

## **Forecasting Stock Market Volatility Using (Non-Linear) Garch Models**

PHILIP HANS FRANSES AND DICK VAN DIJK  
*Erasmus University, Rotterdam, The Netherlands*

### **ABSTRACT**

In this paper we study the performance of the GARCH model and two of its non-linear modifications to forecast weekly stock market volatility. The models are the Quadratic GARCH (Engle and Ng, 1993) and the Glosten, Jagannathan and Runkle (1992) models which have been proposed to describe, for example, the often observed negative skewness in stock market indices. We find that the QGARCH model is best when the estimation sample does not contain extreme observations such as the 1987 stock market crash and that the GJR model cannot be recommended for forecasting.

**KEY WORDS** forecasting volatility; non-linear GARCH

### **INTRODUCTION**

A stylized fact of financial time series is that aberrant observations seem to cluster in the sense that there are periods where volatility is larger than in other periods. Typically, these volatile periods correspond to major (economic) events such as stock market crashes and oil crises. Although most evidence in empirical finance indicates that returns on financial assets seem unforecastable at short horizons (see e.g. Granger, 1992, for a recent survey), the current consensus is that the variance of returns can be predicted using particular time series models. Within this class of models, the Generalized Autoregressive Conditional Heteroscedasticity (GARCH) model proposed by Engle (1982) and Bollerslev (1986) seems to be the most successful (see Bollerslev, Chou and Kroner, 1992, for a survey of GARCH applications). Roughly speaking, in a GARCH process the error variances can be modelled by an Autoregressive Moving Average (ARMA) type process. A useful feature of the GARCH model is that it can effectively remove the excess kurtosis in returns.

A further stylized fact is that the distribution of returns can be skewed. For example, for some stock market indices, returns are skewed to the left, i.e. there are more negative than positive outlying observations. The intrinsically symmetric GARCH model cannot cope with such skewness and, hence, one can expect that forecasts and forecast error variances from a GARCH model may be biased for skewed time series. Recently, a few modifications to the GARCH model have been proposed, which explicitly take account of skewed distributions. In our paper we consider two such modifications, the Quadratic GARCH model (QGARCH) proposed by Engle and Ng (1993) (see also Sentana (1995) for a recent discussion of QGARCH models), and the model advocated in Glosten, Jagannathan and Runkle, 1992, to be abbreviated

as GJR<sup>1</sup>. The focus is the out-of-sample forecasting performance of these two non-linear models relative to the standard GARCH model, and on the performance of all three models relative to the simple random walk forecasting scheme. We limit our analysis to forecasting volatility and not the mean of the time series.

In the next section we present the models used in our modelling and forecasting exercise. In the third section we discuss the stock market data we use in our empirical study. In the fourth section we discuss some within-sample estimation results. In the fifth section we evaluate the forecasting performance of the GARCH, QGARCH and GJR models as well as the Random Walk model. In the final section we present conclusions.

### GARCH MODELS

Consider a stock market index  $I_t$  and its return  $r_t$ , which we construct as  $r_t = \log I_t - \log I_{t-1}$ . In this paper the index  $t$  denotes weekly observations. Although higher-order models are sometimes found to be useful, the dominant empirical model for  $r_t$  is the autoregression of order  $p$  with GARCH(1,1) disturbances (AR( $p$ ) – GARCH(1,1)), which can be expressed as

$$\begin{aligned}\phi_p(B)r_t &= \mu + \varepsilon_t, \text{ with } \phi_p(B) = 1 - \phi_1 B - \dots - \phi_p B^p \\ \varepsilon_t &\sim N(0, h_t) \\ h_t &= \omega + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1}\end{aligned}\tag{1}$$

where  $B$  is the backward shift operator defined by  $B^k x_t = x_{t-k}$ . The parameter  $\mu$  reflects a constant term, which in practice is typically estimated to be close or equal to zero. Furthermore,  $p$  is usually 0 or small, suggesting that there are usually no opportunities to forecast  $r_t$  from its own past. It is assumed that the solutions of the characteristic equation  $\phi(z) = 0$  lie outside the unit circle and that  $\omega, \alpha, \beta > 0$  and  $\alpha + \beta < 1$  (see Bollerslev, 1986).

