARCH(1,1), GARCH(2,1) and GARCH(1,2) perform best only in three, four and five years out-of-sample period at 90% confidence level in the SA region. Almost all of the remaining results show that none of GARCH is superior to others at all confidence levels. GARCH(1,1) and GARCH(2,1) perform best in all of out-of-sample period at 95% and 90% confidence levels in the VIC region. For 99% confidence level, none of GARCH is superior to others in the VIC region, except for 6 years out-of-sample period that reports GARCH(1,1) and GARCH(2,1) are the best.

For monthly series, overall we find that none of GARCH is superior to others. The results are strong for all out-of-sample periods and all 99%, 95% and 90% confidence levels in the NSW and QLD regions. Most similar results are also shown in the SA and VIC regions. The exception is that GARCH(2,1), GARCH(1,2) and GARCH(2,2) perform best for all out-of-sample periods at 90% confidence level in the SA region. Another exception is that GARCH(2,1) and GARCH(2,2) perform best for only 6 years out-of-sample period at 95% confidence level in the VIC region.

Table 4 reports the results of back-testing VaR model for p and q values of various GARCH(p,q) models up to GARCH(2,2) for daily, weekly and monthly data series, at 99%, 95% and 90% confidence levels and for one year, two years, three years, four years, five years and six years out-of-sample periods from retailer (seller) perspective. The reported numbers states the number of VaR violations. From retailer perspective, however, we should note that a VaR violation occurs when a positive return is higher than right-tailed VaR limit in the designated confidence level. We type bold the least VaR violation to indicate the best empirical performance of the respective GARCH(p,q) model.

For daily series, overall we find that GARCH(1,2) tends to perform best. However, we should note the results are also contradict between NSW and QLD regions and SA and VIC regions. GARCH(1,1) and GARCH(2,1) perform best for all of out-of-sample periods at 90% confidence level in the NSW region. For 99% confidence level, however, GARCH(1,1) and GARCH(2,1) perform best only for six years of out-of-sample period while GARCH(1,2) and GARCH(2,2) perform best at most out-of-sample periods (specifically from 1 year to 4 years) in the NSW region. For 95% confidence level, none of GARCH is superior to others at most out-of-sample periods (specifically from one year to four years) in the NSW region. GARCH(1,1) and GARCH(1,2) perform best for most all of out-of-sample

Table 3: Result of back-testing VaR model for P and q values of various GARCH (p, q) models up to GARCH (2,2) for different data frequencies, at different confidence levels and for different out-of-sample periods in the four regions from generator (buyer) perspective

