FEA Working Paper No. 2001-1

ACCURACY OF MODEL SELECTION CRITERIA FOR A CLASS OF AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTIC MODELS

Kwek Kian Teng

Faculty of Economics & Administration

University of Malaya

50603 Kuala Lumpur

MALAYSIA

E-Mail: g2kwek@umcsd.um.edu.my

December 2000

ACCURACY OF MODEL SELECTION CRITERIA FOR A CLASS OF AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTIC MODELS

Abstract

Many empirical studies relying on standard model selection criteria (*e.g.*, Akaike information criterion (AIC) and Schwarz Bayesian information criterion (BIC)) favourably choose the Generalized autoregressive conditional heteroscedastic GARCH(1,1) model for characterizing financial market volatility. This paper offers some insight into the accuracy of selecting a GARCH(1,1) model when the true model is an autoregressive conditional heteroscedastic (ARCH) model. Moreover, Grose and King (1993) found that a selection criterion for autoregressive moving-average (ARMA) models can biasedly select a particular model due to the shape of the likelihood function. Monte Carlo simulations are used to illustrate the unbiased selection of the GARCH(1,1) model. Our study suggests that GARCH(1,1) models are not unduly favoured when ARCH models are the true models.

1. INTRODUCTION

"Do the information criteria (IC)-based procedures favour GARCH(1,1) models?" It has been widely observed that in a lot of empirical work, GARCH(1,1) models seem to be selected regularly. One such study that quoted the GARCH(1,1) model as a favourite among applied research is Bera and Higgins (1993, p.317) who stated:

> *"In applied work, it has been frequently demonstrated that the GARCH(1,1) process is able to represent the majority of financial time*

> *series. A data set which requires a model of order greater than GARCH(1,2) or GARCH(2,1) is very rare."*

Bera and Higgins (1993, p.321) also stated that any model selection procedure, such as AIC and BIC, which penalizes a model for additional parameters, would select the GARCH(1,1) specification over the ARCH(6) specification. Brooks and Burke (1998) also found that BIC picks the GARCH(1,1) model nearly half the time for compounded percentage exchange rate returns from March 1973 to September 1989 on the German Mark. Another quote by Lee and Hansen (1994, p.29) along the same lines is:

> *"The Gaussian GARCH(1,1) model has become the workhorse of the industry, with the largest number of applications."*

Lamoureux and Lastrapes (1991), and Day and Lewis (1992) also found the GARCH(1,1) model being the best in terms of accuracy in the forecast-based comparison of implied options variances. In Bollerslev, Chou and Kroner's (1992, p.20) discussion on West, Edison, and Cho's (1993) paper, they reported that the GARCH(1,1) formulation out-performed the alternative models investigated and thus argued,

> *"... that an investment adviser whose only specialized tool is GARCH may be as worthy of her hire as are professionals currently on Wall Street."*

Bollerslev, Engle and Nelson (1994), and Andersen and Bollerslev (1998) fitted a MA(1)-GARCH(1,1) model to daily data on U.S. dollar/Deutsche-mark exchange rate from January 1981 to July 1992 and Deutsche-mark/U.S. dollar spot exchange rate from March 1979 to September 1993. For other studies on GARCH models see Baillie and Bollerslev (1989), King, Sentana and Wadhwani (1994), and Lamoureux and Lastrapes (1990a and 1990b). To quote Lamoureux and Lastrapes (1990a, p.223):

> *"We restrict our attention to a GARCH(1,1) specification since it has been shown to be a parsimonious representation of conditional variance that adequately fits many time series (e.g., Bollerslev, 1987)."*

So why are GARCH(1,1) models favoured by practitioners? This is an interesting question particularly when one recalls the work of Grose and King (1993), who compared model selection procedures in selecting between autoregressive (AR) and moving average (MA) disturbances in the linear regression model. This is a situation in which the number of parameters is the same for both models and therefore the penalties of the existing IC procedures are always the same. They found MA(1) errors are unduly favoured. In some circumstances the bias was extreme. For example, there is a 90% chance of choosing MA(1) when AR(1) is the true model with a small parameter. Could this be the case for the ARCH(2) and GARCH(1,1) models?

A reason that a particular model can be biasedly preferred is because the maximizing values favour a particular shape or functional form of the log-likelihood. This factor is primarily related to the shape of the likelihood function. As such, this can lead to an unfair favouring of a specific model. This issue motivated us to look for evidence as to whether the shape of the conditional likelihood function would favour a particular model, in particular the GARCH(1,1) model.

I therefore investigated the performance of the model selection procedures using extensive Monte Carlo simulations when a GARCH(1,1) process is included as a possible candidate in the portfolio of ARCH models. Specifically, I only considered the GARCH(1,1) process as the empirical evidence suggests there is a tendency to favour the GARCH(1,1) model above other GARCH models.

An outline of this paper is as follows. §2 reviews some existing information criteria. §3 describes the Monte Carlo experiments to test the various standard model selection criteria, (*i.e.*, whether the GARCH(1,1) model is actually favoured when ARCH(q) models are the true models). §4 discusses the results of our study by comparing the performance of some existing IC procedures. We conclude in §5.

2. INFORMATION CRITERIA

In this study I include seven different existing criteria to investigate whether there exists any bias in the selection of GARCH(1,1) models in the context of ARCH model selection. I considered the IC procedures of Akaike's AIC, Schwarz's BIC, Hannan and Quinn HQIC, Theil's RVC, Amemiya's PC, Hocking's S_p, and GCV or generalized cross validation approach.

Akaike (1973) was the first to propose AIC, a statistic incorporating Kullback-Leibler information with the use of maximum likelihood principles and negative entropy. It was developed from the work of selecting the best order of an AR process according to the minimum final prediction error (FPE) (Akaike, 1969) criterion, *i.e.*, FPE(k') = $(n+k')/(n-k')\hat{\sigma}^2_{k'}$, where n is sample size, k' is the order of the AR process fitted to the data, and $\hat{\sigma}^2_{k'}$ is the ML estimate of the residual variance. Different authors write the AIC in different forms. I give the penalized maximized log-likelihood expression. AIC chooses the value of k, the number of parameters, that maximizes Akaike's information:

$$\text{AIC} = \ell_J(\hat{\theta}) - k, \tag{1.1}$$

where θ is the combined parameter vector to be estimated in the model and $\ell_J(\hat{\theta})$ is the maximized log-likelihood. For a linear regression model, AIC is equivalent to:

$$\min_k \quad \ln(\hat{\sigma}^2_k) + \frac{2k}{n}. \tag{1.2}$$

where $\hat{\sigma}^2_k$ is estimated the maximum likelihood estimate of the residual variance for k parameters.

