

Forecasting exchange rate volatility using conditional variance models selected by information criteria

Article

Accepted Version

Brooks, C. and Burke, S. (1998) Forecasting exchange rate volatility using conditional variance models selected by information criteria. *Economics Letters*, 61 (3). pp. 273-278. ISSN 0165-1765 doi: [https://doi.org/10.1016/S0165-1765\(98\)00178-5](https://doi.org/10.1016/S0165-1765(98)00178-5) Available at <https://centaur.reading.ac.uk/35984/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

Published version at: [http://dx.doi.org/10.1016/S0165-1765\(98\)00178-5](http://dx.doi.org/10.1016/S0165-1765(98)00178-5)

To link to this article DOI: [http://dx.doi.org/10.1016/S0165-1765\(98\)00178-5](http://dx.doi.org/10.1016/S0165-1765(98)00178-5)

Publisher: Elsevier

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

NOTICE: this is the author's version of a work that was accepted for publication in *Economics Letters*. Changes resulting from the publishing process, such as peer review, editing, corrections, structural formatting, and other quality control mechanisms may not be reflected in this document. Changes may have been made to this work since it was submitted for publication. A definitive version was subsequently published in *Economics Letters*, 61.1 (1998), DOI: 10.1016/S0165-1765(98)00178-5

€"

Forecasting Exchange Rate Volatility using Conditional Variance Models Selected by Information Criteria

Chris Brooks¹, ISMA Centre, and Simon P. Burke, Department of Economics
University of Reading

2 July, 1998

Abstract

This paper uses appropriately modified information criteria to select models from the GARCH family, which are subsequently used for predicting US dollar exchange rate return volatility. The out of sample forecast accuracy of models chosen in this manner compares favourably on mean absolute error grounds, although less favourably on mean squared error grounds, with those generated by the commonly used GARCH(1,1) model. An examination of the orders of models selected by the criteria reveals that (1,1) models are typically selected less than 20% of the time.

J.E.L. Classifications: C22, C53, F31

Keywords: forecasting, GARCH, exchange rates, volatility, information criteria

¹ ISMA Centre, Department of Economics, P.O. Box 242, The University of Reading, Whiteknights, Reading RG6 6AA, England; t: (+44) 118 931 8239; f: (+44) 118 931 47 41; e: C.Brooks@rdg.ac.uk

Introduction

The use of GARCH models (a class first proposed by Bollerslev, 1986, and first applied to exchange rates by Hsieh, 1988), for modelling and predicting volatility is now very common in finance (see, for example, Akgiray, 1989, or Day and Lewis, 1992). A typical finding is that these models provide superior forecasts of volatility than those which simply use historical means of squared returns assuming homosceasticity. However, the vast majority of extant studies, restrict the conditionally heteroscedastic model to be GARCH(1,1) (see Bollerslev *et al.*, 1992, for a comprehensive survey of such papers). This approach seems arbitrary, and is certainly not grounded in financial or economic theory. Until recently, the only method of determining the appropriate orders for GARCH(r,m) models was by starting with a “large” model and testing down using a series of likelihood ratio-type restrictions (this procedure is used, for example, by Akgiray, 1989 and Cao and Tsay, 1992). However, Brooks and Burke (1997) have recently proposed a set of information criteria which allow the researcher to select an “optimal” in-sample model from the AR(p)-GARCH(r,m) class, where the maximum permitted orders of p , r , and m are specified in advance. The criteria are based upon estimation of the Kullback-Leibler discrepancy (see Sin and White, 1996). Models from this family can be expressed as

$$y_t = \mu + \sum_{i=1}^p \phi_i y_{t-i} + u_t$$
$$u_t \sim \text{NIID}(0, \sigma_t^2) \tag{1}$$

$$\sigma_t^2 = \alpha_0 + \sum_{j=1}^m \alpha_j u_{t-j}^2 + \sum_{k=1}^r \beta_k \sigma_{t-k}^2$$

The purpose of this paper is to determine whether the new information criteria lead to the selection of models which give improved out of sample forecasting performance compared with GARCH(1,1) models. To this end, we use exactly the same data as West and Cho (1995), henceforth WC, who found that GARCH models gave slightly more accurate forecasts than homoscedastic, IGARCH, autoregressive volatility, or nonparametric models.