Buyer (generator) side																				
Daily series																				
Number of VAR violations																				
Out-of-sample period			**6 years**			**5 years**			**4 years**			**3 years**			**2 years**			**1 year**		
Confidence level			**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**
Region	NSW	GARCH (1,1)	2	11	34	1	9	26	1	5	19	1	5	16	1	4	13	0	2	6
		GARCH (2,1)	2	11	34	1	9	26	1	5	19	1	5	16	1	4	13	0	2	6
		GARCH (1,2)	2	10	39	1	8	27	1	4	20	1	4	17	1	3	11	0	1	4
		GARCH (2,2)	2	10	39	1	8	27	1	4	20	1	4	17	1	3	11	0	1	4
	QLD	GARCH (1,1)	8	22	41	6	17	34	4	13	25	4	11	20	3	6	14	3	5	10
		GARCH (2,1)	8	20	43	7	15	32	5	11	24	4	9	19	3	5	13	3	4	9
		GARCH (1,2)	8	22	41	6	17	34	4	13	25	4	11	20	3	6	14	3	5	10
		GARCH (2,2)	9	19	44	8	15	34	6	11	26	4	9	20	3	5	14	3	4	9
	SA	GARCH (1,1)	17	58	117	13	45	85	11	36	61	10	29	48	8	21	34	3	11	18
		GARCH (2,1)	17	58	117	13	45	85	11	36	61	10	29	48	8	21	34	3	11	18
		GARCH (1,2)	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
		GARCH (2,2)	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
	VIC	GARCH (1,1)	7	38	82	4	26	52	3	19	36	3	15	29	3	11	22	1	4	9
		GARCH (2,1)	7	38	82	4	26	52	3	19	36	3	15	29	3	11	22	1	4	9
		GARCH (1,2)	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
		GARCH (2,2)	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1	0	0	1
Weekly series																				
Number of VAR violations																				
Out-of-sample period			**6 years**			**5 years**			**4 years**			**3 years**			**2 years**			**1 year**		
Confidence level			**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**
Region	NSW	GARCH (1,1)	0	6	12	0	4	8	0	1	5	0	1	4	0	0	3	0	0	1
		GARCH (2,1)	0	8	13	0	5	8	0	2	5	0	2	4	0	1	3	0	0	1
		GARCH (1,2)	0	6	12	0	4	8	0	1	5	0	1	4	0	0	3	0	0	1
		GARCH (2,2)	0	8	13	0	5	8	0	2	5	0	2	4	0	1	3	0	0	1
	QLD	GARCH (1,1)	2	4	7	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1
		GARCH (2,1)	2	4	6	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1
		GARCH (1,2)	2	4	6	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1
		GARCH (2,2)	2	4	6	2	2	3	1	1	1	1	1	1	1	1	1	1	1	1
	SA	GARCH (1,1)	5	9	20	4	6	15	3	4	12	3	3	10	3	3	9	1	1	5
		GARCH (2,1)	5	9	20	4	6	15	3	4	12	3	3	10	3	3	9	1	1	5
		GARCH (1,2)	5	9	19	4	6	15	3	4	12	3	3	10	3	3	9	1	1	5
		GARCH (2,2)	5	9	20	4	6	16	3	4	13	3	3	11	3	3	10	1	1	5
	VIC	GARCH (1,1)	2	6	13	1	3	8	1	2	6	1	2	6	1	1	5	0	0	2
		GARCH (2,1)	2	6	13	1	3	8	1	2	6	1	2	6	1	1	5	0	0	2
		GARCH (1,2)	3	10	17	1	5	11	1	4	8	1	4	8	1	3	7	0	1	3
		GARCH (2,2)	3	10	17	1	5	11	1	4	8	1	4	8	1	3	7	0	1	3
Monthly series																				
Number of VAR violations																				
Out-of-sample period			**6 years**			**5 years**			**4 years**			**3 years**			**2 years**			**1 year**		
Confidence level			**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**	**99**	**95**	**90**
Region	NSW	GARCH (1,1)	0	2	2	0	1	1	0	1	1	0	1	1	0	1	1	0	0	0
		GARCH (2,1)	0	2	2	0	1	1	0	1	1	0	1	1	0	1	1	0	0	0
		GARCH (1,2)	0	2	2	0	1	1	0	1	1	0	1	1	0	1	1	0	0	0
		GARCH (2,2)	0	2	2	0	1	1	0	1	1	0	1	1	0	1	1	0	0	0
	QLD	GARCH (1,1)	0	0	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
		GARCH (2,1)	0	0	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
		GARCH (1,2)	0	0	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
		GARCH (2,2)	0	0	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
	SA	GARCH (1,1)	2	4	6	2	2	4	1	1	2	1	1	2	1	1	2	0	0	1
		GARCH (2,1)	2	4	5	2	2	3	1	1	1	1	1	1	1	1	1	0	0	0
		GARCH (1,2)	2	4	5	2	2	3	1	1	1	1	1	1	1	1	1	0	0	0
		GARCH (2,2)	2	4	5	2	2	3	1	1	1	1	1	1	1	1	1	0	0	0
	VIC	GARCH (1,1)	0	1	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
		GARCH (2,1)	0	0	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
		GARCH (1,2)	0	1	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0
		GARCH (2,2)	0	0	2	0	0	1	0	0	1	0	0	1	0	0	1	0	0	0

VaR: Value-at-risk, GARCH: Generalized autoregressive conditional heteroskedasticity

Table 4: Result of back-testing VaR model for P and q values of various GARCH (p, q) models up to GARCH (2,2) for different data frequencies, at different confidence levels and for different out-of-sample periods in the four regions from retailer (seller) perspective