Schwarz's (1978) criterion is often also referred to as Schwarz's Bayesian Information Criterion or BIC. It provides a Bayesian solution to the model selection problem. Schwarz's BIC was derived using the posterior-probability criterion and evaluating the leading terms of its asymptotic expansion as an alternative estimator for the information quantity $I[\theta_o,\theta]$. BIC chooses the model which maximizes:

$$\text{BIC} = \ell_J(\hat{\theta}) - \frac{1}{2}k \ln n. \tag{1.3}$$

Schwarz assumed a fixed penalty for guessing the wrong model and considered an infinite sequence of nested models, each of which has a non-zero prior probability. When the number of observations is large, BIC penalizes additional parameters much more than AIC, leading to more parsimonious models. In large samples, BIC is equivalent to a Bayesian procedure that selects the model with the highest posterior probability. For the linear regression model, (1.3) is equivalent to: $\min_k \quad \ln(\hat{\sigma}_k^2) + (k\ln n)/n$.

Another estimator for the information quantity $I[\theta_o,\theta]$ that follows a similar approach to Akaike (1969, 1973) is Hannan and Quinn's (1979) penalized maximized log likelihood form which can be expressed as,

$$\text{HQIC} = \ell_J(\hat{\theta}) - k \ln\ln n. \qquad (1.4)$$

Theil's (1961) $\overline{R}^2$ was expressed in the following penalized maximized log likelihood form as:

$$\text{RVC} = \ell_J(\hat{\theta}) + \frac{1}{2} n \ln(n-k), \qquad (1.5)$$

where RVC is residual variance criterion. In this case, Theil's strategy is to choose the model that maximizes RVC.

Amemiya's (1972, 1980) Prediction Criterion (PC) was derived as an alternative method to estimate the variance within a hypothesis testing framework. Expressing PC as the average prediction variance based on regression models, I choose the model that minimizes: $\hat{\sigma}_k^2\left((n+k)/(n-k)\right)$. It is interesting to note that both FPE and PC evaluate the mean squared prediction error of the predictor derived from each model. Thus PC is expressed as:

$$\text{PC} = \ell_J(\hat{\theta}) - \frac{1}{2} n \ln(n+k) + \frac{1}{2} n \ln(n-k). \qquad (1.6)$$

Hocking's (1976) S_p criterion was reviewed by Thompson (1978), and was given an alternative justification by Breiman and Freedman (1983). S_p was developed under the assumption that the regressors are stochastic. Breiman and Freedman (1983) noted that the optimal properties of S_p depend on the assumption that the number of predictors in the model is not finite. In the regression model, it can also be written as, $\min_k \quad \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \big/ [(n-k)(n-k-1)]$. In the logarithmic form, S_p is:

$$\min_k \quad \ln(\hat{\sigma}_k^2) - \ln(n-k)(n-k-1).$$

And can be generalized in the penalized log-likelihood form as,

$$S_p = \ell_J(\hat{\theta}) + \frac{1}{2} n \ln(n - k) + \frac{1}{2} n \ln(n - k - 1). \qquad (1.7)$$

In the regression model, GCV is also equivalent to minimizing $\sum (y-\hat{y})^2 \big/ n(1-(k/n))^2$. In the logarithmic form, GCV chooses the model via

$$\min_k \ \ln(\hat{\sigma}_k^2) - 2\ln(1-\frac{k}{n}) .$$

Also see Golub, Heath and Wahba (1979) for more details on the GCV method. Fox (1995) expressed the generalization of GCV in the penalized likelihood form which was derived from Craven and Wahba (1979) by imposing certain assumptions, *i.e.*, choosing the model which maximizes:

$$GCV = \ell_J(\hat{\theta}) + n \ln(1 - k/n). \qquad (1.8)$$

3. THE MONTE CARLO EXPERIMENTS

A Monte Carlo experiment is constructed by generating four data generating processes (DGPs) consisting of ARCH models, but in the estimation process, a GARCH(1,1) model is also included in the portfolio of ARCH in order to see if any of the existing IC procedures (Akaike’s AIC, Schwarz’s BIC, Hannan and Quinn HQIC, Theil’s RVC, Amemiya’s PC, Hocking’s S_p, GCV or generalized cross validation approach) unduely favour the GARCH(1,1) model even though it is not the true model.

The model in this study is: $y_t = a + \varepsilon_t$, with the error modelled as $\varepsilon_t = \sqrt{h_t}\,\eta_t$, where η_t has a standard normal density and h_t is

$$h_t = \alpha_o + \sum_{j=1}^{q} \alpha_j \varepsilon_{t-j}^2 . \qquad (1.9)$$

Thus, I generated four DGPs of ARCH(q), max(q) = 4, where the autoregressive parameters satisfied the non-negativity and stationarity constraints:

$$\alpha_o > 0,$$

$$\alpha_j \geq 0, \ j = 1,...,q,$$

$$\sum_{j=1}^{q} \alpha_j < 1. \qquad (2.0)$$

The parameter α_j values were drawn using unrestricted ϕ_j values from a $U(0.1,1)$ prior distribution such that the restricted α_j values fulfil the stationarity conditions. I fitted ARCH(1), ARCH(2), ARCH(3), ARCH(4) and GARCH(1,1) models to each of the four simulated data sets generated from DGPs for $j = 1,..., 4$, using the maximum likelihood estimation. These experiments were simulated with 2,000 replications for each DGP. This were conducted for sample size of 60, 80, 100, 150, 200, 300, 400 and 500.

4. SIMULATION RESULTS

With the maximized conditional log-likelihood functions $\ell_j(\hat{\theta}(k))$ obtained from the estimation process, I penalized these likelihoods using AIC, BIC, HQIC, RVC, PC, S_p, and GCV. I then calculated the frequency count of correct and incorrect model selection of these different penalized maximized conditional log-likelihoods. These selection probabilities of correct and incorrect selection are tabulated in Tables 3 - 6 for n = 60, 80, 100, 150, 200, 300, 400, 500.