All odd moments of  $\varepsilon_t$  in model (1) equal zero, and hence  $\varepsilon_t$  and  $r_t$  are symmetric time series with fat tails. For stock market time series, however,  $r_t$  may display significant negative skewness. This empirical stylized result seems to be due to the fact that stock market crashes occur more frequently and more quickly than stock market booms, and that the absolute size of crashes is much larger. Two simple classes of models that can cope with skewed returns are the Quadratic GARCH model proposed by Engle and Ng (1993) and the so-called GJR model, advocated by Glosten, Jagannathan and Runkle (1992). The QGARCH model differs from model (1) by

$$h_t = \omega + \alpha(\varepsilon_{t-1} - \gamma)^2 + \beta h_{t-1}\tag{2}$$

When  $\gamma$  takes a positive value, it is clear from model (2) that a negative  $\varepsilon_{t-1}$  value has a larger impact on  $h_t$ . The GJR model is also similar to model (1), but now the  $h_t$  process is given by

$$h_t = \omega + \alpha \varepsilon_{t-1}^2 + \delta D_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1}\tag{3}$$

<sup>1</sup> An alternative model that can generate skewed time series patterns is the so-called Exponential GARCH model (EGARCH) proposed in Nelson (1990). Although we have considered this model as a possibly useful candidate for our purposes, it was found not to be very useful for repeated forecasting exercises. Given the latter purpose of repeatedly specifying and estimating GARCH models, an obvious requirement of these models is that the estimation method is reasonably simple and that parameter convergence occurs reasonably quick. Unfortunately, it has been our experience while running the estimation procedures that parameter estimation of the EGARCH model can be tedious. In fact, in almost all cases, no quick convergence was obtained. Only when we were able to generate a particular set of starting-values with a precision of 8 digits could we obtain convergence.

where  $D_{t-1}$  is a dummy variable which takes a value of 1 when  $\varepsilon_{t-1} < 0$  and value of 0 when  $\varepsilon_{t-1} \geq 0$ . Similar to the QGARCH model, when  $\delta > 0$ , negative shocks will have a larger impact on  $h_t$  than positive shocks. Stationarity and stability of these models is discussed in the relevant references. The QGARCH and GJR models can improve upon the standard symmetric GARCH since they can cope with negative (or positive) skewness, the latter depending on the sign of the additional parameter.

## THE DATA AND RESEARCH METHOD

The data we analyse in this paper are weekly observed indices for the stock markets in Germany (DAX), The Netherlands (EOE), Spain (MAD), Italy (MIL) and Sweden (VEC). The data span 9 years, with the first observation being week 1 in 1986 and the last observation being week 52 in 1994. In our case, the weekly observations are taken as the values that are recorded on Wednesdays.

A summary of some characteristics of the  $r_t$  series is given in Table I. The number of observations equals 469 for all five stock markets. The mean and variance are all quite small. The excess kurtosis of the series exceeds 0, indicating the necessity of fat-tailed distributions to describe these variables. A key feature of Table I, however, is that the estimated measure of skewness is large and negative. Especially for the DAX and EOE indices, skewness is large in an absolute sense.

The approach taken in this paper is as follows. Since our main goal is to evaluate the volatility forecasting performance of the three GARCH models, we wish to consider a reasonably large hold-out sample. We thus choose four years of observations to estimate the various model parameters. Furthermore, since it is not *a priori* assumed that one model necessarily dominates other models over the whole sample, we repeat our modelling and forecasting exercise for different subsamples. We thus fit the models to a sample of four years, generate a one-step-ahead forecast, delete the first observation from the sample and add the next one, and generate again a one-step-ahead forecast. In order to evaluate possibly changing patterns over the years, we evaluate forecasting performance for the years 1990 to 1994. This results in an evaluation of five times about 52 forecasts. Since we wish to minimize the impact of outlying observations on forecast evaluation, we use the Median of Squared Error (MedSE) instead of the usual Mean SE.

Table I. Summary statistics of data on returns. The sample covers weekly observations for the years 1986 to 1994

Stock market	$n$	Mean ( $\times 10^{-4}$ )	Variance ( $\times 10^{-4}$ )	Skewness	Excess kurtosis
Germany (DAX)	469	9.257	7.889	-0.972	4.865
Holland (EOE)	469	9.335	6.451	-1.455	9.238
Spain (MAD)	469	22.225	9.333	-0.531	3.419
Italy (MIL)	469	6.981	9.937	-0.162	0.532
Sweden (VEC)	469	23.848	9.836	-0.348	5.103

Source of data: Datastream.