Seller (retailer) side																				
Daily series																				
Number of VAR violations																				
Out-of-sample period			6 years			5 years			4 years			3 years			2 years			1 year		
Confidence level			99	95	90	99	95	90	99	95	90	99	95	90	99	95	90	99	95	90
Region	NSW	GARCH (1,1)	22	43	58	17	32	43	16	27	33	15	26	29	13	21	23	9	14	14
		GARCH (2,1)	22	43	58	17	32	43	16	27	33	15	26	29	13	21	23	9	14	14
		GARCH (1,2)	24	45	70	17	33	50	15	27	37	14	26	33	12	21	25	8	14	16
		GARCH (2,2)	24	45	70	17	33	50	15	27	37	14	26	33	12	21	25	8	14	16
	QLD	GARCH (1,1)	42	65	81	31	48	62	22	37	48	13	20	28	11	13	16	4	5	8
		GARCH (2,1)	42	66	84	31	50	64	22	38	51	13	20	30	11	13	19	4	5	9
		GARCH (1,2)	42	65	81	31	48	62	22	37	48	13	20	28	11	13	16	4	5	8
		GARCH (2,2)	42	67	83	31	50	63	22	38	50	13	20	29	11	13	18	4	5	8
	SA	GARCH (1,1)	49	80	132	39	59	95	34	47	75	20	28	50	16	21	35	8	10	18
		GARCH (2,1)	49	80	132	39	59	95	34	47	75	20	28	50	16	21	35	8	10	18
		GARCH (1,2)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
		GARCH (2,2)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
	VIC	GARCH (1,1)	16	47	86	16	33	58	15	27	43	12	21	35	8	13	23	6	11	14
		GARCH (2,1)	16	47	86	16	33	58	15	27	43	12	21	35	8	13	23	6	11	14
		GARCH (1,2)	1	2	2	1	2	2	1	2	2	1	2	2	1	2	2	1	2	2
		GARCH (2,2)	1	2	2	1	2	2	1	2	2	1	2	2	1	2	2	1	2	2
Weekly series																				
Number of VAR violations																				
Out-of-sample period			6 years			5 years			4 years			3 years			2 years			1 year		
Confidence level			99	95	90	99	95	90	99	95	90	99	95	90	99	95	90	99	95	90
Region	NSW	GARCH (1,1)	7	16	21	5	12	17	5	11	16	4	10	15	3	7	9	2	4	5
		GARCH (2,1)	7	15	21	5	11	17	5	10	16	4	9	15	3	7	9	2	4	5
		GARCH (1,2)	7	16	21	5	12	17	5	11	16	4	10	15	3	7	9	2	4	5
		GARCH (2,2)	7	15	21	5	11	17	5	10	16	4	9	15	3	7	9	2	4	5
	QLD	GARCH (1,1)	5	6	13	4	5	8	2	3	6	1	1	3	1	1	1	0	0	0
		GARCH (2,1)	5	6	13	4	5	8	2	3	6	1	1	3	1	1	1	0	0	0
		GARCH (1,2)	5	6	13	4	5	8	2	3	6	1	1	3	1	1	1	0	0	0
		GARCH (2,2)	5	6	13	4	5	8	2	3	6	1	1	3	1	1	1	0	0	0
	SA	GARCH (1,1)	6	13	19	5	7	12	2	3	7	1	2	6	1	2	6	1	2	3
		GARCH (2,1)	6	13	19	5	7	12	2	3	7	1	2	6	1	2	6	1	2	3
		GARCH (1,2)	6	13	20	5	7	13	2	3	8	1	2	7	1	2	7	1	2	4
		GARCH (2,2)	6	13	20	5	7	13	2	3	8	1	2	7	1	2	7	1	2	4
	VIC	GARCH (1,1)	1	15	20	1	10	14	1	8	12	1	7	11	1	5	9	1	3	5
		GARCH (2,1)	1	15	20	1	10	14	1	8	12	1	7	11	1	5	9	1	3	5
		GARCH (1,2)	5	18	28	4	12	19	4	10	17	3	9	15	2	7	12	1	4	6
		GARCH (2,2)	5	18	28	4	12	19	4	10	17	3	9	15	2	7	12	1	4	6
Monthly series																				
Number of VAR violations																				
Out-of-sample period			6 years			5 years			4 years			3 years			2 years			1 year		
Confidence level			99	95	90	99	95	90	99	95	90	99	95	90	99	95	90	99	95	90
Region	NSW	GARCH (1,1)	1	1	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
		GARCH (2,1)	1	1	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
		GARCH (1,2)	1	1	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
		GARCH (2,2)	1	1	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
	QLD	GARCH (1,1)	1	2	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
		GARCH (2,1)	1	2	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
		GARCH (1,2)	1	2	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
		GARCH (2,2)	1	2	3	1	1	2	1	1	2	1	1	2	1	1	1	0	0	0
	SA	GARCH (1,1)	2	5	8	1	4	7	1	2	4	1	2	4	1	2	3	1	1	1
		GARCH (2,1)	2	5	8	1	4	7	1	2	4	1	2	4	1	2	3	1	1	1
		GARCH (1,2)	2	5	8	1	4	7	1	2	4	1	2	4	1	2	3	1	1	1
		GARCH (2,2)	2	5	8	1	4	7	1	2	4	1	2	4	1	2	3	1	1	1
	VIC	GARCH (1,1)	1	2	5	1	2	3	1	2	2	1	2	2	1	1	1	0	0	0
		GARCH (2,1)	1	2	4	1	2	3	1	2	2	1	2	2	1	1	1	0	0	0
		GARCH (1,2)	1	2	5	1	2	3	1	2	2	1	2	2	1	1	1	0	0	0
		GARCH (2,2)	1	2	5	1	2	3	1	2	2	1	2	2	1	1	1	0	0	0