The overall selection probabilities are summarized in Table 1 which gives the average probabilities of correct selection (APCMS) evaluated over all four models and eight samples of different size n, denoted as $APCMS(J,n)$.

Table 2 **Average Probabilities of Correct and Incorrect Model Selection for a Portfolio of ARCH Models with Normal Errors Using *U*(0.1,1) Prior to Generate θs: Overview Selection by AIC, BIC, HQIC, RVC, PC, S_p, and GCV**

True Model	Fitted Model	AIC	BIC	HQIC	RVC	PC	S_p	GCV
Average Probabilities of Model Selection Based on Prior *U*(0.1,1)								
ARCH(1)	ARCH(1)	0.9292	0.9931	0.9748	0.8252	0.9292	0.9320	0.9316
	ARCH(2)	0.0479	0.0063	0.0198	0.0954	0.0479	0.0466	0.0468
	ARCH(3)	0.0179	0.0004	0.0044	0.0494	0.0179	0.0170	0.0171
	ARCH(4)	0.0048	0.0001	0.0010	0.0294	0.0048	0.0042	0.0043
	GARCH(1,1)	0.0002	0.0000	0.0000	0.0006	0.0002	0.0002	0.0002
ARCH(2)	ARCH(1)	0.1782	0.2992	0.2298	0.1234	0.1782	0.1813	0.1809
	ARCH(2)	0.6913	0.6289	0.6751	0.6806	0.6913	0.6909	0.6909
	ARCH(3)	0.0257	0.0029	0.0101	0.0607	0.0257	0.0246	0.0248
	ARCH(4)	0.0138	0.0004	0.0024	0.0476	0.0138	0.0127	0.0128
	GARCH(1,1)	0.0911	0.0686	0.0826	0.0877	0.0911	0.0905	0.0906
ARCH(3)	ARCH(1)	0.1019	0.1996	0.1428	0.0623	0.1019	0.1048	0.1043
	ARCH(2)	0.1866	0.2238	0.2069	0.1591	0.1866	0.1882	0.1883
	ARCH(3)	0.5480	0.3603	0.4648	0.6126	0.5478	0.5425	0.5429
	ARCH(4)	0.0247	0.0017	0.0071	0.0714	0.0246	0.0233	0.0235
	GARCH(1,1)	0.1388	0.2146	0.1784	0.0946	0.1391	0.1413	0.1411
ARCH(4)	ARCH(1)	0.0937	0.1898	0.1353	0.0494	0.0937	0.0962	0.0957
	ARCH(2)	0.1026	0.1315	0.1206	0.0865	0.1026	0.1039	0.1038
	ARCH(3)	0.2053	0.1451	0.1854	0.2134	0.2053	0.2038	0.2040
	ARCH(4)	0.3809	0.1491	0.2579	0.5153	0.3807	0.3714	0.3726
	GARCH(1,1)	0.2175	0.3845	0.3008	0.1353	0.2178	0.2246	0.2240
Grand Total of CMS		2.5494	2.1314	2.3726	2.6337	2.5490	2.5368	2.5380
$APCMS(J,n,f_j(\theta_j))$		0.63735	0.53290	0.59320	0.65843	0.63725	0.63420	0.63450

I summarized these results of averages of averages and ranked the order of selection from the largest to the smallest in Table 2.

Table 2 **Overall Rankings of AIC, BIC, HQIC, RVC, PC, S_p and GCV Based on the Average Probabilities of CMS [*APCMS*($J,n,f_j(\theta_j)$)] for Selecting Between ARCH Models Using Prior *U*(0.1,1) to Generate θs When the GARCH(1,1) Model is Included in the Portfolio of ARCH Models for Selection**

(1)	RVC	(65.843%)
(2)	AIC	(63.735%)
(3)	PC	(63.725%)
(4)	GCV	(63.450%)
(5)	S_p	(63.420%)
(6)	HQIC	(59.320%)
(7)	BIC	(53.290%)

Note: The values in brackets are the percentages of average probabilities of CMS (*APCMS*).

These rankings in Table 2 suggest that RVC out-performed the other selection criteria with a grand average probability of 65.84% of choosing the correct model. Second is AIC with 63.735%, third is PC with 63.725%, fourth is GCV with 63.45%, fifth is S_p with 63.42%, sixth is HQIC with 59.32% and last is BIC with 53.29%. Again AIC and PC differed by a marginal percentage as did GCV and S_p.

These results also suggest that the empirical evidence that favour the selection of the GARCH(1,1) model is not due to the statistical selection problem in IC because all these existing criteria could generally correctly choose the true ARCH processes given that a wrong GARCH(1,1) model was fitted. I highlight below some findings on a comparison between the grand average probabilities of incorrect selection of ARCH(2) and GARCH(1,1) processes when other ARCH processes are the DGPs.

When the DGPs are ARCH(1) and ARCH(3) processes, the probability of incorrectly choosing the ARCH(2) model is larger than the probability of incorrectly choosing the GARCH(1,1) model across all the selection criteria. That is, based on any one selection criterion[1], the average of the grand average probability of incorrectly selecting the ARCH(2) model is 4.44% as against 0.02% for the GARCH(1,1) model when the ARCH(1) model is true. When ARCH (3) is the true model, the probability of

[1] Calculations are based on Table 1.

incorrectly selecting the ARCH(2) model is 19.14% as against 14.97% for the GARCH(1,1) model.

Interestingly, when the true model is ARCH(4), the probability of incorrectly choosing the GARCH(1,1) model (with a grand average probability of 24.35%) outweighs the probability of incorrectly choosing the ARCH(2) model (with an average probability of 10.74%) across all selection criteria.[2] This confirms again the property of parsimonious selection when models are large.

When ARCH(2) is the true model, the grand average probabilities of correct selection are 69.13%, 69.13%, 69.09%, 69.09%, 68.06%, 67.51% and 62.89% for AIC, PC, GCV, S_p, RVC, HQIC and BIC, respectively. The selection probabilities for the GARCH(1,1) model are 9.11%, 9.11%, 9.06%, 9.05%, 8.77%, 8.26% and 6.86%, respectively, when ARCH(2) is the true model. Clearly, these model selection by the different criteria show that ARCH(2) model has been accurately chosen as the true model. Even though GARCH(1,1) also has the same number of parameters as ARCH(2) model, GARCH(1,1) model was never preferred by any of the existing IC.

