Modelling Outliers and Extreme Observations for ARMA-GARCH Processes*

Peter Verhoeven[a] and Michael McAleer[b]

[a)] School of Business, Curtin University of Technology, WA 6845

[b)] Department of Economics, University of Western Australia, WA 6907

[a] pverhoev@ecel.uwa.edu.au

[b]Michael.McAleer@uwa.edu.au, mcaleer@imes.boj.or.jp

Revised: January 2000

Abstract

This paper examines the effects of outliers on the AR(1)-GARCH(1,1) process with regard to: (i) the estimated parameters; (ii) the second and fourth moment conditions; and (iii) the associated forecast errors. It is found that outliers: (i) tend to dominate the (quasi) maximum likelihood estimates, resulting in larger ARCH and smaller GARCH estimates; (ii) may give rise to spurious AR(1) and ARCH effects when the outliers are extreme; (iii) significantly increase the frequency of violation of the fourth moment condition; (iv) significantly decrease the t-ratios of the GARCH estimates; (v) lead to significantly higher and more variable volatility forecasts; and (vi) decrease the persistence measures of the volatility process. Moreover, the effects of outliers on the parameter estimates are largely independent of their location in the estimation period. For all the time series investigated, we find that the i.i.d. assumption of the standardised residuals is not rejected at the 5% significance level. Adjusting outliers makes the distribution of the standardised residuals more normal, reducing negative skewness and excess kurtosis, thereby implying greater efficiency of the QMLE. We also find that the maximum excess kurtosis that can be captured by GARCH(1,1) under the assumption of conditional normality increases with the standard deviation (which is approximately 40). Adjusting outlying observations results in improved mean forecast accuracy for periods of low volatility and poorer mean forecast accuracy for periods of high volatility. These findings are missed by examining only the mean forecast error measures for the whole sample. Overall, the Chen and Liu method for detecting and adjusting outlying observations enabled an improved estimation of AR(1)-GARCH(1,1) processes, with less biased and more efficient parameter estimates, greater frequency of validation of the regularity conditions, and reduced forecast errors for periods of low volatility.

*The authors are most grateful to Shiqing Ling for many helpful discussions, to Dick van Dijk for making available his GAUSS program, and the Australian Research Council for financial support. This paper was revised while the second author was a Visiting Scholar from Abroad at the Institute for Monetary and Economic Studies, Bank of Japan.

1 Introduction

In finance and financial economics, volatility is a fundamental measure of risk, namely the relative movements in the price of a security, such as a stock index, exchange rate, commodity price or futures price, over time. The greater is the variation in price, the greater is the volatility. as the true underlying volatility of a security is unobservable, it must be estimated. While volatility can be expressed in different ways, the typical definition used in finance is given by the standard deviation of the returns of a security over a given period. Volatility is an essential input to the optimisation of financial models describing the expected risk-return trade-off. For example, it is a crucial input to mean-variance portfolio optimisation models and for the pricing of both primary and secondary derivative securities. Improved conditional volatility estimates and forecasts should, therefore, result in more accurate asset prices. It follows that it is of paramount importance to practitioners to use an adequate empirical model to measure the dynamics of volatility in financial securities.

Financial equity returns generally do not match the familiar bell-shaped normal distribution. Instead, long periods of relatively ordinary price movements are frequently interrupted by short periods of relatively large and sudden price movements[1], predominantly consisting of market crashes rather than rallies[2] (see Friedman and Laibson [1989]). A recent example of extreme price movements is

[1]Tsay [1999] shows that large price movements are typically related to business and economic news that occur infrequently and irregularly, such as interest rate changes.

[2]In analysing the occurrence of extreme values in the daily returns of the S&P 500 index from 3 July 1962 to 31 December 1997, Tsay [1999] observed that the daily returns of the S&P 500 index had been moving away from log-normality towards fat-tailed distributions, especially for negative returns.

the October 1997 stock market crash originating in Asia. On 27 October 1997, the Hang Seng Stock Index (HSI) dropped by 14.7%, both the Dow Jones Industrial Average (DJIA) and German Stock Index (DAX) by 7.2%, the Standard & Poor's 500 Composite Index (S&P 500) by 6.9%, and the Japanese Stock Index (Nikkei 225) by 4.4%[3]. A consequence of these outlying and extreme observations is the fat-tailed[4] distribution of returns.

Several models have been developed to accommodate the large numbers of outliers. First, the distribution generating the returns is stable but fat-tailed (for example, the Stable Paretian distribution (Mandelbrot [1963]), the Student-t distribution (Praetz [1972]; Blattberg and Gonedes [1974]), or the generalised error distribution (GED) (Baillie and McMahon [1989]). Second, the returns are generated by a mixture of normal distributions (Clark [1973]; Kon [1984]; Harris [1989]). Third, the returns follow a mixed diffusion jump process (Press [1967]). Finally, the dynamics in the variance of returns causes the fat tails (or heteroskedasticity)(Engle [1982]).

There is also overwhelming evidence that the tail behaviour of equity returns evolves over time[5]. In particular, absolute returns (which are a measure of volatility) have significant positive serial correlation over long lags, implying that they have long term memory (Ding *et al.* [1993]). This is known as volatility clustering,

[3]The fact that the drop in prices on 27 October 1997 was only about one-third as large as that of 19 October 1987 may be attributed to rules and regulations which were introduced after the 1987 crash for the DJIA. For example, circuit breakers, collars and sidecars were triggered for the first time on 27 October 1997, preventing the recurrence of the extreme 1987 outlier.

[4]The distribution is said to be fat-tailed when the tails of the distribution decline by a power not exponentially, so that not all moments are bounded; otherwise, the distribution is thin-tailed.

[5]Volatility in equity (%) returns has not increased significantly over the last thirty years.

in which large (small) returns are more likely to be followed by large (small) returns than by small (large) returns, but of unpredictable sign (Mandelbrot [1963]). Hence, the returns are not independently and identically distributed (i.i.d.) over time. The important implication is that financial market volatility is highly predictable.

Engle [1982] developed the autoregressive conditional heteroskedasticity (ARCH(p)) model, which was subsequently generalised by Bollerslev [1986] (GARCH(p,q)), as the usual ARMA models with constant variance and Gaussian tails are inadequate for most financial time series. In the majority of empirical examples, GARCH(1,1) has been shown to represent adequately the observed intertemporal dependencies in daily returns of most financial time series, and is currently the most popular volatility forecasting model[6]. The GARCH model has been applied extensively to numerous financial time series (see Bollerslev *et al.* [1992] and Bera and Higgins [1993] for detailed reviews of GARCH applications), and its popularity is due to the fact that it: (i) captures the persistence of volatility; (ii) can account partly for the fat-tails of the returns distribution; and (iii) is simple, and also mathematically and computationally straightforward[7].

In recent years, the increased activity in the derivatives markets and the need to price options accurately has led to greater attention being paid to modelling volatility. Since financial data contain many outlying observations, their analysis in time

[6]One frequently voiced criticism of GARCH-type models is that they are typically not based on theory. Thus, although they are successful empirically, they are statistical rather than economic models (Nelson and Foster [1994]).

[7]Nelson [1992] suggests another reason for the success of GARCH(1,1): if the process generating prices is approximately a diffusion, then there is so much information about second moments at high frequencies that even a misspecified GARCH model can produce good estimates of volatility.

series modelling and forecasting has become increasingly important. Outlier analysis in time series is concerned with the detection and accommodation of abnormal and extreme observations in the data, and analysing their influence on the mean and variance of financial time series (for an invaluable review of outliers, see Beckman and Cook [1983]). For example, outliers can have dominating and deleterious effects on the (quasi-) maximum likelihood estimates ((Q)MLE) of the ARMA and GARCH parameters[8], since these models typically assume that the observations are normally distributed[9]. Moreover, outliers may lead to model misspecification, poor forecasts and invalid inferences[10]. Despite the voluminous research on estimating volatility using GARCH-type models, very little attention has been paid to the effects of outlying observations in the data. Although GARCH-type models with normal (or fat-tailed) innovations seem to perform adequately in forecasting standard volatility, they perform poorly in predicting outlying observations.

Three papers that specifically examine outliers for the GARCH(1,1) process are Hotta and Tsay [1998], Franses and Ghijsels [1999] and Franses and van Dijk [1999]. Hotta and Tsay [1998] developed an outlier adjustment method for additive and innovational outliers, while Franses and Ghijsels [1999] and Franses and van Dijk [1999] applied the outlier method for additive outliers, as developed by Chen and Liu [1993a] for ARIMA-type models, to the GARCH(1,1) model. Using weekly returns of several stock indices from 1983 to 1993 (1986-1995), and a mov-

[8]These outliers are also called influential observations.

[9]In some circumstances, a fat-tailed distribution (e.g. a student t distribution with 5 degrees of freedom) may be more resistant to outliers than a normal distribution (Sakata and White [1998]).

[10]There is also increased interest in estimating the parameters of the outliers themselves (extreme value analysis).

ing window of four (ten) years, Franses and Ghijsels [1999] and Franses and van Dijk [1999] found that outliers significantly increase the constant and ARCH estimates, and decreased the GARCH estimates. Moreover, removal of outliers yielded significantly lower (up to 29%) forecast errors, when measured as mean (medium) forecast error (M(ed)SE), mean (medium) absolute error (M(ed)AE) and medium squared percentage error (MedSPE), over GARCH(1,1) models with normal or t-distributed innovations for the original returns. Using 20-minute Mark/USD spot exchange rates from 5 June 1989 to 19 June 1989, Hotta and Tsay [1998] also found that outliers significantly increased the intercept and ARCH estimates, and decreased the GARCH estimates.

This paper examines the effects of outliers and extreme observations on the properties of the AR(1)-GARCH(1,1) model by modelling outliers and extreme observations using the Chen and Liu [1993a] procedure. The analysis differs from previous studies in the following ways: (i) we also examine the effects of extreme observations; (ii) in addition to outliers, we investigate the effects of extreme observations on the regularity conditions, which have previously not been examined; (iii) our data sets contain daily returns which have considerably more outliers and extreme observations than in previous datasets; and (iv) we use three types of financial series, namely stock indices, foreign exchange rates and commodities, which may exhibit different conditional volatility behaviour.

The paper is organised as follows. Section 2 presents the AR(1)-GARCH(1,1) model. Section 3 describes outliers, and methods for accommodating them. Section 4 describes the data, evaluation of forecasting ability, and the impact of outliers on estimation, forecasting and regularity conditions. Some concluding comments are given in Section 5.

2 The AR(1)-GARCH(1,1) model

Consider the AR(1)-GARCH(1,1) model, where the conditional mean (or log-return) is given by

$$y_t = \mu + \varphi y_{t-1} + \varepsilon_t \tag{1}$$

where

$$\varepsilon_t = \eta_t \sqrt{h_t} \tag{2}$$

$\varepsilon_t \sim N(0, h_t)$

$\eta_t \sim i.i.d.N(0, 1)$

and the conditional variance of ε_t is given by

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

Sufficient conditions[11] for positivity of the conditional variance and the GARCH(1,1)

[11] These conditions were imposed by our estimation code.

process to exist are that $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$[12].