VaR: Value-at-risk, GARCH: Generalized autoregressive conditional heteroskedasticity period for 95% confidence level while GARCH(2,1) tends to perform best for 90% confidence level in the QLD region. For 99% confidence level, none of GARCH is superior to others in the QLD region. GARCH(1,2) and GARCH(2,2) also consistently perform best at 99%, 95% and 90% confidence levels in both SA and VIC regions.

For weekly series, overall we find that GARCH(1,1) tends to perform slightly better compare to other GARCH(p,q). The overall results from weekly series indicate that the best statistical GARCH(p,q) model does not tend to generate best empirical performance. None of GARCH is reported to be superior to others at all confidence levels in the QLD region. GARCH(2,1) and GARCH(2,2) only perform best at long term out-of-sample periods (from four years to six years) in the NSW region. GARCH(1,1) and GARCH(2,1) perform best only at 90% confidence level for all of out-of-sample periods in the SA region while GARCH(1,1) and GARCH(2,1) perform best at 99%, 95%, and 90% confidence levels for all of out-of-sample periods in the VIC region.

For monthly series, overall we also find that none of GARCH is superior to others. These consistency findings from both generator (buyer) and retailer (seller) sides can partially be explained from fewer data frequencies in the monthly series (at least compared to daily and weekly series). Therefore, it is not surprising that we obtain indifference results among GARCH models in the monthly series.

Overall, we find that the best GARCH(p,q) model (based on statistical criteria) tends to generate best empirical performance in the Australian interconnected power markets. Our findings are consistent from both generator (buyer) and retailer (seller) sides. However, our findings are strong only in the daily series. We argue that the strong results in daily series occurs because of the nature of GARCH model, which is designed to capture the properties of high frequency data. Several papers (Akgiray, 1989; Baillie and Bollerslev, 2002; Koopman et al., 2007; Lamoureux and Lastrapes, 1990) document that GARCH volatility estimate and forecast are well captured by daily series. Therefore, our findings also indirectly confirm that GARCH(p,q) is appropriate for daily series in power markets.

6. CONCLUSION

This paper investigates whether the best VaR estimate, as determined by statistical criteria, will also perform the best in empirical performance. We use various p and q values in VaR GARCH(p,q) estimation and perform the back testing from both generator (buyer) and retailer (seller) sides, at different confidence levels, and at different out-of-sample periods in the four regions of Australian interconnected power markets.

Using VaR approach, we find that the best GARCH(p,q) model (based on statistical criteria) tends to also generate best empirical performance in the Australian interconnected power markets. Our findings are consistent for both generator (buyer) and retailer (seller) side, at different confidence levels and at different out-of-sample periods. However, our strong results occur only for daily series. Our findings also confirm that GARCH (p,q) is appropriate for daily series in power markets.

Our findings lead to several implications for GARCH modelling in Australia power markets. First, we can continue choosing the best GARCH(p,q) model based on statistical criteria as we document that the best GARCH(p,q) also performs best in reality. Both generator and retailer can do this suggestion since the results are robust for both sides. Second, users of GARCH(p,q) model should be aware that the model is appropriate for estimating the daily series only.