The findings also show that when ARCH(4) is the true model[3], the average probability of incorrectly selecting a GARCH(1,1) model is 24.35% that outweighs that of incorrectly selecting an ARCH(2) model (10.74%) for any given one criterion.

It is worth noting that BIC consistently has largest correct selection for choosing ARCH(1) when it is the true model. Thus, BIC also has the smallest probability of incorrectly choosing the GARCH(1,1) model, that is, when the true model is an ARCH(1) model, BIC has zero probability of selecting the GARCH(1,1) model. This pattern also persists when the true model is an ARCH(2) model, BIC again has the smallest probability of 7.41% of incorrectly choosing a GARCH(1,1) model.

In summary, these findings seem to indicate that when an ARCH(2) model is the true model, the GARCH(1,1) model is not unduly favoured, or rather when it is not the true model, the GARCH(1,1) model will not be chosen excessively. For any one

[2] Calculated from Table 1.

criterion, the grand average probability of correctly selecting an ARCH(2) model is 67.84% as against a 7.30% chance of incorrectly selecting GARCH(1,1) model. So even if the number of parameters are equivalent, these IC can generally correctly pick the largest maximized conditional log-likelihood for the true model. The RVC had the largest grand average probability of correct selection (65.84%) as opposed to BIC, which has the smallest grand average probability of correct selection (53.29%).

5. CONCLUDING REMARKS

Empirical evidence in the ARCH and GARCH literature suggests that the GARCH(1,1) model might possibly be unfairly favoured in the model selection process. We investigated this issue and found that the results of our Monte Carlo study made under the assumption of normality, do not support this claim. That is our Monte Carlo results do not suggest that there is mistreatment for the selection of the true models. In fact, these findings lend support to the relative accuracy of model selection procedures for ARCH models.

In conclusion, the Monte Carlo simulation results show that in no way is a GARCH(1,1) model unduely favoured over an ARCH(2) model. Thus, when an ARCH(1) model is true, the probability of incorrectly choosing an ARCH(2) model is bigger than that for a GARCH(1,1) model. When an ARCH(2) model is true, the probability of choosing accurately a ARCH(2) model is bigger than that of choosing a GARCH(1,1) model, and when a ARCH(3) model is true, the probability of incorrectly choosing an ARCH(2) model is also bigger than that of incorrectly choosing a GARCH(1,1) model.

However, it is not surprising when an ARCH(4) model is true, the probability of incorrectly selecting an ARCH(2) model is less than that for incorrectly picking a GARCH(1,1) model[4]. The reason is because an infinite order ARCH model is better approximated by a finite order GARCH model. For our case, an ARCH(4) model is sometimes better represented by a GARCH(1,1) model. So as the lags get longer, the

---

[3] Also averaged across all criteria and calculated from Table 1.
[4] See Table 1.

ARCH coefficients get smaller and closer to zero values, thus making smaller GARCH processes more preferred by IC as these criteria favour more parsimonious models.

In general, the findings of this study seem to suggest that there is only a tendency to wrongly select a GARCH(1,1) model when the true model is ARCH with long lags. This is basically a model selection problem in choosing models that have fewer parameter but still do a good job of modelling the data.

Table 3 **Average Probabilities of Correct and Incorrect Model Selection by AIC, BIC, HQIC, RVC, PC, S_p, GCV for a Portfolio of ARCH Models with Normal Errors Using *U*(0.1,1) Prior to Generate θs: *n* = 60, 80**

True Model	Fitted Model	AIC	BIC	HQIC	RVC	PC	S_p	GCV
n = 60								
ARCH(1)	ARCH(1)	0.9380	0.9890	0.9650	0.8465	0.9380	0.9435	0.9425
	ARCH(2)	0.0415	0.0100	0.0260	0.0855	0.0415	0.0385	0.0385
	ARCH(3)	0.0140	0.0010	0.0070	0.0405	0.0140	0.0120	0.0130
	ARCH(4)	0.0055	0.0000	0.0020	0.0250	0.0055	0.0050	0.0050
	GARCH(1,1)	0.0010	0.0000	0.0000	0.0025	0.0010	0.0010	0.0010
ARCH(2)	ARCH(1)	0.4105	0.5590	0.4735	0.3090	0.4105	0.4200	0.4190
	ARCH(2)	0.4785	0.3790	0.4435	0.5200	0.4785	0.4765	0.4765
	ARCH(3)	0.0195	0.0025	0.0085	0.0450	0.0195	0.0165	0.0175
	ARCH(4)	0.0115	0.0005	0.0040	0.0370	0.0115	0.0095	0.0095
	GARCH(1,1)	0.0800	0.0590	0.0705	0.0890	0.0800	0.0775	0.0775
ARCH(3)	ARCH(1)	0.2980	0.4740	0.3790	0.1915	0.2985	0.3070	0.3055
	ARCH(2)	0.2445	0.2070	0.2320	0.2480	0.2440	0.2440	0.2445
	ARCH(3)	0.2980	0.1505	0.2235	0.4010	0.2980	0.2860	0.2875
	ARCH(4)	0.0135	0.0010	0.0045	0.0435	0.0125	0.0100	0.0110
	GARCH(1,1)	0.1460	0.1675	0.1610	0.1160	0.1470	0.1530	0.1515
ARCH(4)	ARCH(1)	0.2725	0.4500	0.3580	0.1625	0.2725	0.2820	0.2805
	ARCH(2)	0.1580	0.1450	0.1560	0.1550	0.1580	0.1615	0.1605
	ARCH(3)	0.1700	0.0840	0.1270	0.2225	0.1695	0.1625	0.1640
	ARCH(4)	0.1710	0.0525	0.1075	0.3000	0.1705	0.1545	0.1570
	GARCH(1,1)	0.2285	0.2685	0.2515	0.1600	0.2295	0.2395	0.2380
n = 80								
ARCH(1)	ARCH(1)	0.9305	0.9885	0.9665	0.8375	0.9305	0.9370	0.9365
	ARCH(2)	0.0510	0.0100	0.0295	0.0885	0.0510	0.0490	0.0495
	ARCH(3)	0.0155	0.0010	0.0035	0.0445	0.0155	0.0115	0.0115
	ARCH(4)	0.0030	0.0005	0.0005	0.0285	0.0030	0.0025	0.0025
	GARCH(1,1)	0.0000	0.0000	0.0000	0.0010	0.0000	0.0000	0.0000
ARCH(2)	ARCH(1)	0.3200	0.4695	0.3890	0.2285	0.3200	0.3255	0.3240
	ARCH(2)	0.5590	0.4625	0.5180	0.5840	0.5590	0.5585	0.5590
	ARCH(3)	0.0180	0.0030	0.0075	0.0485	0.0180	0.0155	0.0155
	ARCH(4)	0.0105	0.0010	0.0040	0.0385	0.0105	0.0085	0.0090
	GARCH(1,1)	0.0925	0.0640	0.0815	0.1005	0.0925	0.0920	0.0925
ARCH(3)	ARCH(1)	0.1970	0.3520	0.2625	0.1205	0.1970	0.2020	0.2015
	ARCH(2)	0.2310	0.2245	0.2345	0.2160	0.2315	0.2355	0.2350
	ARCH(3)	0.3850	0.2130	0.3020	0.4915	0.3835	0.3695	0.3710
	ARCH(4)	0.0160	0.0020	0.0055	0.0480	0.0160	0.0140	0.0140
	GARCH(1,1)	0.1710	0.2085	0.1955	0.1240	0.1720	0.1790	0.1785
ARCH(4)	ARCH(1)	0.1670	0.3175	0.2400	0.0895	0.1670	0.1725	0.1710
	ARCH(2)	0.1405	0.1485	0.1450	0.1260	0.1405	0.1430	0.1425
	ARCH(3)	0.2090	0.1200	0.1685	0.2325	0.2090	0.2015	0.2025
	ARCH(4)	0.2400	0.0670	0.1445	0.3890	0.2390	0.2240	0.2250
	GARCH(1,1)	0.2435	0.3470	0.3020	0.1630	0.2445	0.2590	0.2590