In the GARCH(1,1) model, volatility is given as a linear combination of the lagged unconditional and conditional variances. The α (or ARCH) parameter measures the impact of the innovation in the return in the previous period, giving rise to the more abrupt changes in volatility, and the β (or GARCH) parameter measures the impact of the previous conditional variance on its current value, giving rise to the more moderate changes in volatility. The weight on the lagged unconditional variance decays exponentially, that is, the weight attached to observation t-i-1 is a fixed proportion of the weight (β) attached to observation t-i . Hence, both α and β produce the clustering phenomenon, as can easily be seen in an alternative expression of GARCH(1,1):

$$h_t = \frac{\omega}{1-\beta} + \alpha \sum_{i=1}^{\infty} \beta^{i-1} \varepsilon_{t-i}^2. \tag{4}$$

Alternatively, GARCH(1,1) can be expressed as an ARMA(1,1) process for the unconditional variance by substituting the difference between the unconditional and conditional variances of the volatility process, $\nu_t = \varepsilon_t^2 - h_t$, in equation (3) and rearranging, to yield

$$\varepsilon_t^2 = \omega + (\alpha + \beta)\varepsilon_{t-1}^2 + \nu_t - \beta \nu_{t-1}. \tag{5}$$

Several statistical properties have been established for the GARCH(1,1) process in order to define the unconditional moments of $\{\varepsilon_t\}$ (see Bollerslev [1986]). In

[12]Nelson and Cao [1991] have derived some weaker inequality constraints for higher order GARCH(p,q) models.

general, the higher is the moment considered, the stronger is the condition and the less likely is it to be satisfied empirically. From (5), it is clear that a sufficient condition for the second moment of $\{\varepsilon_t\}$ to exist is that $(\alpha + \beta) < 1$. If this condition is met, $\{\varepsilon_t, h_t\}$ is strictly stationary and ergodic, with

$$E(\varepsilon_t) = 0 \tag{6}$$

$$E(\varepsilon_t^2) = \frac{\omega}{(1 - \alpha - \beta)} \tag{7}$$

and

$$cov(\varepsilon_t, \varepsilon_s) = 0 \tag{8}$$

for $t \neq s$.

Diebold [1988] showed that stationary models converge to normality, while non-stationary models do not converge to normality. Violation of the second-order stationarity condition does not necessarily imply non-stationarity of the process. If some weaker requirements (such as the log moment condition) are met, $\{\varepsilon_t, h_t\}$ may still be stationary even though $(\alpha+\beta)$ might be equal to or greater than unity, in which case $E(\varepsilon_t^2) = \infty$ (see Nelson [1990b]; Lee and Hansen [1994]; Lumsdaine [1995]). For example, Nelson [1990b] shows that when $\omega > 0$ and $h_t < \infty$, $\{\varepsilon_t, h_t\}$ is strictly stationary and ergodic if and only if $E[ln(\beta + \alpha\eta_t^2)] < 0$. A practical problem with this condition is that it is difficult to apply in practice because it is the mean value of a distribution of a random variable. A large number of simulations is typically required to obtain statistically significant values for η_t[13].

The term $(\alpha + \beta)$ measures the persistence of the conditional variance. It measures the speed at which future volatility converges to the long-run level, that is, towards the unconditional variance. The closer is the sum to unity, the slower is the innovation to returns to die out and remain important for future forecasts. If the sum equals unity, the volatility process is a unit root process and behaves like a random walk. However, if the persistence exceeds unity, the volatility process is explosive in that $E(\varepsilon_t^2) = \infty$. The second moment condition also implies that for covariance stationarity, there is a trade-off between the α and β coefficients.

A sufficient condition for the existence of the fourth moment of $\{\varepsilon_t\}$ is $(k\alpha^2 + 2\alpha\beta + \beta^2) < 1$ (Bollerslev [1986]) [14], where k is the conditional fourth moment of η_t. Under the assumption of conditional normality, $k \equiv E(\eta_t^4) = 3$, so that the regularity condition becomes $(3\alpha^2 + 2\alpha\beta + \beta^2) < 1$. If the fourth moment exists, then

$$E(\varepsilon_t^4) = \frac{3\alpha^2(1 + \alpha + \beta)}{(1 - \alpha - \beta)(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)} \tag{9}$$

so that the excess kurtosis becomes

$$k_e = (E(\varepsilon_t^4) - 3E(\varepsilon_2^2)^2)E(\varepsilon_t^2)^{-2} = \frac{6\alpha^2}{(1 - \beta^2 - 2\alpha\beta - 3\alpha^2)} \tag{10}$$

which is greater than zero, by assumption, thereby yielding fat-tails. The assump-

[13]This holds because η_t is the true error rather than the estimated error for a given sample.

[14]He and Terasvirta [2000] provide a more detailed characterization of the fourth moment structure of the GARCH(p,q) process. Ling and McAleer [1999b] clarify the necessity and sufficiency of He and Terasvirta's fourth moment condition, and provide the necessary and sufficient conditions for all moments of the general GARCH process, as well as those of Ding *et al.'s* [1993] asymmetric power GARCH process.

tion of normality is used to define the likelihood function, but is not necessary for the asymptotic results[15].

For estimation purposes, if normality is assumed when the true conditional density is not normal, the resulting maximum likelihood estimates (MLE) should be interpreted as quasi-maximum likelihood estimates (QMLE). Weiss [1986] and Bollerslev and Wooldridge (1992) show that, even in the presence of non-normality, the resulting QMLE are asymptotically normally distributed and consistent if the second and fourth moment conditions are satisfied. For symmetric departures from conditional normality, the QMLE are generally close to the MLE. However, for non-symmetric distributions, both the finite and asymptotic loss of efficiency of using an incorrect conditional distribution may be very large (up to 84%) (see also Engle and Gonzalez-Rivera [1991] and Gonzalez-Rivera and Drost [1999]). This means that the precision of the parameter estimates may be very low. Although we may still be able to draw valid inferences, the power of the hypothesis test will be markedly reduced. Furthermore, the precision of the forecast will be reduced, resulting in wider forecast intervals. Ling and McAleer [1999c] show that efficient estimates for non-stationary ARMA models with GARCH errors can be constructed in the absence of knowledge of the conditional distribution through adaptive estimation.

Validation of the regularity conditions of GARCH(1,1) is of interest to forecasting for several reasons. First, only when the regularity conditions are satisfied

[15] Terasvirta [1996] derived the unconditional fourth moment of GARCH(1,1) without the normality assumption.

is it possible to provide forecasts with appropriate standard errors. Second, it makes it possible to examine how well an estimated GARCH(1,1) model reproduces the characteristics in the data. For example, if both the second and fourth moment conditions are satisfied, valid inferences are possible using the t-ratios as to the significance of the parameter estimates. Moreover, these estimates can be used to compute the values for the kurtosis and serial correlation of the unconditional distribution of $\{\varepsilon_t\}$ (Bollerslev [1986]). These computed values can then be compared with the measures calculated directly from the data to check if GARCH(1,1) reproduces the high kurtosis and persistent serial correlation in the squared or absolute values of $\{\varepsilon_t\}$. Moreover, validation of the fourth moment condition permits valid inferences using the asymptotic normal distribution.

Unfortunately, the statistical validity of GARCH(1,1) is generally not checked, despite the fact that the regularity conditions are straightforward. In practice, the t-ratio as usually interpreted as if the regularity conditions are satisfied. This should be of great concern since, in many empirical applications, the parameter estimates of the GARCH(1,1) process imply a violation of the second and/or fourth moment conditions (see Engle and Bollerslev [1986]; Engle and Gonzalez-Rivera [1991]; Bera and Higgins [1993]; Guillaume *et al*. [1994]; and Lumsdaine [1996]).

3 Outliers

Assuming that the time series follows a particular process, it is argued that returns consist of three types of observations: (i) ordinary observations; (ii) extreme observations; and (iii) outlying (extraordinary) observations. The first two types of observations are assumed to arise from a conditional normal (or heavy-tailed) distribution. Ordinary observations are realised the most frequently. Extreme observations arise from the tails of the distribution, occur relatively frequently, and are usually clustered[16]. For example, in analysing the occurrence of extreme values in the daily returns of the S&P 500 index from 3 July 1962 to 31 December 1997 (with a total of 8938 observations) for positive (negative) returns, Tsay [1999] identified 832 (765) clusters exceeding 1.0%, 145 (107) clusters exceeding 2.0%, and 64 (50) clusters exceeding 2.5%. The concept of outliers is generally rather vague, and most analyses of outliers usually lack even an informal definition. For the purpose of this study, outliers are observations that are far removed ($> 3\sigma$) from the mean of the conditional normal (or heavy-tailed) distribution generating the bulk of the data. Although outlying observations could come from a normal (or heavy-tailed) distribution, they should occur only rarely. For example, the normal distribution predicts that, on average, observations will fall outside the range $\pm 4\sigma$ less than 0.003% of the time, where σ denotes the standard deviation. However, in reality, such outlying observations are observed far too frequently to be consistent with normality in the data. For example, for the monthly S&P 500 index from 1926 to 1991 (giving 792 observations), Hotta and Tsay [1998] find three outliers (0.4%) in May 1940, November 1973 and October 1987.

In autoregressive (integrated) moving mean AR(I)MA models (Chang and Tiao [1983]; Burn and Ord [1984]; Hillmer [1984]; Chang *et al.* [1988]; Chuang and

[16] A cluster is defined as at least two consecutive trading days with (absolute) returns above the threshold level.

Abraham [1989]; Ledolter [1989]; Deutsch *et al.* [1990]; Balke [1993]; Hotta [1993]; Balke and Fomby [1991,1994]; Chen and Liu [1993b]; Trivez [1995]; and Thury and Wuger [1996]), it has been found that outliers can result in significantly negatively biased estimates of the AR(1) coefficient and positively biased estimates of the MA(1) coefficient[17], and may also result in model misspecification. Similarly, for the GARCH model, the presence of outliers may lead to problems in estimation, specification, forecasting and interpretation. For example, in the presence of outliers, the estimate of the α (or ARCH) effect in the GARCH(1,1) model can be severely biased upwards, and the estimate of the β (or GARCH) effect severely biased downwards (Friedman and Laibson [1989]; Lamoureux and Lastrapes [1990]; Blair *et al.* [1998]; Hotta and Tsay [1998]; Sakata and White [1998]; Engle and Lee [1999]; Franses and Ghijsels [1999]; and Franses and van Dijk [1999]). Outliers also adversely affect both the size and power of the standard Lagrange Multiplier (LM) test used to detect ARCH effects, leading to model misspecification (Franses *et al.* [1998]; and van Dijk *et al.* [1999]). In particular, a few outliers, either isolated or in patches, may result in spurious ARCH effects, when none, in fact, is present (Balke and Fomby [1994]; and Aggarwal *et al.* [1999]).

Outliers may affect forecasts temporarily through the carry-over effect on the ARCH and GARCH terms, and may have a permanent effect on forecasts through model misspecification and biased parameter estimates. In ARIMA models, the accuracy of the point forecasts immediately following an outlier has been shown to

[17] A 3σ outlier in 100 observations in an AR(1) process with a coefficient of 0.5 led to a negative bias of 0.034. For a 5σ outlier, the bias was -0.095. In an ARIMA(0,1,1) process with an MA(1) coefficient of 0.5, a 3σ outlier led to a positive bias of 0.028, while a 5σ outlier led to a positive bias of 0.066.

be severely impaired. However, point forecasts are significantly less affected by an outlier when such observations are more than two periods away from the forecast origin (Hillmer [1984]; Tsay [1988]; Ledolter [1989]; and Chen and Liu [1993b]). This result arises because the effects of past observations on subsequent forecasts diminish exponentially with the distance from the forecast origin[18]. Moreover, the impacts of outliers on the parameter estimates affects the forecasts substantially less. Outliers have also been found to increase the estimated variance of the observations and, consequently, the width of the forecast intervals (Hillmer [1984]; Chang *et al.* [1988]; Ledolter [1989]; and Trivez [1995]). There is also evidence that GARCH(1,1) overpredicts the persistence of moderate to high observations, and that this effect is considerably exacerbated when outliers are present (Lamoureux and Lastrapes [1990]; Engle and Mustafa [1992]; Hamilton and Susmel [1994], Sentana [1995], and Gray [1996]). For example, when the 1987 stock market crash is retained in the estimation period, the GARCH(1,1) volatility forecasts are overstated indefinitely into the future. Failing to account for the short-lived effects of outliers should, therefore, result in overestimation of subsequent volatility.