The weekly data concern Wednesday observations. The notation ( $10^{-4}$ ) means that the reported value should be multiplied by  $10^{-4}$ .

## WITHIN-SAMPLE ESTIMATION

In this section we present some within-sample estimation results to give an idea of the possible usefulness of (non-linear) GARCH models. To save space we consider the sample covering the period 1986 to 1989, which gives an indication of the typical estimated parameter values. Notice that this period contains the stock market crash in 1987, which effects the skewness of the data.

In Table II we report the relevant parameter estimates for the GARCH model (1). Next to the parameter estimates, we report the value of the AIC and the value of the Log Likelihood (LnL). We present these AIC and LnL values to compare models (1), (2) and (3). We further report the values of Box–Pierce statistics for  $\hat{\varepsilon}_i/h_i^{1/2}$  and  $\hat{\varepsilon}_i^2/h_i$  in order to check the empirical validity of the models. It is clear from Table II that the parameters  $\alpha$  and  $\beta$  in the GARCH(1,1) model are usually significant at the 5% level, and hence that the constant variance model can be rejected, at least within sample. Furthermore, for two cases the sum of the  $\alpha$  and  $\beta$  parameters is close to unity. However, for no index do we observe that  $\alpha + \beta > 1$ .

In Table III we report some estimation results for the non-linear variants of the GARCH(1,1) model. We report the estimates for the  $\gamma$  parameter in the QGARCH model and the  $\delta$  parameter in the GJR model. Although the results should be interpreted with care since not all non-normality may have been removed, the  $t$ -ratios in Table III indicate that the  $\gamma$  and  $\delta$  parameters are significant at the 10% level at least for the DAX, EOE and VEC indices. The possible usefulness of non-linear modifications to the linear GARCH model seems to be confirmed by the LnL values, although the AIC values do not suggest a clear favourite. Unreported estimates of  $\gamma$  and  $\delta$  in other samples shows that these parameters are not always relevant. Hence, the relevance of these parameters can depend on one or two observations being included in or deleted from the sample.

Table II. Estimation results for GARCH models for 1986–89

Index	$p$	Parameter estimates			Diagnostics <sup>a</sup>			
		$\omega$ ( $\times 10^{-3}$ )	$\alpha$	$\beta$	AIC	lnL	$Q_1(10)$	$Q_2(10)$
DAX	0	0.444 (3.55)	0.267 (5.17)	0.212 (1.24)	−7.00	447.93	13.0	10.1
EOE	0	0.185 (2.75)	0.283 (4.65)	0.488 (4.83)	−7.01	461.46	7.69	8.88
MAD	1	0.011 (1.43)	0.128 (3.40)	0.855 (23.14)	−6.78	442.99	2.59	7.49
MIL	1	0.038 (1.40)	0.129 (2.71)	0.820 (12.08)	−6.97	450.57	7.23	8.70
VEC	0	0.273 (2.32)	0.222 (2.71)	0.418 (1.94)	−6.91	456.98	4.11	1.85

**Notes:**

$t$ -ratio are given in parentheses. Estimation is carried out using our own program written in Gauss.

The model is  $\phi(B)r_t = \mu + \varepsilon_t | N(0, h_t)$ ,  $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$  where  $\phi(B)$  is an autoregressive polynomial of order  $p$  and  $r_t$  is the weekly observed return series.

<sup>a</sup>The diagnostics are the Akaike Information Criterion [AIC], which will be used to compare these models with those in Table III, the LogLikelihood (LnL) calculated for similar reasons, and  $Q_1(10)$  and  $Q_2(10)$ , which denote Box–Pierce test statistics for the significance of residual correlations up to lag 10 in the estimated scaled residuals  $\hat{\varepsilon}_i/h_i^{1/2}$  and the squares of these residuals, respectively. The AIC is calculated so that the model with the smallest AIC value is preferred.