Table 4 **Average Probabilities of Correct and Incorrect Model Selection by AIC, BIC, HQIC, RVC, PC, S_p, GCV for a Portfolio of ARCH Models with Normal Errors Using *U*(0.1,1) Prior to Generate θs: *n* = 100, 150**

True Model	Fitted Model	AIC	BIC	HQIC	RVC	PC	S_p	GCV
n = 100								
ARCH(1)	ARCH(1)	0.9315	0.9915	0.9745	0.8355	0.9315	0.9345	0.9340
	ARCH(2)	0.0470	0.0075	0.0200	0.0920	0.0470	0.0460	0.0460
	ARCH(3)	0.0175	0.0005	0.0040	0.0460	0.0175	0.0165	0.0165
	ARCH(4)	0.0035	0.0005	0.0015	0.0250	0.0035	0.0025	0.0030
	GARCH(1,1)	0.0005	0.0000	0.0000	0.0015	0.0005	0.0005	0.0005
ARCH(2)	ARCH(1)	0.2610	0.4085	0.3235	0.1855	0.2610	0.2650	0.2650
	ARCH(2)	0.6170	0.5285	0.5910	0.6305	0.6170	0.6155	0.6155
	ARCH(3)	0.0190	0.0015	0.0060	0.0465	0.0190	0.0185	0.0185
	ARCH(4)	0.0120	0.0000	0.0010	0.0465	0.0120	0.0105	0.0105
	GARCH(1,1)	0.0910	0.0615	0.0785	0.0910	0.0910	0.0905	0.0905
ARCH(3)	ARCH(1)	0.1360	0.2790	0.1970	0.0765	0.1360	0.1410	0.1400
	ARCH(2)	0.2245	0.2375	0.2355	0.2005	0.2245	0.2275	0.2275
	ARCH(3)	0.4525	0.2595	0.3570	0.5550	0.4525	0.4455	0.4455
	ARCH(4)	0.0180	0.0015	0.0060	0.0520	0.0180	0.0175	0.0175
	GARCH(1,1)	0.1690	0.2225	0.2045	0.1160	0.1690	0.1685	0.1695
ARCH(4)	ARCH(1)	0.1160	0.2475	0.1695	0.0510	0.1160	0.1180	0.1175
	ARCH(2)	0.1195	0.1390	0.1345	0.1050	0.1195	0.1200	0.1200
	ARCH(3)	0.2115	0.1315	0.1835	0.2340	0.2115	0.2115	0.2110
	ARCH(4)	0.3030	0.1005	0.1890	0.4450	0.3030	0.2935	0.2955
	GARCH(1,1)	0.2500	0.3815	0.3235	0.1650	0.2500	0.2570	0.2560
n = 150								
ARCH(1)	ARCH(1)	0.9325	0.9930	0.9715	0.8365	0.9325	0.9335	0.9335
	ARCH(2)	0.0450	0.0060	0.0210	0.0885	0.0450	0.0445	0.0445
	ARCH(3)	0.0185	0.0010	0.0060	0.0485	0.0185	0.0185	0.0185
	ARCH(4)	0.0040	0.0000	0.0015	0.0265	0.0040	0.0035	0.0035
	GARCH(1,1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH(2)	ARCH(1)	0.1635	0.3030	0.2255	0.1110	0.1635	0.1680	0.1675
	ARCH(2)	0.7050	0.6240	0.6810	0.7020	0.7050	0.7025	0.7030
	ARCH(3)	0.0240	0.0030	0.0080	0.0545	0.0240	0.0240	0.0240
	ARCH(4)	0.0145	0.0005	0.0020	0.0465	0.0145	0.0135	0.0135
	GARCH(1,1)	0.0930	0.0695	0.0835	0.0860	0.0930	0.0920	0.0920
ARCH(3)	ARCH(1)	0.0610	0.1620	0.1035	0.0330	0.0610	0.0635	0.0625
	ARCH(2)	0.1860	0.2320	0.2095	0.1450	0.1860	0.1855	0.1865
	ARCH(3)	0.5830	0.3625	0.4835	0.6520	0.5825	0.5790	0.5790
	ARCH(4)	0.0180	0.0005	0.0050	0.0730	0.0180	0.0165	0.0165
	GARCH(1,1)	0.1520	0.2430	0.1985	0.0970	0.1525	0.1555	0.1555
ARCH(4)	ARCH(1)	0.0410	0.1175	0.0725	0.0190	0.0410	0.0420	0.0420
	ARCH(2)	0.0765	0.1130	0.0995	0.0535	0.0765	0.0790	0.0790
	ARCH(3)	0.2020	0.1460	0.1860	0.2030	0.2020	0.2030	0.2025
	ARCH(4)	0.4520	0.1590	0.2955	0.6000	0.4520	0.4405	0.4415
	GARCH(1,1)	0.2285	0.4645	0.3465	0.1245	0.2285	0.2355	0.2350