It is argued in this paper that outlying observations are not generated by an AR(1)-GARCH(1,1) process, which explains the majority of observations in the data set, and should be treated as contamination. Such contamination is demonstrated by de Haan *et al.* [1989] and Danielsson and de Vries [1997], who find that volatility clustering vanishes for outlying observations, that the GARCH

[18] In an ARIMA(0,1,1) process, when the outlier (of size 3σ) is positioned at the forecast origin, it causes a 441% increase in the mean square error (MSE), compared with a 40% increase in the MSE when the outlier is one step prior to the forecast origin, and a 3.6% increase in the MSE when the outlier is positioned two steps prior to the forecast origin.

effect is not particularly important, and that it suffices empirically to assume that the outlying observations are i.i.d.. Consequently, these outlying observations should not be used in estimating the AR(1)-GARCH(1,1) parameters, as this leads to biased estimates, model misspecification and inaccurate forecasts. It is our intention to fit the AR(1)-GARCH(1,1) model to the ordinary and extreme observations but not to the outlying events. The principal idea behind outlier adjustments is to apply the non-robust maximum likelihood procedure (MLE)[19] to the "cleansed" returns $\{y_t^*\}$, rather than to the original returns $\{y_t\}$, where cleansing refers to an outlier detection and down-weighting procedure (Kleiner *et al.* [1979])[20].

As the location of outliers is commonly judged to be obvious, the practice of visual inspection of the data, and deletion or adjustment of those observations that are arbitrarily judged to be outlying, has been commonly used for over 200 years (Bernoulli, 1777). However, the most obvious outliers do not necessarily have the largest impact on the parameter estimates (Sakata and White [1998]). Moreover, the greater is the complexity of the process (such as multivariate), the more difficult it is to identify outliers visually (Thury and Wuger [1996]). Conse-

[19]The MLE procedure under the assumption of normally distributed errors is highly sensitive to departures from normality, such as outliers and extreme observations (Sakata and White [1998]).

[20]Others have proposed robust methods to estimate the parameters, which avoids the problem of identifying outliers. For example, Schwert [1989a] and Taylor [1997] suggest a GARCH model that uses the lagged absolute residuals instead of the squared residuals, as this may be more robust to outliers. Friedman and Laibson [1989] suggest truncating the effect of large observations by using the sine of the scaled squared residuals and smooth pasting it to the constant 1 at $\frac{\pi}{2}$. Sakata and White [1998] suggest a new class of estimators that are robust to outliers. Although these methods are relatively straightforward, there are several disadvantages: (i) Chuang and Abraham [1989] demonstrate that robust methods perform well in some cases, but poorly in others; (ii) since the outlying observations are not adjusted, the impact of outliers on the forecasts remains; (iii) in most cases, only limited information on the outlier may be obtained (e.g. from the weights applied to the residuals).

quently, more objective and sophisticated outlier detection procedures are required to analyse those outliers that cannot be easily determined by visual inspection.

The estimated residuals of a volatility model usually provide some information for detecting outliers. Using this approach, outliers are defined with respect to an assumed econometric model. A common and simple approach for identifying a single outlier in a linear regression model is selecting the observation with the largest absolute studentised residual, or the observation with the largest impact on the residual sum of squares upon deletion (Beckman and Cook [1983]). These methods are not appropriate for detecting outliers in time series, since the observations are usually neither normally nor independently distributed.

Outlier analysis in time series is undertaken by specifying particular outlier models. Fox [1972] is the first to have proposed models for additive outliers (AO) and innovation outliers (IO), the two most important and frequently encountered outliers in time series[21]. An AO at time t affects only a single observation, is not persistent, and has no impact on volatility. It can arise because of a recording or measurement error, or as a result of an unusual event that does not carry over to subsequent observations. An IO at time t has an effect on several observations and also on the volatility of the observed data. For example, changes in daily U.S. interest rates have a profound long-term effect on the tail behaviour of the daily stock returns (Tsay [1999]). The 1987 stock market crash shows up in the S&P 500 stock return index as two outliers[22], an extreme negative outlier (20.4%) on

[21]Hotta and Tsay [1998] refer to AO and IO as level outliers (LO) and volatility outliers (VO), respectively.

[22]Similarly, the 1997 stock market crash shows up in the S&P 500 return as two outliers, namely a negative returns outlier of 6.9% on 27 October 1997, followed by a positive returns outlier of 5.1% on

19 October 1987, followed by an extreme positive outlier (5.3%) on 20 October 1987[23], and these have also been classified as IO (Hotta and Tsay [1998]) [24]. In a stationary ARMA process, AO have a decaying effect on the subsequent residuals and on the estimated underlying volatility process (see equation (3)). IO affect only the level of the standardised residual given in equation (2) and have no effect on the estimated underlying volatility process in equation (3) because the outlier is part of the innovation, with the impact carrying over to subsequent observations (Chang *et al.* [1988]; Tsay [1988]; Chen and Liu [1993a]; and Hotta and Tsay [1998]).

Based on these outlier types, Fox [1972] proposed a likelihood ratio test for their detection in autoregressive models at an unknown point in time. Abraham and Box [1979] proposed some simple but realistic dynamic intervention models, such as that of Box and Tiao [1975], to accommodate outliers within the model as interventions in the deterministic part of the process. In this way, it is possible to separate the outlier effects from the stochastic component to be explained, without having to model the stochastic behaviour of the outliers. Based on these ideas, procedures have been developed which: (i) estimate an ARMA model; (ii) iterate sequentially among outlier detection, adjustment of the residual series, and re-estimation of the variance of the residuals; and (iii) estimate a dynamic intervention model which includes intervention variables to remove the effects of

28 October 1997.

[23]The second outlier in the returns is a direct result of the extreme negative single outlier that occurred on 19 October 1987.

[24]Hotta and Tsay [1998] also found that the 1987 had a larger impact on negative returns than on positive returns. The annual average size of negative returns over 1.5% remained high after the 1987 crash, but had little effect on positive returns.

the detected outliers (Kleiner *et al.* [1979]; Chang [1982]; Chang and Tiao [1983]; Hillmer *et al.* [1983]; Tsay [1986, 1988]; Chang *et al.* [1988]; Chen and Tiao [1990]; Balke [1993]; and Justel *et al.* [1998]. This approach has several drawbacks, one of which is that the presence of outliers may lead to incorrectly identified and inefficiently estimated ARMA parameters, making this process vulnerable to swamping[25] and masking effects[26] during outlier detection.

Chen and Liu [1993a] developed a procedure which yields the joint, rather than sequential, detection of multiple outliers and estimation of the parameters, and appears more robust against spurious outliers and masking effects. This is an iterative method, which consists of specification-estimation-detection-removal cycles to accommodate individually the most significant outliers, as follows: (i) an ARMA model is specified and estimated; (ii) the impact of the outliers on the residuals of the model is investigated; (iii) the observed series are adjusted according to the impact of the largest outlier, and the model is re-estimated. The process is repeated until no more outliers are detected[27].

The general model for an observed time series $\{y_t\}$ which is contaminated with outliers can be expressed as follows:

$$y_t = y_t^* + f(t) \tag{11}$$

where $\{y_t^*\}$ is the unobserved uncontaminated series, which is assumed to follow

[25]Swamping effects occur if ordinary observations are declared as outliers.

[26]Masking effects occur if arbitrary large observations are not detected as outliers when one outlier obscures the presence of another.

[27]Since the variance of the residuals itself is affected by outliers, moving between iterations of adjusting only the largest outlier and re-estimating the model, and hence the variance of the residuals, reduces the masking effect.

an ARMA process, namely

$$y_t^* = \frac{\theta(L)}{\phi(L)}\varepsilon_t^* \tag{12}$$

$\theta(L) = 1-\theta_1 L - .. - \theta_l L^l$ is the moving mean operator, $\phi(L) = 1-\phi_1 L - .. - \phi_k L^k$ is the autoregressive operator, L is the lag operator, $L^k y_t = y_{t-k}$, and $\varepsilon_t^* \sim N(0, h_t)$.

In equation (11), $f(t)$ is a parametric outlier model, in the class of intervention models of Box and Tiao [1975], representing the exogenous disturbances of y_t^* which may affect both the model parameters and the variance of ε_t^*, h_t. The general form of the outlier model is

$$f(t) = \xi_0 \frac{\xi(L)}{\delta(L)} I_{td} \tag{13}$$

where ξ_0 is the initial impact of the outlier (intervention parameter), namely $E(y_t) = E(y_t^*) + \xi_0$, $\frac{\xi(L)}{\delta(L)}$ is the dynamic pattern of the outlier, $\xi(L) = 1 - \xi_1 L - .. - \xi_l L^l$, $\delta(L) = 1 - \delta_1 L - .. - \delta_k L^k$, and I_{td} is an indicator function (that is, $I_{td} = 1$ if t = d and $I_{td} = 0$ if $t \neq d$) signifying the occurrence of an outlier at time d. The specific outlier models are as follows:

(i) for AO, $\frac{\xi(L)}{\delta(L)} = 1$ and

$$f(t) = \xi_0 I_{td} \tag{14}$$

while (ii) for IO, $\frac{\xi(L)}{\delta(L)} = \frac{\phi(L)}{\theta(L)} = \pi(L)$ and

$$f(t) = \pi(L)\xi_0 I_{td}. \tag{15}$$

The estimated contaminated residuals can then be expressed as:

$$\hat{\varepsilon}_t = \varepsilon_t^* + \xi_0 x_t \tag{16}$$

where

$x_t = 0$ for $t < d$

$x_t = 1$ for $t = d$

$x_{t+k} = \pi_k$ for AO (or 0 for IO) for $t > d$ and $k = 1, 2, ...$

It is easy to see that, although the AO have an immediate effect on the observed series, they have a dynamic effect on ε_t for every $t > d$, depending on the correlation structure, $\pi(L)$, of the model, and consequently on the underlying volatility process (equation (3)). In an ARMA model, the IO affect all observations, but affect only one residual at $t = d$.

Since the GARCH(1,1) model can also be expressed as an ARMA(1,1) process, the Chen and Liu [1993a] procedure can be applied to $\{\varepsilon_t^2\}$, where $\{\nu_t\}$ is the residual series in

$$\varepsilon_t^{*2} = \omega + (\alpha + \beta)\varepsilon_{t-1}^{*2} + \nu_t^* - \beta\nu_{t-1}^*. \tag{17}$$

Therefore,

$$\nu_t^* = \psi(L)\varepsilon_t^{*2} - \frac{\omega}{(1 - \beta L)} \tag{18}$$

where

$$\psi(L) = \frac{1 - (\alpha + \beta)L}{(1 - \beta L)}. \tag{19}$$

For AO, it follows that

$$\varepsilon_t^{*2} = \varepsilon_t^2 - \xi_0 I_{td} \tag{20}$$

in which case

$$\nu_t^* = \psi(L)(\varepsilon_t^2 - \xi_0 I_{td}) - \frac{\omega}{(1-\beta L)} = \nu_t - \psi(L)\xi_0 I_{td} = \nu_t - \xi_0 x_t. \tag{21}$$

Equation (21) can be interpreted as a regression model for ν_t^*.

