Time Series Analysis Using ARCH Models: A Case Analysis of Australian Stock Index

Prabhat Mittal

Satyawati College Eve, University of Delhi, Delhi

Email Id: profmittal@yahoo.co.in

Abstract. Australian All Ordinaries Stock Index has been in the headline since 1997 for its tear jerking effect on the stock exchange. Present work attempts to develop a realistic time-series model to explain the behavior of the stock price data during 2 January 1997 to 29 December 2006 collected from www.yahoofinance.com. To begin with residual analysis reveals that assumption of constant one period ahead forecast variance does not hold true. Accordingly, a new class of stochastic processes, called Autoregressive Conditional Heteroscedastic (ARCH) is studied. To this end, Computer programs on Ms-Excel have been used to fit the ARCH model.

Introduction

Researchers engaged in forecasting time series, such as stock prices, inflation rates foreign exchange rates etc. have observed that their ability to forecast such variables varies considerably from one time period to another. This variability could very well be due to Volatility in financial markets, sensitive as they are to rumors, political upheavals, changes in Government monetary fiscal policies, and the like. This would suggest that the variance of forecast errors is not constant but varies from period to period, that is there is some kind of autocorrelations in the variance of forecast errors.

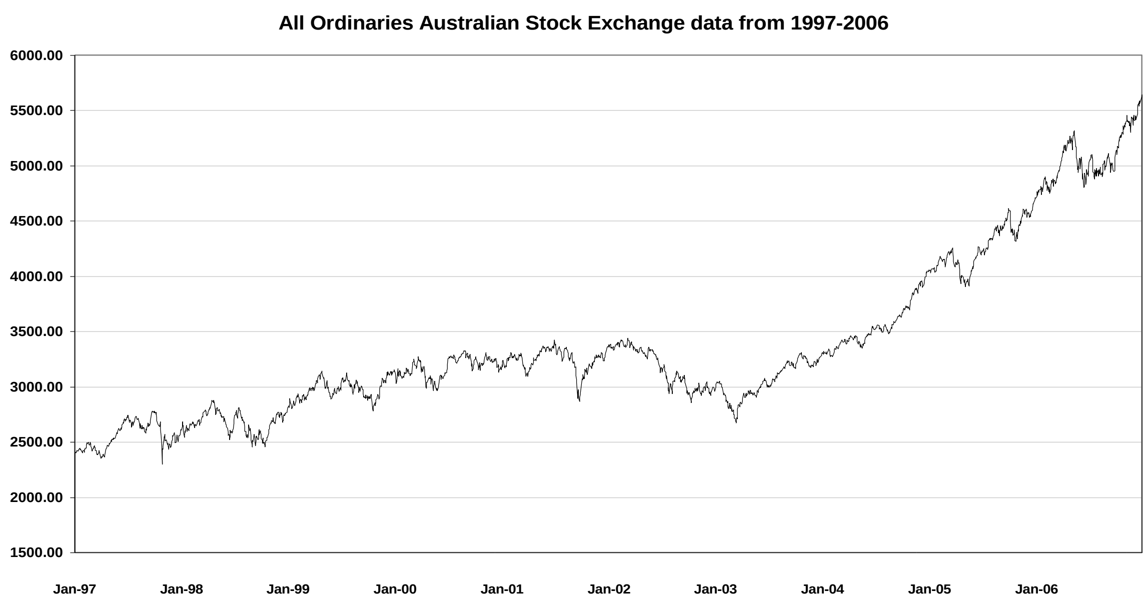
To capture this correlation Engle 1982 defined a *stochastic process* whose variables have conditional mean Zero and Conditional variance given by a linear function of previous squared variables known to be as **Autoregressive Conditional Heteroscedastic (ARCH) model.** The key idea of ARCH is that the variance of u at time t ($=\sigma^2_t$) depends on the size of the squared error term at time (t-1) i.e. on u^2_{t-1}.

There is a multitude of ARCH specifications, the best known being GARCH (generalized). The model is compatible with major stylized facts for asset returns and uses efficient methods for estimating model parameters and calculating forecasts for future volatility. ARCH models are defined by conditional density functions that provides the Likelihood function of data set, which can be maximized to give optimal parameter estimates. The GARCH(1,1) model with conditional normal distribution is the most popular ARCH specification in empirical research, particularly when modeling daily returns. The distribution of the return r_t for the period t, conditional in all previous return is defined by $r_t \mid r_{t-1}, r_{t-2\ldots} \sim N(\mu, h_t)$ with the conditional variances defined recursively by

$$h_t = \omega + \alpha e_{t-1}^2 + \beta h_{t-1} \text{ where } e_{t-1} = r_{t-1} - \mu .$$

& $$r_t = \log(p_t / p_{t-1}) = \mu + h_t^{1/2} z_t \ldots\ldots\ldots\ldots\ldots(A)$$

There are four parameters μ, α, β, ώ with ώ ≥ 0, α ≥ 0, β ≥ 0 to ensure that the conditional variance is never negative. The model is styled GARCH(1,1) because one previous squared residual and one previous value of the conditional variance are used to define the conditional variance for period t. To ensure stationary in GARCH process, α + β < 1. *The autocorrelation estimates can be used to test the hypothesis that the process generating standardized residuals is a series of independent & identically distributed (i.i.d.) random variables*. The i.i.d. can be tested using **Q Statistic**. The following graph shows the index level movement of the data.