Table 5 **Average Probabilities of Correct and Incorrect Model Selection by AIC, BIC, HQIC, RVC, PC, S_p, GCV for a Portfolio of ARCH Models with Normal Errors Using *U*(0.1,1) Prior to Generate θs: *n* = 200, 300**

True Model	Fitted Model	AIC	BIC	HQIC	RVC	PC	S_p	GCV
n = 200								
ARCH(1)	ARCH(1)	0.9290	0.9930	0.9765	0.8235	0.9290	0.9310	0.9310
	ARCH(2)	0.0475	0.0070	0.0175	0.0970	0.0475	0.0460	0.0460
	ARCH(3)	0.0200	0.0000	0.0050	0.0510	0.0200	0.0195	0.0195
	ARCH(4)	0.0035	0.0000	0.0010	0.0285	0.0035	0.0035	0.0035
	GARCH(1,1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH(2)	ARCH(1)	0.1035	0.2365	0.1565	0.0620	0.1035	0.1045	0.1045
	ARCH(2)	0.7665	0.6945	0.7485	0.7465	0.7665	0.7670	0.7665
	ARCH(3)	0.0250	0.0040	0.0110	0.0610	0.0250	0.0245	0.0245
	ARCH(4)	0.0160	0.0005	0.0035	0.0530	0.0160	0.0155	0.0160
	GARCH(1,1)	0.0890	0.0645	0.0805	0.0775	0.0890	0.0885	0.0885
ARCH(3)	ARCH(1)	0.0280	0.0890	0.0490	0.0110	0.0280	0.0285	0.0285
	ARCH(2)	0.1490	0.2065	0.1830	0.1135	0.1490	0.1525	0.1520
	ARCH(3)	0.6640	0.4620	0.5730	0.7175	0.6640	0.6605	0.6610
	ARCH(4)	0.0240	0.0015	0.0070	0.0735	0.0240	0.0225	0.0225
	GARCH(1,1)	0.1350	0.2410	0.1880	0.0845	0.1350	0.1360	0.1360
ARCH(4)	ARCH(1)	0.0050	0.0255	0.0140	0.0025	0.0050	0.0050	0.0050
	ARCH(2)	0.0510	0.0935	0.0790	0.0365	0.0510	0.0515	0.0515
	ARCH(3)	0.1795	0.1610	0.1800	0.1425	0.1795	0.1805	0.1810
	ARCH(4)	0.5650	0.2305	0.3910	0.7105	0.5650	0.5570	0.5575
	GARCH(1,1)	0.1995	0.4895	0.3360	0.1080	0.1995	0.2060	0.2050
n = 300								
ARCH(1)	ARCH(1)	0.9265	0.9945	0.9800	0.8180	0.9265	0.9290	0.9285
	ARCH(2)	0.0445	0.0055	0.0140	0.0985	0.0445	0.0425	0.0430
	ARCH(3)	0.0210	0.0000	0.0055	0.0560	0.0210	0.0215	0.0215
	ARCH(4)	0.0080	0.0000	0.0005	0.0275	0.0080	0.0070	0.0070
	GARCH(1,1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH(2)	ARCH(1)	0.0560	0.1475	0.0950	0.0295	0.0560	0.0565	0.0565
	ARCH(2)	0.8110	0.7800	0.8095	0.7710	0.8110	0.8115	0.8115
	ARCH(3)	0.0340	0.0035	0.0135	0.0710	0.0340	0.0335	0.0335
	ARCH(4)	0.0155	0.0005	0.0030	0.0505	0.0155	0.0145	0.0145
	GARCH(1,1)	0.0835	0.0685	0.0790	0.0780	0.0835	0.0840	0.0840
ARCH(3)	ARCH(1)	0.0075	0.0405	0.0190	0.0040	0.0075	0.0085	0.0085
	ARCH(2)	0.0975	0.1570	0.1220	0.0690	0.0975	0.0990	0.0990
	ARCH(3)	0.7625	0.5690	0.6930	0.7805	0.7625	0.7605	0.7605
	ARCH(4)	0.0360	0.0040	0.0110	0.0865	0.0360	0.0355	0.0355
	GARCH(1,1)	0.0965	0.2295	0.1550	0.0600	0.0965	0.0965	0.0965
ARCH(4)	ARCH(1)	0.0030	0.0200	0.0060	0.0005	0.0030	0.0030	0.0030
	ARCH(2)	0.0185	0.0490	0.0305	0.0105	0.0185	0.0185	0.0185
	ARCH(3)	0.1160	0.1490	0.1455	0.0885	0.1160	0.1170	0.1170
	ARCH(4)	0.7290	0.3685	0.5665	0.8320	0.7290	0.7255	0.7255
	GARCH(1,1)	0.1335	0.4135	0.2515	0.0685	0.1335	0.1360	0.1360

Table 6 **Average Probabilities of Correct and Incorrect Model Selection by AIC, BIC, HQIC, RVC, PC, S_p, GCV for a Portfolio of ARCH Models with Normal Errors Using *U*(0.1,1) Prior to Generate θs: *n* = 400, 500**