In practice, ξ_0, I_{td}, and the type of outlier are unknown, and hence must be estimated. If the parameters of the ARMA-GARCH model are known, the least squares estimate for the initial impact, $\hat{\xi}_0$, of a single outlier on the estimated residuals, $\hat{\nu}_t$, at t = d, may be expressed as:

$$\hat{\xi}_0 = \frac{\sum_{t=d}^{N} \hat{\nu}_t x_t}{\sum_{t=d}^{N} x_t^2} \tag{22}$$

and

$$var(\hat{\xi}_0) = \frac{\hat{h}_t}{\sum_{t=d}^{N} x_t^2} \tag{23}$$

where N is the number of observations. For each iteration, the maximum standardised test statistic for $\hat{\xi}_{0t}$, for all observations $\hat{\tau}_{0t}$ is estimated to detect empirically the outliers with the largest initial impact (Chang *et al.* [1988]):

$$\hat{\tau}_0 = \frac{\hat{\xi}_0}{\hat{h}_t}\sqrt{\sum_{t=d}^{N} x_t^2} \tag{24}$$

where $\hat{h}_t$, the estimated standard deviation of the residuals, is robust to outliers[28].

[28]Since the standard deviation itself is affected by outliers, this is how outliers can be masked.

Furthermore, all standardised test statistics are assumed to be approximately normally distributed with mean zero and unit variance. The maximum across all types of outliers is examined and compared with a pre-specified critical value.

The procedure is given as follows: (i) estimate the ARMA model for $\{y_t\}$ to obtain $\{\hat{\varepsilon}_t\}$; (ii) estimate the GARCH(1,1) model for $\{\hat{\varepsilon}_t\}$ to obtain $\{\hat{h}_t\}$ and $\{\hat{\nu}_t\} = \{\hat{\varepsilon}_t^2 - \hat{h}_t\}$; (iii) obtain the least squares estimates $\hat{\xi}_{0t}$ for all observations and compute the standardised test statistic. The observations with the maximum test statistic exceeding the critical value are then adjusted, giving the estimated uncontaminated residuals:

$$\hat{\varepsilon}_t^* \equiv sign(\hat{\varepsilon}_t)\sqrt{\hat{\varepsilon}_t^2 - \hat{\xi}_0}. \tag{25}$$

The AR(1)-GARCH(1,1) model is then re-estimated using the adjusted observed returns and the procedure above is repeated. This cycle is continued until the test statistics for all observations lie below the critical value [29].

Chen and Liu [1993a,b] observe that the estimated parameters can also be affected by the choice of the critical value. When the critical value is too large, the method may not identify any outliers. Although lower critical values result in an increased power of outlier detection, the chances of detecting spurious outliers are also increased, which may result in overadjustment of spurious outliers and biased parameter estimates. Furthermore, AO are the easiest to detect, while outliers that have the same shape as the basic data set, such as level shifts[30], are

[29] The distinction between outliers and extreme observations depends on this critical value.

[30] With level shifts, the data are drawn from the same normal distribution but with a mean shift.

the hardest to detect. Moreover, the closer is the distance of outliers from the basic data, the harder they are to detect.

4 Empirical Results

4.1 Data

In this section we use the AR(1)-GARCH(1,1) model to analyse unadjusted returns $\{y_t\}$ and outlier-adjusted returns $\{y_t^*\}$, as this is the most commonly used parametric model for examining time-varying volatility[31]. The inclusion of the AR(1) term is based on the observed (weak) positive first-order serial correlation[32] in stock index returns which is attributed, in part, to non-synchronous trading and bid-ask spreads (Safvenblad [1997]).

The effects of outliers on the AR(1)-GARCH(1,1) model are evaluated using 1000 trading days of five financial time series, comprising stock, currency and commodity data, namely the daily close-to-close logarithmic index returns of S&P 500, Nikkei 225 and HSI, the noon (Pacific time) British Pound-U.S. Dollar (GBP/USD) spot exchange rate, and the closing (London) Gold Bullion (GB) spot rate. Datastream provided the close-to-close index returns and the noon GB spot rate, while the GBP/USD spot exchange rate was obtained from the Pacific Exchange Rate Service. For each series, highly volatile time periods containing a large number of outliers and extreme observations were selected, namely 22 March 1995 to 1 June 1999 for S&P 500, 10 July 1995 to 26 july 1999 for HSI, 3 August

[31]Misspecification of the conditional mean equation appears to have very little influence on the estimated conditional variance in continuous (Nelson [1990a,b]; Gannon [1996] as well as discrete time (McKenzie [1997]).

[32]Koutmos [1998] finds that stock returns are asymmetrically related to past returns and that the positive first-order serial correlation is mostly due to the persistence of positive returns.

1984 to 28 June 1988 for Nikkei 225, 1 June 1988 to 13 May 1992 for GBP/USD, and 5 April 1979 to 7 June 1983 for GB (see Figures 1-5).

Mean values of the parameter estimates, moment conditions and forecast errors were calculated using 500 one-day ahead volatility forecasts. The first 500 trading days were used to estimate the model, which yielded the one-day ahead forecasts for h_t. Then the estimation time interval was moved one-day ahead into the future by deleting the first trading day and adding an extra day at the end of the sample period. The parameters of the model were re-estimated and the one-day ahead forecasts re-generated. This procedure was repeated 500 times. For the outlier test statistic, we used critical values of 10, 8, 6 and 4.

4.2 Evaluation of forecasting ability

Tests as to the predictive content of volatility models are based on examining the differences between the out-of-sample volatility forecasts and the subsequently realised volatility. Unfortunately, volatility cannot be observed directly, so there is the added complexity of finding an appropriate definition of conditional volatility against which the GARCH(1,1) forecasts can be compared. In this paper, the following definition for realised volatility is used:

$$\sigma_t = \mid y_t - \bar{y} \mid \tag{26}$$

where the daily logarithmic returns are defined as $y_t = ln(\frac{P_t}{P_{t-1}})$, $\bar{y}$ is the conditional sample mean[33] of y_t given the previous values $y_{t-k}, k \geq 1$, and P_t denotes

[33]No significant differences in the volatility values are observed if the conditional sample median is used, or if the mean is taken to be zero.

price in period t[34].

The following descriptive statistics are used to quantify the mean of the out-of-sample forecast errors: mean error (ME), mean squared error (MSE), root mean squared error (RMSE), mean absolute error (MAE), and smoothed mean absolute percentage error (SMAPE). A weighted version of the SMAPE measure (SMWAPE) that weights larger values of volatility more heavily than smaller values is also calculated, as are the median versions of the descriptive statistics. These forecast statistics are a simple method for comparing the influence of outliers on GARCH(1,1) volatility forecasts.

4.3 The effects of outliers on the unconditional returns distributions

Table 1 reports the descriptive statistics for each of the unconditional distributions of the five time series for a total of 1000 observations. The Lagrange multiplier statistics for normality (LM(N)) indicate that none of the unadjusted time series is normally distributed. While the skewness of all the returns distributions is relatively small and negative (except for HSI), the kurtosis is large, implying that much of the departure from normality is due to leptokurtosis. GB spot rates are the most volatile series, having the largest standard deviation, followed by HSI, Nikkei 225, S&P 500 and GBP/USD.

Table 2 summarises the statistics for the detected outliers and extreme observations. Consistent with the evidence of leptokurtosis, outliers and extreme

[34] Andersen and Bollerslev [1998] acknowledge that, while absolute (or squared) daily returns provide unbiased estimates of the underlying unobservable volatility, they are very noisy estimators of daily movements in volatility due to the large idiosyncratic error term.

observations appear in all five time series. Outliers ($> 3\sigma$) account for between 1.2% - 1.6% of the observations, while extreme observations (of between 2σ and 3σ) account for twice the number of outliers observed. The largest numbers of outliers are observed for the Nikkei 225 series, with 5 observations larger than 5σ, 8 observation larger than 4σ, 12 observations larger than 3σ, 33 observations larger than 2σ, and 141 observations larger than 1σ, followed by HSI (4, 6, 16, 56 and 182, respectively), S&P 500 (2, 4, 13, 45 and 230, respectively), GB (2, 6, 12, 46 and 233, respectively), and GBP/USD (0, 2, 12, 52 and 266, respectively). Each of these numbers is significantly greater than expected from a normal distribution. The Nikkei 225 series also has the highest kurtosis measure (51.06), and the largest (relative) positive (9.05σ) and negative (-13.55σ) outliers, which correspond to the 19 October 1987 stock market crash. Although a single negative outlier may result in a significant loss in returns, a sequence of smaller outliers and/or extreme observations often leads to an even greater loss. For example, although the largest negative outlier for HSI or GB is between 15-18%, in each of these series there are at least two clusters where the total decrease in returns over a three-day period is between 15-21%.

Negative observations ($> 1\sigma$) occur up to 20% more frequently than positive observations, giving rise to the observed negative skewness in the returns distribution. Although both positive and negative outliers are observed, negative outliers also occur more frequently, which is consistent with previous findings (see, for example, Jansen and de Vries [1991]). Moreover, in absolute value, the largest outlier is usually negative, which is often subsequently followed by a large (but smaller absolute) positive outlier[35]. Similar findings have been reported by others (see, for example, Longin [1996]). Of the 35 outliers detected across the five series with critical value 10, 21 (60%) are negative. Of these, five (24%) are followed the next day by a positive outlier and only two negative outliers (5%) occur consecutively. Of the 14 positive outliers, only one (7%) is subsequently followed by a negative outlier and no positive outliers occur consecutively. The probability of a negative outlier being followed by another negative outlier is also greater than for a positive outlier to be followed by another positive outlier. For example, for a critical value of 8, of the 42 negative outliers detected, 8 (19%) occur consecutively, compared with only two (7%) of the 28 positive outliers. In total, 11 negative outliers are followed by another outlier of either sign, compared with only three for positive outliers. This is in agreement with the well-documented asymmetric sign effect in financial data, where large negative shocks lead to greater volatility than positive shocks of equivalent magnitude.

Outlying observations are frequently clustered and do not appear to be i.i.d.. For example, of the 70 outliers detected for a critical value of 8, there are eight clusters of two or more containing 22 (31%) observations, while 33 (47%) of the observations occur within two days of each other. Of the remaining outliers, there are usually several weeks between subsequent observations. The clustering increases for extreme observations. Of the 315 outliers and extreme observations detected for a critical value of 4, there are 41 clusters of two or more containing a total of

[35] On 29 October 1929, 19 October 1987, and 27 October 1997, the sharp fall in returns of both S&P 500 and DJIA rebounded on the following day.

136 (43%) observations. As in the findings of Blanchard and Watson [1986], Balke and Fomby [1994], Danielsson and de Vries [1997] and Tsay [1999]), among others, outliers and extreme observations are often found to occur together, implying a period of increased volatility. For example, for a critical value of 8, more than 55 (78%) of the outliers detected are clustered with extreme observations, which implies that outliers and extreme observations are not independently distributed.

Outliers also coincide with the turning points of the business cycle for HSI and GB, but not for the other three series. Schwert [1989b] also finds that there is a negative correlation between conditional volatility of the nominal asset returns and the business cycle.

The Chen and Liu [1993a] method of detecting the most significant individual outliers seems to work well with financial data. For a critical value of 10, outlying observations greater than 4σ are accommodated. Although the distinction between outliers and extreme observations may be arbitrary, the Chen and Liu [1993a] method begins to accommodate extreme observations (of between 2σ and 3σ) somewhere between a critical value of 6 and 8. The mean number of observations which are transformed (or down-weighted) each estimation is approximately doubled with each two-step decrease in the critical value. As in the findings of others, such as Chen and Liu [1993a] and Thury and Wuger [1996], it is found that an increased number of spurious extreme observations is detected while actual extreme observations are missed for a critical value of 4. This result may be due to the high correlation between neigbouring test statistics, as well as an increased risk of type I error. Moreover, the larger is the outlier, the greater is the estimated impact parameter, $\hat{\xi}_0$, and thus the greater is the adjustment to the observation. For example, large outliers are down-weighted, on average, by 70%, compared with 35% for extreme observations.