Table III. Estimation results of non-linear GARCH models for 1986–9

Index	QGARCH			GJR		
	$\gamma$ ( $\times 10^{-1}$ )	AIC <sup>a</sup>	LnL <sup>a</sup>	$\delta$	AIC	LnL
DAX	0.022 (1.82)	-7.00	450.06	0.482 (4.52)	-7.00	456.20
EOE	0.115 (1.73)	-7.01	463.86	0.202 (1.85)	-7.01	463.04
MAD	-0.125 (-1.57)	-6.78	444.12	-0.104 (-1.05)	-6.78	443.20
MIL	-0.012 (-0.23)	-6.97	450.61	-0.023 (-0.32)	-6.97	450.42
VEC	0.199 (2.19)	-6.91	458.07	0.274 (2.02)	-6.91	457.79

Note:

The QGARCH model is given by  $\phi(B)r_t = \mu + \varepsilon_t$ ,  $\varepsilon_t \sim N(0, h_t)$ ,  $h_t = \omega + \alpha(\varepsilon_{t-1} - \gamma)^2 + \beta h_{t-1}$  and the GJR model by  $\phi(B)r_t = \mu + \varepsilon_t$ ,  $\varepsilon_t \sim N(0, h_t)$ ,  $h_t = \omega + \alpha \varepsilon_{t-1}^2 + \delta D_{t-1} - \varepsilon_{t-1}^2 + \beta h_{t-1}$ , where  $D_{t-1} = 1$  when  $\varepsilon_{t-1} \leq 0$  and  $D_{t-1} = 0$  when  $\varepsilon_{t-1} > 0$ , and where  $\phi(B)$  is an autoregressive polynomial of order  $p$  and  $r_t$  is the weekly observed return series.

$t$ -ratios are given in parentheses. Estimation is carried out using our own program written in Gauss.

<sup>a</sup> See Table II.

## OUT-OF-SAMPLE FORECASTING

To evaluate the ability of (non-linear variants of) GARCH models to adequately forecast the volatility in financial time series, we need a measure of the 'true volatility'. As is standard in related studies on forecasting volatility (see Day and Lewis, 1992; Pagan and Schwert, 1990), we use the measure  $w_t$ , defined by

$$w_t = (r_t - \bar{r})^2 \quad (4)$$

where  $\bar{r}$  is the average return over the last four years of weekly observations. As a measure of the one-step-ahead forecast error  $u_{t+1}$  we use

$$u_{t+1} = w_{t+1} - \hat{h}_{t+1} \quad (5)$$

where  $\hat{h}_{t+1}$  is generated using the  $h_t$  equations in (1), (2) and (3) with the estimates for the parameters  $\omega$ ,  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$  substituted. To compare the forecasts from the GARCH models with a naive no-change forecast, we also evaluate the forecast errors for the Random Walk (RW) model, for which forecasts are given by  $\hat{h}_{t+1} = w_t$ . Notice that we calculate one-step-ahead forecasts, i.e. we use the estimated models to forecast for  $t+1$  while the models are calibrated for four years of observations until and including  $t$ .

In Table IV we report the MedSE for the years 1990 to 1994. The results indicate that the QGARCH model is preferred in 11 of the 25 cases, and that the RW, GARCH and GJR models are preferred in 12, 2 and 0 cases, respectively. The results for the GJR model suggest that it is not a useful tool for forecasting. Overall, it seems that RW and QGARCH perform about equally well.

The forecasts for 1990 and 1991 are based on models which are fitted from estimation samples which include the 1987 stock market crash observation. From Table IV, we notice that for these two years, the RW model outperforms the other models in 8 of the 10 cases. On the other hand, when we consider the years 1992, 1993 and 1994, we find that the QGARCH model

Table IV. Out-of-sample forecasting performance of ARCH, QGARCH, GJR and random walk models for the volatility of stock market indices

Index	Model	Forecast period				
		1990	1991	1992	1993	1994
DAX	GARCH	4.395	3.359	1.520	1.775	2.854
	QGARCH	3.146	3.322	1.425	0.877	2.780
	GJR	178.417	13.821	5.585	6.717	75.495
	RW	3.566	0.585	0.230	1.013	3.219
EOE	GARCH	1.305	1.050	0.848	0.767	0.928
	QGARCH	1.497	1.076	0.707	0.552	1.006
	GJR	126.656	61.713	3.779	2.545	11.725
	RW	1.146	0.148	0.090	0.633	0.983
MAD	GARCH	3.249	1.118	2.717	2.076	2.423
	QGARCH	3.298	1.085	1.839	1.008	1.962
	GJR	116.116	125.261	23.442	16.193	47.612
	RW	1.112	1.062	0.859	1.214	5.021
MIL	GARCH	1.130	2.358	3.351	8.895	8.444
	QGARCH	1.083	1.943	2.349	6.047	8.479
	GJR	21.923	10.170	65.221	71.364	35.197
	RW	1.173	1.664	6.736	19.851	10.803
VEC	GARCH	3.027	2.396	3.040	6.067	3.098
	QGARCH	3.462	2.978	2.279	5.357	3.200
	GJR	48.520	218.769	67.610	3.497	4.476
	RW	2.483	1.104	2.554	1.621	3.103