GARCH (1,1) Model

The GARCH(1,1) model, with conditional normal distributions, is defined by equations A. Prices p_t, returns r_t, conditional variances h_t, and standardized residuals z_t along with some summary statistics like mean, standard deviation kurtosis, range for the set of 2532 returns are obtained using different functions in Ms-Excel. (Worksheet1). The parameters μ, α, β, ώ are estimated from the returns data by seeking the values giving the maximum likelihood $\log L = \sum_{t=1}^{2532} l_t$ with $l_t = -\frac{1}{2}[\log(2\pi) + \log(h_t) + z_t^2]$ where each term l_t is a function of μ, α, β, ώ.

Summary Stats					**Parameter**	**Value**		
	Returns r(t)				**Model**			
Mean	0.000336				**μ**	**0.000598**		
Standard Dev	0.007545				**ώ**	**0.000001**		
Skewness	-0.474941				**α**	**0.096070**		
Kurtosis	6.648217				**β**	**0.883569**		
Lag 1 correlation	-0.006920				**Reparameterized**			
					μ*1000	**0.597787**		
Minimum	-0.074487				**α**	**0.096070**		
Maximum	0.060666				**α+β**	**0.979638**		
					σsquare*10000	**0.621817**		
					Log L	**9001.8603**		
Date	Price	Return	Variance	Standardized Residual		Log density	Annualized percent Voltatility	Percentage return
	p(t)	r(t)	h(t)	z(t)		l(t)		
2-Jan-97	2411.20							
3-Jan-97	2399.50	-0.004864	0.000057	-0.723964		3.7059	12.00	-0.49
6-Jan-97	2409.80	0.004283	0.000054	0.499585		3.8656	11.73	0.43
7-Jan-97	2400.70	-0.003783	0.000051	-0.615550		3.8368	11.32	-0.38