For all series, removing outlying observations reduces the standard deviation of the unconditional distribution of the returns by up to 40%, and also reduces the kurtosis measure, but the distribution remains non-normal. Similarly, S&P 500, Nikkei 225, HSI and GB spot rates all become more symmetric, suggesting that the outliers are primarily responsible for the asymmetries in these series. In contrast, the negative skewness observed for the GBP/USD spot exchange rate does not approach zero, implying that the skewness in this series is caused by ordinary returns. When all of the outliers and most of the extreme observations are down-weighted, and all return observations are below $2\sigma - 2.5\sigma$, the returns distribution becomes platykurtic.

4.4 The effects of outliers on the parameter estimates

For the five time series, table 3 reports the mean values of the parameter estimates of the AR(1)-GARCH(1,1) model, their standard deviations and their mean t-ratios (using the standard errors of Bollerslev and Wooldridge [1992], which are designed to be robust to non-normality, especially the presence of outliers). Diagnostic checks for the normality and i.i.d. assumptions are also reported.

Before examining the parameter estimates of the AR(1)-GARCH(1,1) model, it is necessary to check whether the model is an adequate description of the data. A minimum requirement that the empirical model is satisfactory is that the estimated

η_t are i.i.d.. The Ljung-Box statistics do not indicate any violations of the i.i.d. assumption for η_t. For both the adjusted and unadjusted series, none of the Ljung Box statistics for serial correlation up to lag 12 is significant at the 5% level, which indicates that there is no remaining serial correlations in the first and second moments of η_t. Generally, when outliers are removed, the Ljung-Box test statistics become less significant, implying that a few equi-distant outliers may lead to spurious autocorrelation. Second, as indicated by the Jarque-Bera LM(N) statistics, the assumption of conditional normality of η_t with mean zero and unit variance is not rejected for the unadjusted series, whereas conditional normality is rejected for the unadjusted series. The non-normality is due to both excess kurtosis and negative skewness. Moreover, although the standard deviation is close to unity, the mean is always negative. Examination of the standardised residuals reveals that the large (negative) outliers remain.

The empirical finding that GARCH(1,1) with normalised residuals reduces the leptokurtosis but does not accommodate all the outlying or extreme observations is very common (see, for example, Engle and Gonzalez-Rivera [1991] and Terasvirta [1996]). As the true conditional distribution may not be normal, Bollerslev and Wooldridge's [1992] robust t-statistics can be used, provided that the fourth moment condition is satisfied (so that the QMLE will be asymptotically normally distributed). The drawback is a loss of efficiency in parameter estimation and in point forecasting. Adjusting outliers results in a better approximation to normality of the standardised residuals, thereby reducing the excess kurtosis and skewness of $\{\eta_t\}$, and yielding a greater efficiency for the QMLE. Moreover, when the ratio of excess kurtosis to the standard deviation of the unconditional returns distribution reaches 40, the assumption of normality of η_t is essentially satisfied. This result implies that the amount of excess kurtosis induced by GARCH(1,1) under the normality assumption increases with the standard deviation of the series, and usually occurs when the critical value is less than 8[36]. Although alternative symmetric distributions for η_t may induce greater excess kurtosis (such as the t-distribution (Bollerslev [1987)] and GED (Hsieh [1989]; Nelson [1991]; Koutmos [1998]), these distributions cannot accommodate the negative skewness. Such negative skewness in the returns distribution is partly accommodated by asymmetric GARCH models (Franses and van Dijk [1996]) such as EGARCH (Nelson [1989]), GJR-GARCH (Glosten *et al.* [1992]), or QGARCH (Sentana [1991]).

For HSI, Nikkei 225 and GBP/USD, the daily returns are positively correlated, while the daily returns are negatively correlated for GB[37]. Outliers generally yield larger mean returns, μ, larger mean volatility, ω, and larger and more significant first-order serial correlation in the returns, φ. For example, when outliers are down-weighted, the mean φ estimate for HSI and Nikkei 225 fall by up to 40% and become insignificant. Extremely large (positive or negative) outliers result in substantially negatively biased φ estimates, leading to spurious AR(1) effects. These findings are consistent with those of Ledolter [1989].

For both the adjusted and unadjusted data, the ARCH and GARCH estimates

[36]The excess kurtosis calculated using the α and β estimates is highly sensitive to the presence of outliers, but approaches the kurtosis measure calculated directly from the data when outliers are removed.

[37]Sentana and Wadhwani [1992] demonstrate that when volatility is high (low), stock returns at short horizons exhibit negative (positive) serial correlation.

are significantly different from zero, with the β estimates having a substantially higher statistical significance. The mean α estimates are also much smaller than the mean β estimates, indicating that, on average, there is a relatively weak reaction of conditional volatility to innovations (ARCH effects), but with a long-term memory (GARCH effects).

The results show that the φ and α estimates, in particular, and to a lesser extent, the β estimates, vary considerably over time. Interestingly, the α estimates are substantially larger (smaller) when volatility is high (low), while the β estimates are substantially smaller (larger), implying that larger innovations have larger ARCH (short-run) effects but smaller GARCH (long-run) effects. For example, when outliers are down-weighted for HSI, the mean ($\frac{1}{1-\beta}$) and median ($\frac{log2}{log\beta}$) lags of the GARCH process more than double from 6.0 to 12.3 days and from 3.8 to 8.2 days, respectively. Such evidence that large shocks have only short-lived effects and are considerably less persistent than small shocks has also been reported by Poterba and Summers [1986], Friedman and Laibson [1989], Stein, [1989], Schwert [1990], Engle and Mustafa [1992], Hamilton and Susmel [1994], Gray [1996], Klaassen [1998], Sakata and White [1998], Susmel [1998], Engle and Lee [1999], Franses and Ghijsels [1999], and Schwert [2000]. This result implies that large shocks have only a temporary effect, so that financial time series are rather robust to outliers and revert to their long-term trends relatively quickly.

As a result of the second moment condition for covariance stationarity and the resulting trade-off between the α and β coefficients, the standard GARCH(1,1) model is unable to capture both outlying (extreme) observations and low volatility persistence (Hamilton and Susmel [1994]; Kim and Kim [1996]; Mandelbrot *et al.* [1997]; Susmel [1998]). Consequently, when the same specifications are used, the interpretion of the persistence measure of the GARCH(1,1) model is unlikely to be valid. Failing to accommodate for the short-lived effects of outlying and extreme observations should lead to an overestimation of subsequent volatility.

To the extent that both outlying and low volatility innovations have different memory effects and cannot be modelled using the standard linear GARCH specification, non-linear GARCH models have been developed. For example, Cai [1994], Hamilton and Susmel [1994], Gray [1996], Kim and Kim [1996], Susmel [1998], and Bollen *et al.* [2000], using variations of the standard Markov-Switching (MS) model, in which the volatility process switches from one (G)ARCH process to another, depending on the probability of being in a particular regime (or state of the stock market), namely a low-volatility, moderate-volatility, or a high- (or unusual) volatility regime, show that such models provide much better forecasts than a GARCH(1,1) model. Such improvement is due mainly to its ability to model the different degrees of persistence of low-, moderate- and high-volatility regimes, and does not attribute a large degree of persistence to the effects of extreme and outlying observations. In particular, the outlying observations tend to be collected in the unusual volatility state, which plays the role of an intervention dummy. Friedman and Laibson [1989] and Terasvirta and Anderson [1992] show that a threshold ARCH model may be more appropriate in modelling the subdued volatility as a result of outlying observations, where the process is governed by a series of distinct ARCH processes and the ARCH process generating the volatility depends upon the value taken by a 'threshold'. Others have introduced jump processes or other independent random shocks directly to the variance, which allows the modelling of sharp increases in volatility, without further resticting the allowable range of β (Mandelbrot *et al.* [1997]).

The MLE is dominated by outlying observations, with the short-term memory of the outlying observations overwhelming the long-term memory of the smaller observations. For example, removing the seven largest outliers in the second least volatile series, S&P 500, which results in a decrease in the size of the maximum absolute outlier from 6.85σ to 4.35σ, causes the mean α estimate to decrease by 0.064 (from 0.123 to 0.059), while the mean β estimate increases by 0.085 (from 0.834 to 0.919). The most significant changes in the parameter estimates are observed for Nikkei 225, which has the largest outliers. For this series, when the largest eight outliers are down-weighted and the maximum absolute outlier decreases from 13.55σ to 5.10σ, the mean α estimate decreases by 0.279 (from 0.435 to 0.156) while the mean β estimate increases by 0.398 (from 0.377 to 0.775).

Setting the critical value too low has adverse effects on the parameter estimates in that the α estimates converge to zero, while β estimates converge to unity, with the volatility process resembling a random walk. The reason for this outcome is that the volatility clustering is removed, with the data more closely resembling a white noise process.

For the individual time series, outliers cause level shifts in the parameter estimates. For example, a sharp increase in the α estimate and a simultaneous and almost equivalent sharp decrease in the β estimate follows when an outlier enters the estimation period. This effect is sustained while the outlier remains in the estimation period, and its location in the estimation period appears to have little effect on its magnitude[38]. When the outlier is eventually omitted from the estimation period, the effect on the parameter estimates is reversed. The larger is the size of the outlier (measured in σ), the larger is its effect on the parameter estimates, which is particularly evident for Nikkei 225, which includes the 1987 stock market crash. When the extremely large negative outlier ($> 13\sigma$) of 19 October 1987 enters the estimation period, this causes a sharp and dramatic increase (0.571) in the α estimate and an almost equivalent sharp decrease (0.582) in the β estimate. Moreover, when the extremely large positive outlier ($> 9\sigma$) of 20 October 1987 enters the estimation period, no significant additional effect on the parameter estimates is observed. This illustrates the fact that the occurrence of a single extreme outlier can yield spurious ARCH effects, which is similar to the findings of, for example, Lastrapes [1989], Lamoureux and Lastrapes [1990], Balke and Fomby [1994], van Dijk *et al.* [1999], and Aggarwal *et al.* [1999]).

Removing outliers results in a significant increased stability of the β estimates over time, but a decreased stability of the φ and α estimates. In addition to their dramatic negative effect on the β estimates, outliers have a similar effect on their t-ratios. Interestingly, although outliers usually have a dramatic positive effect on the α estimates, no such effect is observed on their t-ratios. For example, when the largest outlier (6.9σ) of the S&P 500 series enters the estimation period, this

[38]This must not be confused with the location of the outlier relative to the observations in its immediate surroundings which appears to be a significant determinant of the effect of the direction and magnitude of the outlier on the parameter estimates.

causes a sharp and dramatic decrease (from 23.58 to 9.82) in the t-ratio of the β estimate while the t-ratio of the α estimate decreases only slightly (from 1.05 to 1.01). The effects on the t-ratios of the β estimates are sustained while the outlier remains in the estimation sample.

4.5 The effects of outliers on the moment conditions

Table 4 reports the estimated values of $(\alpha + \beta)$ and $(3\alpha^2 + 2\alpha\beta + \beta^2)$, their standard deviations in parentheses, and the frequency of violation of the second- and fourth-order moment conditions.

The mean estimated value of $(\alpha + \beta)$ is usually very close to, but less than, unity, implying that the volatility process is highly persistent and close to a unit root. On average, outliers cause a decrease in the persistence measure, $(\alpha + \beta)$, of the volatility process, which is similar to the results of Friedman and Laibson [1989], Blair *et al.* [1998], Sakata and White [1998], and Engle and Lee [1999]. Note that $(\alpha+\beta)$ is not very sensitive to outliers because these observations usually have opposite but equivalent effects on the estimates of the GARCH(1,1) parameters.