*Note:*

As a measure of 'true volatility' we use  $(r_t - \bar{r})^2$ , where  $\bar{r}$  is the average return over  $t - 1, \dots, t$ -four years.

The models are estimated for samples of four years, where in each new sample the first observation is deleted and the next observation is included, and one-step ahead forecasts are generated for 1990, ..., 1994.

Each cell contains the Median of Squared Errors ( $\times 10^{-7}$ ).

outperforms its rivals in 9 of the 15 cases. Hence, the forecasting performance of the GARCH type models appear sensitive to extreme within-sample observations.

## CONCLUSIONS

Our forecasting results for five weekly stock market indices show that the QGARCH model can significantly improve on the linear GARCH model and the no-change forecasting model, in cases when the forecasting models are calibrated on data which exclude such extreme events as the 1987 stock market crash. The GJR model, on the other hand, cannot be recommended.

## ACKNOWLEDGEMENTS

A previous version of this paper was presented at the INFORMS meeting in Singapore (July 1995). We thank Olaf van Thull for his help with the data and computing efforts. The first author thanks the Royal Netherlands Academy of Arts and Sciences for its financial support. We also gratefully acknowledge the help of Tim Bollerslev and Hans Ole Mikkelsen with the computer programs. Finally, we thank the Editor and an anonymous referee for useful suggestions.

## REFERENCES

- Bollerslev, T., 'Generalized autoregressive conditional heteroskedasticity', *Journal of Econometrics*, **31** (1986), 307–27.
- Bollerslev, T., Chou, R. and Kroner, K., 'ARCH modeling in finance: a review of the theory and empirical evidence', *Journal of Econometrics*, **52** (1992), 5–59.
- Day, T. and Lewis, C., 'Stock market volatility and the information content of stock index options', *Journal of Econometrics*, **52** (1992), 267–87.
- Engle, R. F., 'Autoregressive conditional heteroskedasticity with estimates of the variance of UK inflation', *Econometrica*, **50** (1982), 987–1008.
- Engle, R. F. and Ng, V., 'Measuring and testing the impact of news on volatility', *Journal of Finance*, **48** (1993), 1749–78.
- Glosten, L., Jagannathan, R. and Runkle, D., 'On the relation between the expected value and the volatility of nominal excess return on stocks', *Journal of Finance*, **46** (1992), 1779–801.
- Granger, C. W. J., 'Forecasting stock market prices: Lessons for forecasters', *International Journal of Forecasting*, **8** (1992), 3–13.
- Nelson, D., 'Conditional heteroskedasticity in asset returns: a new approach', *Econometrica*, **59** (1990), 347–70.
- Pagan, A. R. and Schwert, G. W., 'Alternative models for conditional stock market volatility', *Journal of Econometrics*, **45** (1990), 267–90.
- Sentana, E., 'Quadratic ARCH models', *Review of Economic Studies*, **62** (1995), 639–61.

*Authors' biographies:*

**Philip Hans Franses** is a research fellow of the Royal Netherlands Academy of Arts and Sciences, affiliated to the Econometric Institute, Erasmus University Rotterdam. His research interests are model selection and forecasting in time series. He has published on these topics in, e.g., the *Journal of Forecasting*, *Journal of Marketing Research*, *Journal of Econometrics*, and the *Journal of Business and Economic Statistics*.

**Dick van Dijk** is a PhD student at the Tinbergen Institute, Erasmus University Rotterdam. His research topic is model selection in non-linear time series.

*Authors' address:*

**Philip Hans Franses** (Econometric Institute) and **Dick van Dijk** (Tinbergen Institute), Erasmus University, Rotterdam, PO Box 1738, NL–3000 DR Rotterdam, The Netherlands.