True Model	Fitted Model	AIC	BIC	HQIC	RVC	PC	S_p	GCV
n = 400								
ARCH(1)	ARCH(1)	0.9290	0.9975	0.9835	0.8045	0.9290	0.9300	0.9295
	ARCH(2)	0.0495	0.0025	0.0140	0.1025	0.0495	0.0490	0.0495
	ARCH(3)	0.0160	0.0000	0.0020	0.0525	0.0160	0.0160	0.0160
	ARCH(4)	0.0055	0.0000	0.0005	0.0405	0.0055	0.0050	0.0050
	GARCH(1,1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH(2)	ARCH(1)	0.0335	0.1000	0.0645	0.0170	0.0335	0.0335	0.0335
	ARCH(2)	0.8345	0.8310	0.8455	0.7750	0.8345	0.8350	0.8350
	ARCH(3)	0.0355	0.0025	0.0170	0.0855	0.0355	0.0350	0.0350
	ARCH(4)	0.0165	0.0000	0.0015	0.0535	0.0165	0.0165	0.0165
	GARCH(1,1)	0.0800	0.0665	0.0715	0.0690	0.0800	0.0800	0.0800
ARCH(3)	ARCH(1)	0.0030	0.0220	0.0085	0.0020	0.0030	0.0030	0.0030
	ARCH(2)	0.0580	0.1220	0.0895	0.0375	0.0580	0.0580	0.0580
	ARCH(3)	0.8225	0.6610	0.7650	0.8145	0.8225	0.8230	0.8230
	ARCH(4)	0.0365	0.0020	0.0085	0.1010	0.0365	0.0355	0.0355
	GARCH(1,1)	0.0800	0.1930	0.1285	0.0450	0.0800	0.0805	0.0805
ARCH(4)	ARCH(1)	0.1160	0.2475	0.1695	0.0510	0.1160	0.1180	0.1175
	ARCH(2)	0.1195	0.1390	0.1345	0.1050	0.1195	0.1200	0.1200
	ARCH(3)	0.2115	0.1315	0.1835	0.2340	0.2115	0.2115	0.2110
	ARCH(4)	0.3030	0.1005	0.1890	0.4450	0.3030	0.2935	0.2955
	GARCH(1,1)	0.2500	0.3815	0.3235	0.1650	0.2500	0.2570	0.2560
n = 500								
ARCH(1)	ARCH(1)	0.9165	0.9980	0.9810	0.7995	0.9165	0.9175	0.9175
	ARCH(2)	0.0575	0.0020	0.0165	0.1105	0.0575	0.0575	0.0575
	ARCH(3)	0.0210	0.0000	0.0020	0.0560	0.0210	0.0205	0.0205
	ARCH(4)	0.0050	0.0000	0.0005	0.0340	0.0050	0.0045	0.0045
	GARCH(1,1)	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
ARCH(2)	ARCH(1)	0.0775	0.1695	0.1105	0.0450	0.0775	0.0775	0.0775
	ARCH(2)	0.7590	0.7315	0.7640	0.7155	0.7590	0.7605	0.7605
	ARCH(3)	0.0305	0.0035	0.0090	0.0735	0.0305	0.0295	0.0295
	ARCH(4)	0.0135	0.0000	0.0005	0.0555	0.0135	0.0130	0.0130
	GARCH(1,1)	0.1195	0.0955	0.1160	0.1105	0.1195	0.1195	0.1195
ARCH(3)	ARCH(1)	0.0845	0.1780	0.1235	0.0595	0.0845	0.0845	0.0845
	ARCH(2)	0.3025	0.4040	0.3495	0.2435	0.3025	0.3035	0.3035
	ARCH(3)	0.4165	0.2050	0.3215	0.4890	0.4165	0.4160	0.4155
	ARCH(4)	0.0355	0.0010	0.0095	0.0935	0.0355	0.0350	0.0355
	GARCH(1,1)	0.1610	0.2120	0.1960	0.1145	0.1610	0.1610	0.1610
ARCH(4)	ARCH(1)	0.0290	0.0925	0.0530	0.0195	0.0290	0.0290	0.0290
	ARCH(2)	0.1375	0.2250	0.1855	0.1005	0.1375	0.1380	0.1380
	ARCH(3)	0.3430	0.2380	0.3090	0.3505	0.3430	0.3430	0.3430
	ARCH(4)	0.2840	0.1145	0.1805	0.4010	0.2840	0.2830	0.2830
	GARCH(1,1)	0.2065	0.3300	0.2720	0.1285	0.2065	0.2070	0.2070

References

Akaike, H., 1969. "Fitting Autoregressive Models for Prediction", *Annals of the Institute of Statistical Mathematics*, 21: 243-247.

Akaike, H., 1973. "Information Theory and an Extension of the Maximum Likelihood Principle" in B.N. Petrov and F. Csaki (eds.), *Second International Symposium on Information Theory*, (Akademiai Kiado: Budapest), 267-281. Also in S. Kotz, and N.L. Johnson (eds.), *Breakthroughs in Statistics Volume 1: Foundations and Basic Theory*, (Springer-Verlag: New York), 601-609.

Amemiya, T., 1972. Lecture Notes on Econometrics, Mimeo, Department of Economics, Stanford University, Stanford, CA.

Amemiya, T., 1980. "Selection of Regressors", *International Economic Review*, 21: 331-354.

Andersen, T.G., and T. Bollerslev, 1998. "Deutsche Mark-Dollar Volatility: Intra-day Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies", *Journal of Finance*, 53: 219-265.

Baillie, R.T., and T. Bollerslev, 1989. "The Message in Daily Exchange Rates: A Conditional-Variance Tale", *Journal of Business and Economic Statistics*, 7: 297-305.

Bera, A.K., and M.L. Higgins, 1993. "ARCH Models: Properties, Estimation and Testing", *Journal of Economic Surveys*, 7: 305-366.

Bollerslev, T., R.Y. Chou, and K.F. Kroner, 1992. "ARCH Modelling in Finance, A Review of the Theory and Empirical Evidence", *Journal of Econometrics*, 52: 5-59.

Bollerslev, T., R.F. Engle, and D.B. Nelson, 1994. "ARCH Models", in R.F. Engle and D.L. McFadden (eds.), *Handbook of Econometrics*, *4*, (Elsevier Science: Amsterdam), 2959-3038.

Breiman, L., and D. Freedman, 1983. "How Many Variables Should be Entered in a Regression Equation?", *Journal of the American Statistical Association*, 78: 131-136.

Brooks, C., and S.P. Burke, 1997. "Large and Small Sample Information Criteria for GARCH Models Based on Estimation of the Kullback-Leibler", Discussion Papers in Quantitative Economics and Computing, No. 53, University of Reading, Reading, Berkshire.