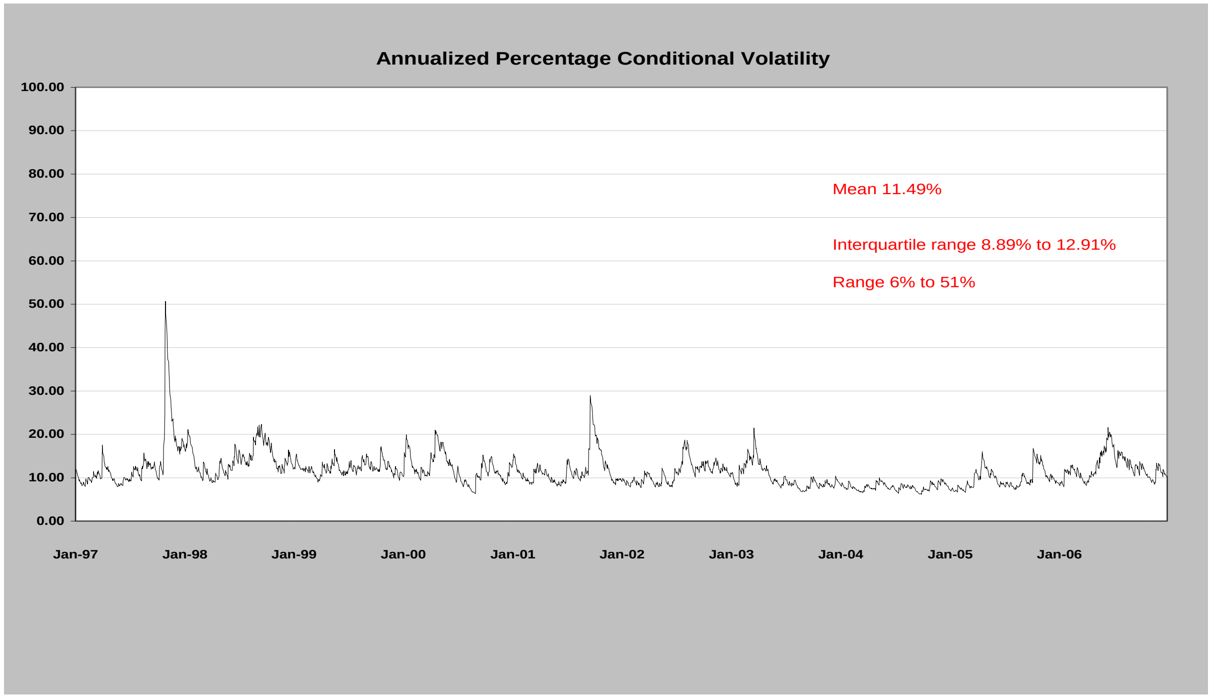


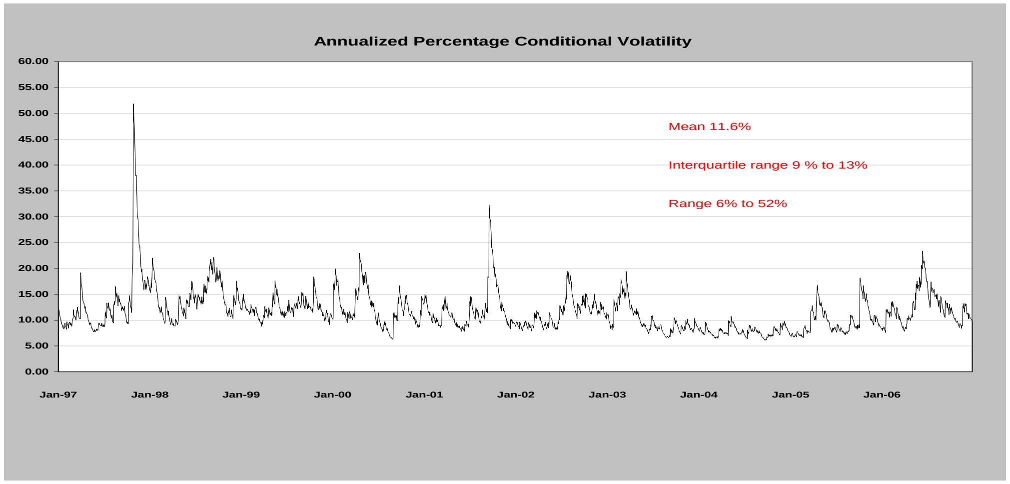
Figure shows the ten-year time series of volatility estimates from 1997 to 2006 given by the annualized conditional standard deviations, $\sigma_t = \sqrt{253h_t}$. The volatility estimates are potted as percentages and range from 6% to 51%. Half of the estimates are inside the inter quartile range, from 8.89% to 12.91%. The median and mean values are 10.97% & 11.49% respectively. The forecast function converges to 12.5% and half life is estimated to be 34 periods

GJR-GARCH (1,1) model states that μ_t is a constant and h_t is a linear function of e_{t-1}^2 and h_{t-1}. GJR-GARCH model introduces asymmetry in to GARCH model by weighting e_{t-1}^2 differently for negative and positive residuals; thus,

$$h_t = \omega + \alpha e_{t-1}^2 + \alpha^- S_{t-1} e_{t-1}^2 + \beta h_{t-1} \text{ with } S_{t-1} = \begin{cases} 1 & e_{t-1} < 0 \\ 0 & e_{t-1} \geq 0 \end{cases}$$

The conditional distribution is supposed to be normal and the parameters ($\mu, \lambda, \omega, \alpha, \alpha^{-1}, \beta$) are estimated by maximum the log-likelihood.

Summary Stats					Parameter	Value				
	Returns r(t)				*Model*					
Mean	0.000336				**μ**	0.000476				
Standard Deviation	0.007545				**λ**	0.003842				
Skewness	-0.474941				**θ**	-0.020001				
Kurtosis	6.648217				**ώ**	0.000001				
Lag 1 correlation	-0.006920				**α**	0.061707				
					α-	0.055261				
Minimum	-0.074487				**β**	0.889281				
Maximum	0.060666									
					Reparametrized					
					λ	0.003842				
					θ	-0.020001				
					α	0.061707				
					Asymmetry A	1.895530				
					Persistence φ	0.978618				
					Log L	9022.1437				
Date	Price	Return	mean	Variance	Residual	St resid	Sign	Log density	Annualized percent Voltatility	Percentage return
	p(t)	**r(t)**	**μ(t)**	h(t)	e(t)	z(t)	S(t)	L(t)		
2-Jan-97	2411.20									
3-Jan-97	2399.50	-0.004864	0.000336	0.000057	-0.005200	-0.6893	1	3.7305	12.00	-0.49
6-Jan-97	2409.80	0.004283	0.000504	0.000055	0.003779	0.5098	0	3.8556	11.93	0.43