Only for the two series with the largest outliers, namely Nikkei 225 and HSI, is the second moment condition violated, although the largest estimated value is only 1.02. When extreme outliers are down-weighted, the second moment condition is satisfied. However, for HSI, adjusting too many of the outliers and extreme observations eventually causes an increase in the second moment condition. Violation of the second moment condition might appear to suggest non-stationarity in the GARCH(1,1) process. However, if some weaker requirements (such as the log-moment condition) are met, the volatility process can still be stationary, even though $(\alpha+\beta)$ might exceed unity (Nelson [1990b]; Lee and Hansen [1994]; Lumsdaine [1995]).

The mean $(3\alpha^2+2\alpha\beta+\beta^2)$ estimates of the fourth moment condition are also close to unity, but are usually slightly smaller than the mean $(\alpha+\beta)$ estimates. Outliers have a dramatic effect on the $(3\alpha^2+2\alpha\beta+\beta^2)$ estimates. For example, when the extreme outlier of 19 October 1987 enters the estimation period for Nikkei 225, this causes a doubling in its value (from 1.023 to 2.07). It is clear that the fourth moment regularity condition is more stringent than the second moment condition when outliers are present. For example, for Nikkei 225 and HSI, the fourth moment condition is violated more than half the time. Removing outliers usually results in an increased stability of the $(3\alpha^2+2\alpha\beta+\beta^2)$ estimates over time, and a lower frequency of violation of the fourth moment condition.

Violation of the fourth moment regularity condition means that the assumption of asymptotic normality may not hold, so that inferences based on normality may not be valid. For example, the unconditional kurtosis measure for the return series may be meaningless. In addition, the confidence intervals for forecasts may not be meaningful.

4.6 The effects of outliers on the volatility forecasts

Table 5 reports the forecast errors of the AR(1)-GARCH(1,1) model for the five time series calculated for the entire sample, while Table 6 reports the forecast errors calculated separately for low ($< 1\sigma$) and high ($\geq 1\sigma$) volatilities.

In accordance with the previous findings, based on MAE, RMSE or RMedSE measures calculated over the entire sample, there is a substantial reduction in the mean forecast errors (of up to 17%) when outliers (larger than $\pm 4\sigma$) are down-weighted. When the sample is split into low and high volatility periods, reduced mean forecast errors are observed only for low volatility, whereas the mean forecast errors are substantially larger for high volatility. For example, removing the 9 largest outliers from HSI reduces RMSE for low volatility by 15%, but increases RMSE for high volatility by 11%. Based on a relative error measure such as SMAPE, the mean forecast error is reduced by about 3% for low volatility periods but increased by about 10% for high volatility periods. The deterioration in the forecast accuracy for high volatility periods can be explained by the fact that outlying observations are frequently clustered with other large observations, so that removing an outlier will reduce the forecast accuracy of the subsequent large observations.

When the RMSE measure is calculated separately for positive and negative forecast errors, removing outliers reduces RMSE for positive forecast errors but increases it for negative forecast errors. This is observed for both low and high volatility periods.

For the whole sample, the empirical results also show that the GARCH(1,1) forecasts are strongly positively biased and, on average, overpredict volatility 70% of the time. Moreover, the forecast errors are highly negatively skewed and leptokurtic. When the sample is split into low and high volatility periods, it is evident that GARCH(1,1) markedly overpredicts low volatility periods (80% of the time) but underpredicts high volatility periods (90% of the time)[39] Hence, consistent with the results in section 4.4, the AR(1)-GARCH(1,1) process is unable to model extreme tail behaviour.

When outliers are down-weighted, for low volatility periods this results in a dramatic decrease in the rate of overprediction, negative skewness, and excess kurtosis of the distribution of the forecast errors. For periods of high volatility clustering, this results in a dramatic increase in the rate of underprediction, negative skewness, and excess kurtosis of the distribution of the forecast errors. The higher rate of overprediction when outliers are present, for low volatility periods, arises partly through the higher α estimates. A larger proportion of the innovations is allowed to have an impact on the volatility process, thereby making the subsequent forecasts for low volatility periods inappropriately high. This is consistent with the evidence that GARCH(1,1) overforecasts the persistence of high volatility (see, for example, Engle and Mustafa [1992]; Lamoureux and Lastrapes [1993]; Hamilton and Susmel [1994]; Sentana [1995]; Gray [1996]). Engle and Mustafa [1992] and Sakata and White [1998] show that when the 1987 stock market crash is included in the estimation period, the GARCH(1,1) forecasts are substantially larger in its level and more variable.

The Pesaran and Timmermann test statistic (PTTEST), which computes a non-parametric association between the forecasted and realised volatility, implies that there is a strong association between the GARCH(1,1) forecasted volatil-

[39] The GARCH(1,1) volatility forecast fits remarkably well the actual volatility when this is based on the logarithm of the days high to the days low price. This logarithmic range measure is much higher and less noisy than the volatility measure based on returns.

ity and the realised volatility. However, this is only observed for periods of low volatility.

A simple gauge of the predictive power of GARCH(1,1) is the R^2 value which is obtained by regressing the ex-post volatility on the forecast volatility. The empirical results show that the adjusted R^2 values are relatively low (never exceeding 18%) for all five series[40]. These regressions are based on the observed absolute returns as a measure of the realised volatility. Anderson and Bollerslev [1998] attribute the poor fit to the inherent noise in realised volatility[41]. When outliers are down-weighted, the out-of-sample explanatory power of GARCH(1,1) decreases substantially for all series.

Overall, the results suggest that when outliers are down weighted, the AR(1)-GARCH(1,1) volatility forecasts are improved substantially for periods of low volatility clustering, but not for periods of high volatility clustering.

5 Conclusion

Using several financial time series, we have shown that outliers and extreme observations significantly affect the parameter estimates of the AR(1)-GARCH(1,1) model. Outliers dominate the maximum likelihood estimates, causing a significant increase in the AR(1) and ARCH parameter estimates and a significant decrease in the GARCH parameter estimates. Furthermore, the presence of extreme outliers

[40] Care must be taken with using R^2 as a measure of the goodness of fit as the volatility series usually display strong trends.

[41] Anderson and Bollerslev [1998] show that when the volatility measure is based on high-frequency intra-day (5-minute) returns rather than on daily returns, there is a significant reduction in the noise component. This leads to a substantial improvement in the intertemporal stability of volatility and a marked increase in the predictive power of GARCH.

may give rise to spurious AR(1) and ARCH effects when, in fact, none is present.

For all the data sets examined, the independent and identically distributed assumption of the standardised residuals cannot be rejected at the 5% significance level. In contrast, the assumption of normality of the standardised residuals was rejected for all the time series. The maximum excess kurtosis that is captured by GARCH(1,1) under the assumption of conditional normality increases with the standard deviation (which is approximately 40). Adjusting outliers makes the distribution of the standardised residuals more normal, thereby implying greater efficiency of the QMLE of the parameters.

It is also found that the regularity conditions, in particular the fourth moment condition, are harder to satisfy when outliers or extreme observations are present. Rejection of the fourth moment condition means that the conditional innovations may not be asymptotically normally distributed, and that the t-ratios of the parameter estimates may not be asymptotically normal. Thus, it may not be possible to construct meaningful confidence intervals for the forecasts.

Removing outlying and extreme observations significantly reduces the size and variation in the GARCH(1,1) forecasts, resulting in significantly lower out-of-sample predictive power as measured by R^2. Such removal results in significantly improved forecasts for periods of low volatility and poorer forecasts for periods of high volatility. These observations would be missed by examining only the mean forecast error measures for the entire sample.

Overall, the Chen and Liu [1993a] method for detecting and removing outlying and extreme observations enabled an improved estimation of AR(1)-GARCH(1,1) processes, with a greater frequency of validation of the regularity conditions, less biased estimates, and reduced forecast errors for periods of low volatility. An added advantage of this method, as compared with robust estimation methods, is that valuable information can be obtained regarding the characteristics of outliers, such as their size, type and position.

6 References

Abraham, B., and G.E.P. Box, Bayesian analysis of some outlier problems in time series analysis, Biometrika, 66, 229-236, 1979.

Aggarwal, R., C. Inclan, and R. Leal, Volatility in emerging stock markets, Working Paper, Georgetown University, Washington, 1999.

Andersen, T.G., and T. Bollerslev, Answering the skeptics: yes, standard volatility models do provide accurate forecasts, International Economic Review, 39, 885-905, 1998.

Baillie, R.T., and P.C. McMahon, The foreign exchange market, Cambridge University Press, Cambridge, 1989.

Balke, N.S., Detecting level shifts in time series, Journal of Business & Economic Statistics, 1, 81-92, 1993.

Balke, N.S., and T. Fomby, Shifting trends, segemented trends, and infrequent permanent shocks, Journal of Monetary Economics, 28, 61-85, 1991.

Balke, N.S., and T. Fomby, Large innovations, small innovations, and economic fluctuations: outliers in macroeconomic time series, Journal of Applied Econometrics, 9, 181-200, 1994.

Beckman, R.J., and R.D. Cook, Outlier...s, Technometrics, 25, 119-149, 1983.

Bera, A.K., and M.L. Higgins, ARCH models: properties, estimation and testing, Journal of Economic Surveys, 7, 305-362, 1993.

Bernoulli, D., The most probable choice between several discrepant observations and the formation therefrom of the most likely induction, in C. G. Allen, Biometrika, 48, 3-13, 1961.

Blair, B., S.H Poon, and S.J. Taylor, Asymmetric and crash effects in stock volatility for the S&P 100 index and its constituents, Working Paper No. 98/003, Department of Accounting & Finance, Lancaster University, 1998.

Blanchard, O.J., and M.W. Watson, Are business cycles alike?, in The American Business Cycle: Continuity and Change, R.J. Gordon (ed.), University of Chicago Press, Chicago, 1986.

Blattberg, R.C., and N.J. Gonedes, A comparison of the stable and student distributions as statistical models for stock prices, Journal of Business, 47, 244-280, 1974.

Bollen, N.P.B., S.F. Gray, and R.E. Whaley, Regime-switching in foreign exchange rates: evidence from currency option prices, Journal of Econometrics, 94, 239-276, 2000.

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

Bollerslev, T., A conditional heteroskedastic time series for speculative prices and rates of return, Review of Economics and Statistics, 69, 542-547, 1987.

Bollerslev, T., R.Y Chou, and K.F. Kroner, ARCH modelling in finance: a selective review of the theory and empirical evidence, Journal of Econometrics, 52, 5-59, 1992.

Bollerslev, T., and R.F. Engle, modelling the persistence of conditional variances, Econometric Reviews, 5, 1-50, 1986.

Bollerslev, T., and J.M. Wooldridge, Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances, Econometric Reviews, 11, 143-172, 1992.

Box, G.E.P., and G.C. Tiao, Intervention analysis with applications to environmental and economic problems, Journal of the American Statistical Association, 70, 70-79, 1975.

Burn, D.A., and J.K. Ord, The effect of time series outliers on the sample autocorrelation coefficient, Proceedings of the Business and Economic Statistics Section, American Statistical Association, 581-585, 1984.

Cai, J., A Markov model of switching-regime ARCH, Journal of Business & Economic Statistics, 12, 309-316, 1994.

Chang, I., Outliers in time series, Unpublished Ph.D. thesis, Department of Statistics, University of Wisconsin-Madison, 1982.

Chang, I., and G.C. Tiao, Effect of exogenous interventions on the estimates of time series parameters, Proceedings of the Business and Economic Statistics Section, American Statistical Association, 532-537, 1983.

Chang, I., G.C. Tiao, and C. Chen, Estimation of time series parameters in the presence of outliers, Technometrics, 3, 193-204, 1988.

Chen, C., and L.M. Liu, Joint Estimation of model parameters and outlier effects in time series, Journal of the American Statistical Association, 88, 284-297, 1993a.