Data and Methodology

The data are a set of weekly continuously compounded percentage exchange rate returns on the Canadian dollar, German mark, and Japanese yen, all against the US dollar². The sample covers the period 21 March 1973 - 20 September 1989, making a total of 863 observations. More details and summary statistics are provided by WC.

Following WC, the first half of the data, 432 observations, are used for in-sample model estimation, while the remainder of the observations are retained for out of sample forecasting. The information criteria select the appropriate model order (from AR(0)-GARCH(0,0) up to AR(5)-GARCH(5,5)) for the sample, and then forecasts are generated using these chosen models. The criteria thus have a total of 186 models from which to choose. Moreover, the criteria are jointly selecting the conditional mean and variance model orders, whereas WC do not allow for any structure beyond a constant in the mean. The criteria thus have a total of 186 models from which to choose. A fixed window of length 432 is then rolled through the data one point at a time, optimal model orders are determined and conditional forecasts generated again, and so on³. All models are estimated using quasi-maximum likelihood. In common with WC and Many other studies in this area, the *ex post* realised volatility measure that the models are trying to forecast is the square of the daily log-returns. The results from computing 1, 12, and 24 step ahead forecasts generated in this manner are compared, on mean squared error (MSE) and mean absolute error (MAE) grounds, to those generated using a GARCH(1,1) model - WC's best model overall. A total of 431 1-step ahead, 420 2-step ahead, and 408 24-step ahead forecasts are constructed. Although it seems sensible to evaluate conditional mean

² WC also consider the Italian lira / dollar exchange rate, but subsequently drop it as a consequence of the estimated GARCH models exhibiting non-stationarity in the conditional variance (i.e. the sum of the α_i and β_k coefficients in equation (1) being greater than 1). We also encounter explosive model parameters for some of the models for the French franc and British pound, which leads to the multi-step ahead forecasts blowing up as the forecast horizon is extended. Hence for the same reason, the results for these exchange rates which were in WCs original data set, are not displayed here.

³ So 186 different orders of AR-GARCH model are being estimated and then forecasts constructed for 431 different samples of length 432 observations. This forecasting exercise therefore represents an extremely computer-intensive procedure which took over a week to estimate on a Silicon Graphics Onyx 2 supercomputer.

forecasts using a mean squared error-type measure, since this is the loss function that has been employed in-sample to determine the model parameters, the same is not necessarily true for the evaluation of conditional variance forecasts. Since the *ex post* measure of volatility which the models attempt to forecast is the square of the return, MSE in a sense represents the square of a square, so it is no longer clear that MAE is any less appropriate for forecast evaluation. Indeed, Makridakis and Hibon (1995) recommend MAPE as an evaluation measure over fourteen others for comparing the results from different forecasting models.

The criteria use given by equations (2) & (3)

$$HSIC = \sum_{t=1}^T \log(\tilde{\sigma}_t^2) + g \log(T) \quad (2)$$

$$HHQIC = \sum_{t=1}^T \log(\tilde{\sigma}_t^2) + 2g \log(\log(T)) \quad (3)$$

where $\tilde{\sigma}_t^2$ are the estimated conditional variances using the specified GARCH model, T is the sample size, and $g=p+r+m+2$ is the total number of estimated parameters. The criteria are modifications of the traditional criteria (Schwarz, 1978; Hannan and Quinn, 1979) used to select appropriate dimensions for linear models allowing for the increased number of parameters estimated (the penalty term), and the heteroscedastic nature of the data (denoted by an H in front of the acronym for the information criteria) - see Brooks and Burke (1997) for more details.