Brooks, C., and S.P. Burke, 1998. "Forecasting Exchange Rate Volatility Using Conditional Variance Models Selected by Information Criteria", *Economics Letters*, 61: 273-278.

Craven, P., and G. Wahba, 1979. "Smoothing Noisy Data with Spline Functions: Estimating the Correct Degree of Smoothing by the Method of Generalized Cross-Validation", *Numerische Mathematik*, 31: 377-403.

Day, T.E., and C.M. Lewis, 1992. "Stock Market Volatility and the Information Content of Stock Index Options", *Journal of Econometrics*, 52: 267-287.

Fox, K.J., 1995. "Model Selection Criteria: A Reference Source", Unpublished Manuscript, Department of Economics, University of British Columbia and School of Economics, University of New South Wales, N.S.W.

Golub, G.H., M. Heath, and G. Wahba, 1979. "Generalized Cross-Validation as a Method for Choosing a Good Ridge Parameter", *Technometrics*, 21: 215-223.

Grose, S.D., and M.L. King, 1993. "The Use of Information Criteria for Model Selection between Models with Equal Numbers of Parameters", Working Paper No 20/93, Department of Econometrics, Monash University, Clayton, Victoria.

Hannan, E.J., and B.G. Quinn, 1979. "The Determination of the Order of an Autoregression", *Journal of Royal Statistical, Society, Series B*, 41: 190-195.

Hocking, R.R., 1976. "The Analysis and Selection of Variables in Linear Regression", *Biometrics*, 32: 1-49.

King, M., E. Sentana, and S. Wadhwani, 1994. "Volatility Links between National Stock Markets", *Econometrica*, 62: 901-933.

Kwek, K.T., and M.L. King, 1998. "Information Criteria in 'Conditional Heteroscedastic' Models: A Bayesian Prior Approach to Penalty Function Building", in C.L. Skeels (ed.), *Proceedings of the Econometric Society Australasian*, CD Rom, The Australian National University, Canberra.

Lamoureux, C.G., and W.D. Lastrapes, 1990a. "Heteroscedasticity in Stock Return Data: Volume versus GARCH Effects", *Journal of Finance*, 45: 221-229.

Lamoureux, C.G., and W.D. Lastrapes, 1990b. "Persistence in Variance, Structural Change, and the GARCH Model", *Journal of Business and Economic Statistics*, 8: 225-234.

Lamoureux, C.G., and W.D. Lastrapes, 1991. "Forecasting Stock Return Variance: Toward an Understanding of Stochastic Implied Volatilities", Unpublished Manuscript, Washington University, St. Louis, MO.

Lee, S.W., and B.E. Hansen, 1994. "Asymptotic Theory for the GARCH(1,1) Quasi-Maximum Likelihood Estimator", *Econometric Theory*, 10: 29-52.

Schwarz, G.W., 1978. "Estimating the Dimension of a Model", *Annals of Statistics*, 6: 461-464.

Theil, H., 1961. *Economic Forecasts and Policy*, 2nd ed. (North Holland Publishing Company: Amsterdam).

Thompson, M.L., 1978. "Selection of Variables in Multiple Regression", *International Statistical Review*, 46: 1-21 and 129-146.

West, K.D., H.J. Edison, and D. Cho, 1993. "A Utility-Based Comparison of Some Models of Exchange Rate Volatility", *Journal of International Economics*, 35: 23-45.

FEA Working Paper Series*

2000-1 Jones, G.W. "Education, Equity and Exuberant Expectations: Reflections on South-East Asia", September 1999.

2000-2 Goh, K.L. "Problems of the Wald Test in Non-Linear Settings and Some Solutions", October 1999.

2000-3 Begum, N. and M.L. King. "A New Approach to Testing a Composite Null against a Composite Alternative", January 2000.

2000-4 Noor Azina, I. and A.N. Pettitt. "Monitoring Hospital Outcomes Using Markov Chain Theory", July 2000.

2000-5 Khudayberganov, N. and S.F. Yap. "A Proposed Framework of Analysis for the Single Economic Environment: Prerequisites, Impact and Analysis in the Context of the European Integration", August 2000.

2001-1 Kwek, K.T. "Accuracy of Model Selection Criteria for a Class of Autoregressive Conditional Heteroscedastic Models", December 2000.

2001-2 Wong, H.K. and K.S. Jomo, "The Impact of Foreign Capital Inflows on the Malaysian Economy, 1966-96", January 2001.

* Papers are available at: http://www.cc.um.edu.my/FEP/

FEA Working Paper Series

Objective and Scope:

The Faculty of Economics and Administration (FEA) Working Paper Series is published to encourage the dissemination and facilitate discussion of research findings related to economics, development, public policies, administration and statistics. Both empirical and theoretical studies will be considered. The FEA Working Paper Series serves mainly as an outlet for research on Malaysia and other ASEAN countries. However, works on other regions that bear important implications or policy lessons for countries in this region are also acceptable.

Information to Paper Contributors:

1) Two copies of the manuscript should be submitted to:
   Chairperson
   Publications Committee
   Faculty of Economics and Administration
   University of Malaya
   50603 Kuala Lumpur
   MALAYSIA

2) The manuscript must be typed in double spacing throughout on one side of the paper only, and should preferably not exceed 30 pages of A4 size paper, including tables, diagrams, footnotes and references.

3) The first page of the manuscript should contain
   (i) the title,
   (ii) the name(s) and institutional affiliation(s) of the author(s), and
   (iii) the postal and email address of the corresponding author.
   This cover page will be part of the working paper document.

4) The electronic file of the manuscript must be submitted. The file can be a Word, Word Perfect, pdf or post-script document. This will be posted at the Faculty's website (http://www.cc.um.edu.my/FEP/) for public access.

5) Contents of the manuscript shall be the sole responsibility of the authors and publication does not imply the concurrence of the FEA or any of its agents. Manuscripts must be carefully edited for language by the authors. Manuscripts are vetted and edited, if necessary, but not refereed. The author is, in fact, encouraged to submit a concise version for publication in academic journals.

6) When published, the copyright of the manuscript remains with the authors. Submission of the manuscript will be taken to imply permission accorded by the authors for FEA to publicise and distribute the manuscript as a FEA Working Paper, in its hardcopy as well as electronic form.