Chen, C., and L.M. Liu, Forecasting time series with outliers, Journal of Forecasting, 12, 13-35, 1993b.

Chen, C., and G.C. Tiao, Random level-shift time series models, ARIMA approximations, and level shift detection, Journal of Business & Economic Statistics, 8, 83-97, 1990.

Chuang, A., and B. Abraham, Comparison of parameter estimation methods in time series with outliers: a simulation study, Proceedings of the Business and Economic Statistics Section, American Statistical Association, 83-92, 1989.

Clark, P., A subordinated stochastic processes model with finite variance for speculative prices, Econometrica, 41, 135-156, 1973.

Danielsson, J., and C.G. de Vries, Value-at-risk and extreme returns, Discussion Paper 98-017/2, Tinbergen Institute, Erasmus University Rotterdam, 1997.

Deutsch, S.J., J.E. Richards, and J.J. Swain, Effects of a single outlier on ARMA identification, Communications in Statistics: Theory and Methods, 19, 2207-2227, 1990.

Diebold, F.X., Empirical modelling of exchange rate dynamics., Springer Verlag, New York, 1988.

van Dijk, D., P.H. Franses, and A. Lucas, Testing for ARCH in the presence of additive outliers, Journal of Applied Econometrics, 14, 539-562, 1999.

Ding, Z., C.W.J. Granger, and R.F. Engle, A long memory property of stock market returns and a new model, Journal of Empirical Finance, 1, 83-106, 1993.

Dueker, M.J., Markov switching in GARCH processes and mean-reverting stock-market volatility, Journal of Business & Economic Statistics, 15, 26-34, 1997.

Ederington, L.H., and W. Guan, Forecasting volatility, Working Paper, Price College of Business, University of Oklahoma, 1998.

Engle, R. F., Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, Econometrica, 50, 987-1007, 1982.

Engle, R.F., and G. Gonzalez-Rivera, Semiparametric ARCH models, Journal of Business & Economic Statistics, 9, 345-359, 1991.

Engle, R.F., and G.G.J. Lee, A permanent and transitory component model of stock return volatility, in Cointegration, Causality, and Forecasting: A Festschrift in Honor of Clive W.J. Granger, R. Engle, and H. White (eds.), Oxford University Press, 475-497, 1999.

Engle, R. F., and C. Mustafa, Implied ARCH models from option prices, Journal of Econometrics, 52, 289-311, 1992.

Fox, A. J., Outliers in time series, Journal of the Royal Statistical Society B, 34, 350-363, 1972.

Franses, P.H., and H. Ghijsels, Additive outliers, GARCH and forecasting volatility, International Journal of Forecasting, 15, 1-9, 1999.

Franses, P.H., and D. van Dijk, Forecasting stock market volatility using (non-linear) GARCH models, Journal of Forecasting, 15, 229-235, 1996.

Franses, P.H., and D. van Dijk, Outlier detection in the GARCH(1,1) model, Discussion Paper EI-9926/A, Tinbergen Institute, Erasmus University Rotterdam, 1999.

Franses, P.H., D. van Dijk, and A. Lucas, Short patches of outliers, ARCH and volatility modelling, Discussion Paper 98-057/4, Tinbergen Institute, Erasmus University Rotterdam, 1998.

Friedman, B.M., and D.I. Laibson, Economic implications of extraordinary movements in stock prices, Brooking Papers on Economic Activity, 2, 137-189, 1989.

Gannon, G.L., Unconditional first and conditional second moment effects: index portfolios and index futures, Research in Finance Supplement to 1996, A.H. Chen, and K.C. Chan (eds.), 143-158, 1996.

Ghose, D., and K. F. Kroner, The relationship between GARCH and symmetric stable processes: finding the source of fat tails in financial data, Journal of Empirical Finance, 2, 225-251.

Glosten, L., R. Jagannathan, and D. Runkle, On the relation between the expected value and volatility on nominal excess return and innovations, Journal of Finance, 46, 1779-1801, 1992.

Gray, S.F., Modelling the conditional distribution of interest rates as a regime-switching process, Journal of Financial Economics, 42, 27-62.

Guillaume, D.M., O.V. Pictet, and M.M. Dacorogna, On the intra-daily performance of GARCH processes, Internal Document DMG. 1994-07-31, Olsen & Associates, Research Institute for Applied Economics, Zurich, Switzerland, 1994.

de Haan, L., S.I. Resnick, H. Rootzen, and C.G. de Vries, Extremal behaviour of solutions to a stochastic differential equation with apllications to ARCH processes, Stochastic Process and Their Applications, 32, 213-224, 1989.

Hamilton, J.D., and R. Susmel, Autoregressive conditional heteroskedasticity and changes in regime, Journal of Econometrics, 64, 307-333, 1994.

Harris, L., S&P 500 cash stock price volatilities, Journal of Finance, 44, 1155-1175, 1989.

He, C., and T. Terasvirta, Fourth moment structure of the GARCH(p,q) process, to appear in Econometric Theory, 16, 2000.

Hillmer, S.C., Monitoring and adjusting forecasts in the presence of additive outliers, Journal of Forecasting, 3, 205-215, 1984.

Hillmer, S.C, W.R. Bell, and G.C. Tiao, Modelling considerations in the seasonal adjustment of economic time series, A. Zellner (ed.), Applied Time Series Analysis of Economic Data, Washington, D.C., U.S. Bureau of the Census, 74-100, 1983.

Hotta, L.K., The effect of additive outliers on the estimates from aggregated and disaggregated ARIMA models, International Journal of Forecasting, 9, 85-93, 1993.

Hotta, L.K., and R.S. Tsay, Outliers in GARCH processes, Working Paper, Graduate School of Business, University of Chicago, 1998.

Hsieh, D.A., Modelling heteroskedasticity in daily foreign-exchange rates, Journal of Business & Economic Statistics, 7, 307-317, 1989.

Jansen, D.W., and C.G. de Vries, On the frequency of large stock returns: putting booms and busts into perspective, Review of Economics and Statistics, 73, 18-24, 1991.

Justel, A., D. Pena, and R. S. Tsay, Detection of outlier patches in autoregressive time series, Working Paper, Graduate School of Business, University of Chicago, 1998.

Kim, C.J., and M.J. Kim, Transient fads and the crash of '87, Journal of Applied Econometrics, 11, 41-58, 1996.

Klaassen, F, Improving GARCH volatility forecasts, Working Paper, Department of Econometrics, Tilburg University, 1998.

Kleiner, B., Martin, R.D. and D.J. Thomson, Robust estimation of power spectra, Journal of the Royal Statistical Society B, 41, 313-351, 1979.

Kon, S., Models of stock returns: A comparison, Journal of Finance, 39, 147-165, 1984.

Koutmos, G., Asymmetries in the conditional mean and the conditional variance: evidence from nine stock markets, Journal of Economics and Business, 50, 277-290, 1998.

Lamoureux, C.G., and W.D. Lastrapes, Persistence in variance, structural change, and the GARCH model, Journal of Business & Economic Statistics, 8, 225-234, 1990.

Lamoureux, C.G., and W.D. Lastrapes, Forecasting stock-return variance: toward an understanding of stochastic implied volatilities, Review of Financial studies, 6, 293-326, 1993.

Lastrapes, W.D., Exchange rate volatility and the U.S. monetary policy: an ARCH application, Journal of Money, Credit, and Banking, 21, 66-77, 1989.

Ledolter, J., The effects of additive outliers on the forecasts from ARIMA models, International Journal of Forecasting, 5, 231-240, 1989.

Lee, S.W., and B.E. Hansen, Asymptotic theory for the GARCH(1,1) quasi-maximum likelihood estimator, Econometric Theory, 10, 29-52, 1994.

Ling, S., and M. McAleer, Non-nested tests of GARCH and E-GARCH, Working Paper, Department of Economics, University of Western Australia, 1999a.

Ling, S., and M. McAleer, Necessary and sufficient moment conditions for the

GARCH(r,s) and asymmetric power GARCH(r,s) models, 1999b. Submitted.

Ling, S., and M. McAleer, On adaptive estimation in nonstationary ARMA models with GARCH errors, 1999c. Submitted.

Liu, L.M., Dynamic relationship analysis of U.S. gasoline and crude oil prices, Journal of Forecasting, 10, 521-547, 1991.

Longin, F.M., The asymptotic distribution of extreme stock market returns, Journal of Business, 69, 383-408, 1996.

Lumsdaine, R.L., Finite-sample properties of the maximum likelihood estimator in GARCH(1,1) and IGARCH(1,1) models: a Monte-Carlo investigation, Journal of Business & Economic Statistics, 13, 1-10, 1995.

Lumsdaine, R.L., Consistency and asymptotic normality of the quasi-maximum likelihood estimator in IGARCH(1,1) and covariance stationary GARCH(1,1) models, Econometrica, 64, 575-596, 1996.

Mandelbrot, B., The variation of certain speculative prices, Journal of Business, 36, 394-419, 1963.

Mandelbrot, B., Fisher, A., and L. Calvet, A multifractional model of asset returns., Cowles Foundation Discussion Paper No. 1164, Yale University, 1997.

McKenzie, M.D., ARCH modelling of Australian bilateral exchange rate data, Applied Economics, 7, 147-164, 1997.

Nelson, D.B., Modelling stock market volatility changes, Proceedings of the Business and Economics Statistics Section, American Statistical Association, 93-98, 1989.

Nelson, D.B., ARCH models as diffusion approximations, Journal of econometrics, 45, 7-38, 1990a.

Nelson, D.B., Stationarity and persistence in the GARCH(1,1) model, Journal of Econometrics, 45, 7-38, 1990b.

Nelson, D.B., Conditional heteroskedasticity in asset returns: a new approach, Econometrica, 59, 347-370, 1991.

Nelson, D.B., Filtering and forecasting with misspecified ARCH models I. Getting the right variance with the wrong model, Journal of Econometrics, 52, 61-90, 1992.

Nelson, P.B., and C.Q. Cao, A note on the inequality constraints in the univariate GARCH model, Discussion Paper, University of Chicago, 1991.

Nelson, D.B., and D. P. Foster, Asymptotic filtering theory for univariate ARCH models, Econometrica, 62, 1-41, 1994.

Nelson, D.B., and D.P. Foster, Filtering and forecasting with misspecified ARCH models II: making the right forecast with the wrong model, Journal of Econometrics, 67, 303-335, 1995.

Poterba, J. M., and L. H. Summers, The persistence of volatility and stock market fluctuations, American Economic Review, 76, 1142-1151, 1986.

Praetz, P., The distribution of share price changes, Journal of Business, 45, 49-55, 1972.

Press, S.J., A compound events model for security prices, Journal of Business, 40, 317-335, 1967.

Safvenblad, P., Trading volume and serial correlation: empirical evidence from the Stockholm stock exchange, Working Paper Series in Economics and Finance

No. 191, Stockholm School of Economics, 1997.

Sakata, S., and H. White, High breakdown point conditional dispersion estimation with application to S&P 500 daily returns volatility, Econometrica, 66, 529-567, 1998.

Schwert, G.W., Why does stock market volatility change over time?, Journal of Finance, 44, 1115-1153, 1989a.

Schwert, G.W., Business cycles, financial crisis and stock volatility, Carnegie-Rochester Conference Series on Public Policy, 31, 83-125, 1989b.

Schwert, G.W., Stock volatility and the crash of '87, Review of Financial Studies, 3, 77-102, 1990.

Schwert, G.W., Stock market volatility: ten years after the crash, to appear in Journal of Finance, 2000.

Sentana, E., Quadratic ARCH models: a potential re-interpretation of ARCH models, Financial Markets Group Discussion Paper 122, London School of Economics, 1991.