Results

Table 1 shows the MSE for the 1, 12, and 24 step forecasting horizons for models selected using both of the modified information criteria⁴, together with the MSE for forecasts generated always using a GARCH(1,1) model. It is clear on MSE grounds that the GARCH(1,1) model always outperforms those selected by the information criteria, irrespective of the forecast horizon. Typically, the MSE of

⁴ A similar modification of Akaike's information criterion was also computed, but did not yield very different results from the Hannan-Quinn criterion, and hence the results are not shown due to space constraints.

the models selected using information criteria are 20-30% higher than the GARCH(1,1). There is very little to choose between the criteria.

The mean absolute prediction errors shown in table 2, on the other hand, favour the new criteria. For one and twelve step ahead forecast horizons, the fixed (1,1) order model gives the poorest overall performance, although again there is very little evidence that either of the criteria should be preferred to the other. The criteria are not good at selecting models which can provide good long term (6 months ahead) forecasts, which is hardly surprising since the criteria optimise in-sample model selection, and as one forecasts further into the future, using in-sample information so heavily in determining model orders is likely to lead to performance degradation. The use of information criteria does, however, lead to an improvement of around 5% in the MAE for the one week ahead forecasts.

The first two rows of table 3 give the percentage of times that a GARCH(1,1) model is selected (irrespective of the AR order in the conditional mean) by the HSIC and HHQ criteria respectively for each series. It is clearly evident that a model other than GARCH(1,1) is always chosen more than 75% of the time, except for the Schwarz criterion which picks a GARCH(1,1) model nearly half the time for Germany. The final two rows of table 3 present the most frequently selected models, and then the percentage of times that this model was chosen for each of the two modified criteria. A notable feature of the results is the frequency with which an ARCH(4) model is chosen - often 40% of the time out of a total of 31 models (again ignoring conditional mean model orders) that the criteria could have chosen.

Conclusions

A set of modified information criteria have been used to select appropriate model orders for forecasting the conditional variance of weekly exchange rate returns. The criteria lead to models which generally provide more accurate forecasts on mean absolute error grounds at short forecasting

horizons than a fixed GARCH(1,1) model, although the GARCH(1,1) model is still preferable if the forecasts are evaluated using mean squared error. We consider that the results presented here suggest that more research into attempting to optimally select GARCH model orders rather than limiting the set of models for consideration to the (1,1) case, which is the current standard practice, could be a fruitful avenue for further research.

References

- Akgiray, V., 1989, Conditional Heteroscedasticity in Time Series of Stock Returns *Journal of Business* 62, 55-80
- Bollerslev, T. (1986) Generalised Autoregressive Conditional Heteroskedasticity *Journal of Econometrics* 31, 307-327
- Bollerslev, T., Chou, R.Y. and Kroner, K.F. (1992) ARCH Modelling in Finance: A Review of the Theory and Empirical Evidence *Journal of Econometrics* 52(5), 5-59
- Brooks, C. and S.P. Burke, 1997, Large and Small Sample Information Criteria for GARCH Models Based on Estimation of the Kullback-Liebler, Discussion Papers in Quantitative Economics and Computing No. 53, University of Reading
- Cao, C.Q. and R.S. Tsay, 1992, Nonlinear Time Series Analysis of Stock Volatilities *Journal of Applied Econometrics* 7 Supplement, S165-S185
- 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
- Hannan, E.J. and B.G. Quinn, 1979, The determination of the order of an autoregression. *Journal of the Royal Statistical Society, Series B*, 41, 190-195.
- Hsieh, D.A., 1988 The Statistical Properties of Daily Foreign Exchange Rates 1974-1983 *Journal of International Economics* 24, 129-145.
- Makridakis, S. M. and Hibon, 1995, Evaluating Accuracy, or Error, Measures INSEAD Working Paper 95/18/TM
- Schwarz, G., 1978, Estimating the dimension of a model. *Annals of Statistics*, 6, 461-464.
- Sin, C.-Y., and H. White, 1996, Information criteria for selecting possibly misspecified parametric models. *Journal of Econometrics*, 71, 207-225.
- West, K.D. and D. Cho, 1995, The Predictive Ability of Several Models of Exchange Rate Volatility *Journal of Econometrics* 69, 367-391

Acknowledgement

We would like to thank an anonymous referee for comments on an earlier version of this paper. We are also grateful to Kenneth West for providing the data used in this study, and for helpful discussions. The usual disclaimer applies.