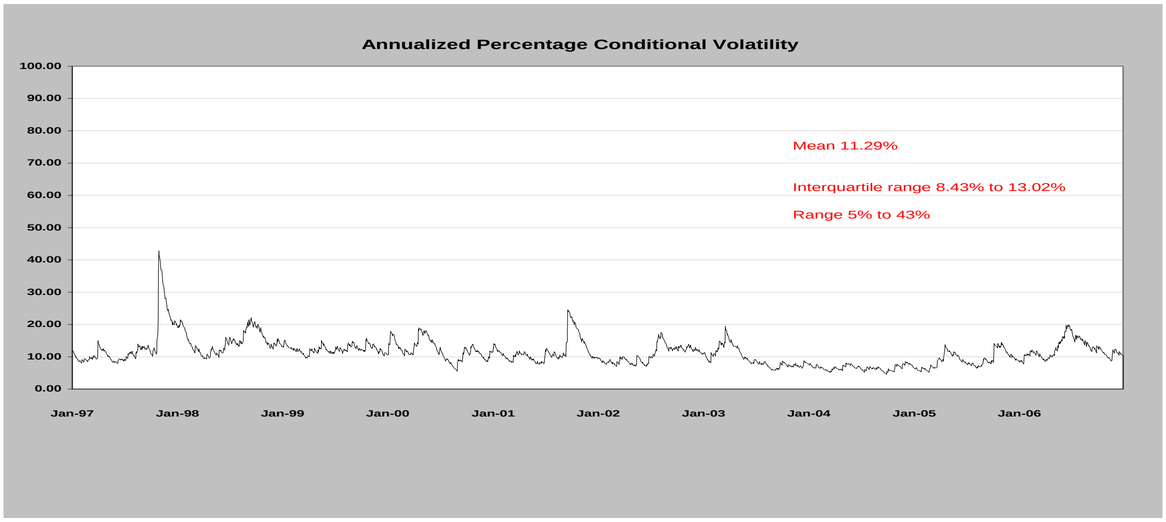


The volatility estimate shown in the adjacent figure ranges from 6% to 52%. Half of the estimates are inside the inter quartile range, from 9% to 13%. The median and mean values are 10.54% and 11.59% respectively. The forecast function converges to 11.9% and half life H is estimated to be 32 trading periods.

***Quadratic*-GARACH** (1,1) model states that μ_t is a constant and h_t is a linear function of e^2_{t-1} and h_{t-1}. Quadratic-GARCH model includes the squared residual multiplied by a function of residual. It is defined as model of Sentana (1995).

$$h_t = \omega + \lambda e_{t-1} + \alpha e^2_{t-1} + \beta h_{t-1} = h_t = \omega - \frac{\gamma^2}{4\alpha} + \alpha\left[e_{t-1} + \frac{\gamma}{2\alpha}\right]^2 + \beta h_{t-1}$$

	Returns r(t)				*Model*			
Mean	0.000336		*Reparameterized*		**μ**	0.000000		
Standard Dev	0.007545		**γ**	0.000000	**ώ**	0.000000		
Skewness	-0.474941		**α**	0.064228	**α**	0.064228		
Kurtosis	6.648217		**α+β**	0.999900	**β**	0.935672		
Lag 1 correlation	-0.006920		**Log L**	**8969.4262**	*Reparameterized*			
Date	Price	Return	Variance	Std. Residual		Log density	Annualized percent Voltatility	Percentage return
	p(t)	r(t)	h(t)	z(t)		l(t)		
2-Jan-97	2411.20							
3-Jan-97	2399.50	-0.004864	0.000057	-0.644730		3.7602	12.00	-0.49
6-Jan-97	2409.80	0.004283	0.000055	0.578711		3.8197	11.77	0.43



The volatility estimate shown in the above figure ranges from 5% to 43%. Half of the estimates are inside the inter quartile range, from 8.43% to 13.02%. The median and mean values are 10.18% and 11.29% respectively. The forecast function converges to 12% .

Conclusions

Summary statistics were obtained on annualized volatilities as shown by the above respective graphs. The average annual volatility of 11.59% was found to be highest due to GJR(1,1) model. 50% of the annual volatilities lies in the approx. range of 8% to 13% in all the models.

On examining the three ARCH models, the parameters estimates were obtained as shown in their respective worksheets. The maximum log-likelihood values were determined as **9001.8603 in GARCH (1,1),** 9022.1437 in GJR (1,1) & **8969.4262 in Quadratic GARCH (1,1).**

To test the Null Hypothesis H_0: Standardized residuals are i.i.d., Autocorrelation estimates using the portmanteau Q-statistic of Box and Pierce (1970) were used and the values were 6.7946, 6.8091 and 6.0728 for GARCH (1,1), GJR(1,1) and quadratic GARCH (1,1) models respectively. Later these values were compared by $\chi 2$ (value 11.0705) with 5 d.o.f. and were found insignificant at 5% significance level. **Thus we do not reject the null hypothesis that the standardized residuals of all the models are i.i.d.**