REFERENCES

Akgiray, V. (1989), Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. The Journal of Business, 62(1), 55-80.

Andriosopoulos, K., Nomikos, N. (2012), Risk Management in the Energy Markets and Value-at-risk Modelling: A Hybrid Approach, EU Working Paper.

Baillie, R.T., Bollerslev, T. (2002), The message in daily exchange rates: A conditional-variance tale. Journal of Business and Economic Statistics, 20(1), 60-68.

Basak, S., Pavlova, A. (2016), A model of financialization of commodities. The Journal of Finance, 71(4), 1511-1556.

Bollerslev, T. (1986), Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics, 31(3), 307-327.

Cabrera, B.L., Schulz, F. (2016), Volatility linkages between energy and agricultural commodity prices. Energy Economics, 54, 190-203.

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, 283-300.

Chinn, M.D., Coibion, O. (2014), The predictive content of commodity futures. The Journal of Futures Markets, 34(7), 607-636.

Chkili, W., Hammoudeh, S., Nguyen, D.K. (2014), Volatility forecasting and risk management for commodity markets in the presence of asymmetry and long memory. Energy Economics, 41, 1-18.

Danielsson, J. (2011), Financial Risk Forecasting. West Sussex UK: John Wiley & Sons Ltd.

Efimova, O., Serletis, A. (2014), Energy markets volatility modelling using GARCH. Energy Economics, 43, 264-273.

Frauendorfer, K., Vinarski, A. (2007), Risk Measurement in Electricity Markets. Universität St. Gallen Working Paper.

Gorton, G., Rouwenhorst, K.G. (2006), Facts and fantasies about commodity futures. Financial Analysts Journal, 62(2), 47-68.

Handika, R., Triandaru, S. (2016), Retailer value-at-risk in interconnected power markets: An Australian empirical analysis. International Journal of Economics and Financial Issue, 6(6S), 6-9.

Higgs, H., Worthington, A. (2008), Stochastic prices modeling of high volatility, mean reverting, spike-prone commodities: The Australia wholesale spot electricity market. Energy Economics, 30, 3172-3185.

Joskow, P.L., Kahn, E. (2001), Commodity Index Investing and Commodity Futures Prices, MIT Center for Energy and Environmental Policy Research Working Paper.

Koopman, S.J., Ooms, M., Carnero, M.A. (2007), Periodic seasonal Reg-ARFIMA-GARCH models for daily electricity spot prices. Journal of the American Statistical Association, 102(477), 16-27.

Lamoureux, C.G., Lastrapes, W.D. (1990), Heteroskedasticity in stock return data: Volume versus GARCH effects. The Journal of Finance, 45(1), 221-229.

Ledoit, O., Wolf, M. (2008), Robust performance hypothesis testing with the sharpe ratio. Journal of Empirical Finance, 15, 850-859.

Narayan, P.K., Narayan, S., Sharma, S.S. (2013), An analysis of commodity markets: What gain for investors? Journal of Banking and Finance, 37(10), 3878-3889.

Perignon, C., Deng, Z.Y., Wang, Z.J. (2008), Do banks overstate their value-at-risk? Journal of Banking and Finance, 32(5), 783-794.

Rossi, B. (2012), The changing relationship between commodity prices and equity prices in commodity exporting countries. IMF Economic Review, 60(4), 533-569.

Stoll, H.A., Whaley, R.E. (2010), Commodity index investing and commodity futures prices. Journal of Applied Finance, 20(1), 7-46.

Tang, K., Xiong, W. (2012), Index investment and the financialization of commodities. Financial Analysts Journal, 68(6), 54-74.

Walls, W.D., Zhang, W. (2006), Using extreme value theory to model electricity price risk with an application to the Alberta power market. Energy, Exploration and Exploitation, 23(5), 375-404.

Wong, W.K. (2010), Backtesting value-at-risk based on tail losses. Journal of Empirical Finance, 17(3), 526-538.

Youssef, M., Belkacemb, L., Mokni, K. (2015), Value-at-risk estimation of energy commodities: A long-memory GARCH–EVT approach. Energy Economics, 51, 99-110.