Table 1: Mean Squared Prediction Errors

| | <u>Canada</u> | | <u>Germany</u> | | <u>Japan</u> | | <u>Overall</u> |
|--------------------------|---------------|--------|----------------|------|--------------|------|----------------|
| | Rank | MSE | Rank | MSE | Rank | MSE | Rank |
| Panel A: 1-week Horizon | | | | | | | |
| GARCH(1,1) | 1 | 0.0482 | 1 | 2.28 | 1 | 1.92 | 1 |
| HSIC | 2 | 0.0601 | 2= | 2.80 | 2= | 2.47 | 2 |
| HHQIC | 3 | 0.0604 | 2= | 2.80 | 2= | 2.47 | 3 |
| Panel B: 12-week Horizon | | | | | | | |
| GARCH(1,1) | 1 | 0.0491 | 1 | 2.31 | 1 | 1.98 | 1 |
| HSIC | 2= | 0.0582 | 2= | 2.82 | 2= | 2.40 | 2= |
| HHQIC | 2= | 0.0582 | 2= | 2.82 | 2= | 2.40 | 2= |
| Panel C: 24-week Horizon | | | | | | | |
| GARCH(1,1) | 1 | 0.0505 | 1 | 2.02 | 1 | 2.05 | 1 |
| HSIC | 3 | 0.0586 | 2 | 2.59 | 2 | 10.8 | 2 |
| HHQIC | 2 | 0.0581 | 3 | 2.60 | 3 | 14.6 | 3 |

Notes: MSE shown has been multiplied by a factor of 10^7 .**Table 2: Mean Absolute Prediction Errors**

| | <u>Canada</u> | | <u>Germany</u> | | <u>Japan</u> | | <u>Overall</u> |
|--------------------------|---------------|-------|----------------|------|--------------|------|----------------|
| | Rank | MAE | Rank | MAE | Rank | MAE | Rank |
| Panel A: 1-week Horizon | | | | | | | |
| GARCH(1,1) | 3 | 0.378 | 3 | 2.66 | 3 | 2.45 | 3 |
| HSIC | 1 | 0.355 | 1= | 2.56 | 1= | 2.32 | 1 |
| HHQIC | 2 | 0.356 | 1= | 2.56 | 1= | 2.32 | 2 |
| Panel B: 12-week Horizon | | | | | | | |
| GARCH(1,1) | 3 | 0.404 | 3 | 2.63 | 3 | 2.50 | 3 |
| HSIC | 1= | 0.341 | 1= | 2.57 | 1= | 2.25 | 1= |
| HHQIC | 1= | 0.341 | 1= | 2.57 | 1= | 2.25 | 1= |
| Panel C: 24-week Horizon | | | | | | | |
| GARCH(1,1) | 3 | 0.408 | 1= | 2.46 | 1 | 2.58 | 1= |
| HSIC | 2 | 0.348 | 1= | 2.46 | 2 | 3.82 | 1= |
| HHQIC | 1 | 0.347 | 3 | 2.48 | 3 | 4.20 | 3 |

Notes: MAE shown has been multiplied by a factor of 10^4 .**Table 3: Most Frequently Selected Conditional Variance Model Orders and Proportion of Times the Chosen Model is not a GARCH(1,1)**

| | <u>Canada</u> | <u>Germany</u> | <u>Japan</u> |
|---|---------------|----------------|--------------|
| % times GARCH(1,1) chosen by HSIC | 18.9 | 45.7 | 14.4 |
| % times GARCH(1,1) chosen by HHQIC | 13.9 | 23.2 | 13.2 |
| (GARCH Model order chosen most often by HSIC) | (0,3) | (1,1) | (0,4) |
| % times chosen | 37.8 | 45.7 | 49.0 |
| (GARCH Model order chosen most often by HHQIC) | (0,4) | (0,4) | (0,4) |
| % times chosen | 39.0 | 33.6 | 50.1 |