Sentana, E., Quadratic ARCH models, Review of Economic Studies, 62, 639-661, 1995.

Sentana, E. and S.B. Wadhwani, Feedback traders and stock return autocorrelations: evidence from a century of daily data, Economic Journal, 102, 415-425, 1992.

Stein, J., Overreactions in the options market, Journal of Finance, 44, 1011-1023, 1989.

Susmel, R., Switching volatility in Latin American emerging equity markets, Emerging Market Quarterly, 2, 44-56, 1998.

Taylor, S.J., Modelling financial time series, 2nd ed., Wiley, New York, 1997.

Terasvirta, T., Two stylized facts and the GARCH(1,1) model, Working Paper Series in Economics and Finance No. 96, Stockholm School of Economics, 1996.

Terasvirta, T., and H.M. Anderson, Characterising nonlinearities in business cycles using smooth transition autoregressive models, Journal of Applied Econometrics, 7, S119-S136, 1992.

Thury, G., and M. Wuger, Outlier detection and adjustment: an emprical analysis for Austrian data, Empirica, 19, 71-93, 1996.

Trivez, F.J., Level shifts, temporary changes and forecasting, Journal of Forecasting, 14, 543-550, 1995.

Tsay, R.S., Time series model specification in the presence of outliers, Journal of the American Statistical Association, 81, 132-141, 1986.

Tsay, R.S., Outliers, level shifts, and variance change in time series, Journal of Forecasting, 7, 1-20, 1988.

Tsay, R.S., Extreme value analysis of financial data, Working Paper, Graduate School of Business, University of Chicago, 1999.

Weiss, A.A., Asymptotic theory for ARCH models: estimation and testing, Econometric Theory, 2, 107-131, 1986.

Table 1. Summary statistics for the original and adjusted daily time series for 1000 observations

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
S&P 500 (22/3/95-1/6/99)					
Mean	8.807e-4	9.985e-4	9.857e-4	9.977e-4	1.058e-3
Median	1.052e-3	1.052e-3	1.052e-3	1.052e-3	1.052e-3
σ	1.038e-2	9.476e-3	9.294e-3	8.885e-3	7.707e-3
Maximum(σ)	4.807	4.314	4.399	3.267	3.194
Minimum(σ)	-6.854	-3.270	-3.334	-3.441	-2.937
SR(σ)	11.661	7.585	7.733	6.708	6.131
# Obs.(+) $> 1/2/3/4/5\sigma$	116/19/5/2/0	136/25/4/1/0	137/25/3/1/0	143/24/2/0/0	157/18/1/0/0
# Obs.(-) $> 1/2/3/4/5\sigma$	114/26/8/2/2	126/35/4/0/0	129/36/4/0/0	136/40/3/0/0	145/28/2/0/0
Skewness	-0.577	-0.035	-0.077	-0.122	-0.0175
Kurtosis	8.839	4.086	3.907	3.382	2.908
LM(N)	1476.03	49.35	35.25	8.54	0.40
$\frac{K_e}{\sigma}$	519.17	114.61	97.59	42.99	
HSI (10/7/95-26/7/99)					
Mean	2.895e-4	3.269e-4	4.406e-4	5.117e-6	4.483e-4
Median	4.984e-4	4.984e-4	4.984e-4	4.984e-4	4.984e-4
σ	2.074e-2	1.829e-2	1.741e-2	1.619e-2	1.380e-2
Maximum(σ)	8.316	4.709	4.944	4.262	3.413
Minimum(σ)	-7.105	-4.019	-4.220	-3.639	-3.126
SR(σ)	15.421	8.727	9.164	7.901	6.539
# Obs.(+) $> 1/2/3/4/5\sigma$	87/29/9/3/2	105/31/11/3/0	109/30/11/2/0	126/29/6/2/0	134/24/4/0/0
# Obs.(-) $> 1/2/3/4/5\sigma$	95/27/7/3/2	111/36/7/2/0	120/35/6/1/0	131/33/6/0/0	152/29/6/0/0
Skewness	0.318*	0.132*	0.155*	0.0959*	-0.0095
Kurtosis	14.06*	5.826*	5.594*	4.713*	3.458*
LM(N)	5109.07*	335.66*	284.32*	123.86*	8.77*
$\frac{K_e}{\sigma}$	533.30	154.51	150.00	44.04	33.19
Nikkei 225 (3/8/84-28/6/88)					
Mean	9.992e-4	1.100e-3	1.168e-3	1.156e-3	nc
Median	1.224e-3	1.224e-3	1.224e-3	1.224e-3	
σ	1.191e-2	8.982e-3	8.434e-3	7.579e-3	
Maximum(σ)	9.054	4.561	4.857	4.306	
Minimum(σ)	-13.553	-5.103	-3.936	-4.009	
SR(σ)	22.607	9.664	8.793	8.315	
# Obs.(+) $> 1/2/3/4/5\sigma$	65/12/5/3/2	111/22/8/3/0	117/21/6/2/0	126/20/3/1/0	
# Obs.(-) $> 1/2/3/4/5\sigma$	76/21/7/5/3	76/21/7/5/3	121/34/12/1/0	145/33/6/1/0	
Skewness	-2.002*	-0.212*	-0.076	-0.133*	
Kurtosis	51.057*	6.126*	5.204*	3.928*	
LM(N)	96895.80*	414.63*	203.33*	38.85*	
$\frac{K_e}{\sigma}$	4035.01	348.03	261.32	122.44	

Table 1 (continued)

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
	GBP/USD Spot Exchange Rates (1/6/88 to 13/5/92)				
Mean	-1.375e-5	7.877e-6	2.685e-5	4.832e-5	2.779e-4
Median	1.804e-4	1.804e-4	1.804e-4	1.804e-4	1.804e-4
σ	6.990e-3	6.935e-3	6.608e-3	6.261e-3	5.392e-3
Maximum(σ)	4.093	3.986	3.513	2.922	2.650
Minimum(σ)	-3.954	-4.087	-3.469	-3.215	-2.556
SR(σ)	8.047	8.072	6.983	6.137	5.206
# Obs.(+) > 1/2/3/4/5σ	125/17/4/0/0	125/17/4/0/0	136/16/3/0/0	142/18/1/0	149/15/0/0/0
# Obs.(-) > 1/2/3/4/5σ	141/35/8/2/0	146/35/7/1/0	151/37/5/0/0	158/33/2/0	165/20/0/0/0
Skewness	-0.301*	-0.248*	-0.257*	-0.253*	0.0044
Kurtosis	4.608*	4.473*	3.712*	3.226*	2.480*
LM(N)	123.93*	100.64*	32.17*	12.77*	11.28*
$\frac{K_e}{\sigma}$	230.04	212.40	107.88	36.10	
	GB Spot Rates (5/4/79 to 7/6/83)				
Mean	5.461e-4	6.413e-4	5.144e-4	4.741e-4	3.524e-4
Median	1.305e-3	1.305e-3	1.305e-3	1.305e-3	1.305e-3
σ	2.399e-2	2.143e-2	2.111e-2	2.048e-2	1.777e-2
Maximum(σ)	5.089	4.355	3.553	3.257	2.695
Minimum(σ)	-7.449	-3.128	-3.176	-3.151	-3.062
SR(σ)	12.538	7.483	6.729	6.408	5.758
# Obs.(+) > 1/2/3/4/5σ	106/20/8/2/1	128/28/6/1/0	129/27/4/0/0	136/26/2/0/0	145/19/0/0/0/0
# Obs.(-) > 1/2/3/4/5σ	127/26/4/4/1	152/32/3/0/0	155/32/3/0/0	161/32/1/0/0	168/27/1/0/0/0
Skewness	-0.322*	0.0779	-0.032	-0.063	-0.129*
Kurtosis	8.538*	3.875*	3.566*	3.321*	2.681*
LM(N)	1294.98*	32.90*	13.508*	4.953*	7.000*
$\frac{K_e}{\sigma}$	230.85	40.83	26.81	15.67	

* Significant at the 5% level. SR(σ)is the Studentised Range of σ and is calculated as (max(σ)-min(σ)). LM(N) is the Jarque-Bera Lagrange multiplier test statistic for normality of the returns, which is asymptotically χ^2 distributed with two degrees of freedom under the null hypothesis of normality. nc signifies that the method did not converge at least once. K_e is the excess kurtosis.

Table 2. Summary of outlier statistics for the adjusted series for 1000 observations

	cv = 10	cv = 8	cv = 6	cv = 4
	S&P 500 (22/3/95-1/6/99)			
Total #	7	11	21	83
Mean#[a]	1.86	2.94	8.12	27.26
# Positive	2	4	8	34
Mean(σ)	4.795	4.052	3.596	2.313
Adjustment(σ)	(3.435)	(2.877)	(2.485)	(1.46)
# Negative	5	7	13	49
Mean(σ)	-4.797	-4.161	-3.380	-2.209
Adjustment(σ)	(3.645)	(3.090)	(2.400)	(1.364)
# Consecutive days	2	4	4	20
# Within 2-days	4	5	5	39
# Cluster of 2	1	2	2	5
# Cluster of 3	0	0	0	0
# Cluster > 3	0	0	0	0
# (+/-)	0	0	0	1
# (-/+)	1	2	2	3
# (+/+)	0	0	0	0
# (-/-)	0	0	0	1
	HSI(10/7/95-26/7/99)			
Total #	9	22	39	91
Mean #	3.49	5.00	11.39	40.15
# Positive	4	8	14	40
Mean(σ)	5.326	4.056	3.623	2.487
Adjustment(σ)	(3.466)	(2.653)	(2.273)	(1.537)
# Negative	5	14	25	51
Mean(σ)	-4.705	-3.140	-3.623	-2.243
Adjustment(σ)	(3.133)	(2.036)	(1.701)	(1.355)
# Consecutive days	5	11	19	34
# Within 2-days	5	15	22	45
# Cluster of 2	1	1	2	4
# Cluster of 3	1	1	1	1
# Cluster > 3	0	1	2	3
# (+/-)	0	1	4	5
# (-/+)	2	3	3	7
# (+/+)	0	0	1	5
# (-/-)	1	4	6	11
	Nikkei 225 (3/8/84-28/6/88)			
Total #	8	17	38	nc
Mean #	2.96	4.49	11.31	
# Positive	3	6	16	
Mean(σ)	7.159	5.062	3.290	
Adjustment(σ)	(5.383)	(3.733)	(2.392)	
# Negative	5	11	22	
Mean(σ)	-6.832	-4.569	-3.302	
Adjustment(σ)	(4.919)	(3.323)	(2.338)	
# Consecutive days	5	7	12	
# Within 2-days	6	8	17	
# Cluster of 2	1	2	3	
# Cluster of 3	1	1	2	
# Cluster > 3	0	0	0	
# (+/-)	1	1	2	
# (-/+)	2	2	4	
# (+/+)	0	1	1	
# (-/-)	0	0	0	

Table 2 (continued)

	cv = 10	cv = 8	cv = 6	cv = 4
GBP/USD Spot Exchange Rates (1/6/88 to 13/5/92)				
Total #	1	8	20	72
Mean #	0.052	3.36	8.27	26.20
# Positive	0	3	8	21
Mean(σ)	0	3.816	3.268	2.588
Adjustment(σ)	(0)	(2.741)	(2.294)	(1.719)
# Negative	1	5	12	51
Mean(σ)	-4.093	-3.739	-3.249	-2.379
Adjustment(σ)	(3.094)	(2.806)	(2.269)	(1.526)
# Consecutive days	0	0	0	20
# Within 2-days	0	0	0	25
# Consecutive days	5	11	19	34
# Within 2-days	5	15	22	45
# Cluster of 2	0	0	0	10
# Cluster of 3	0	0	0	0
# Cluster > 3	0	0	0	0
# (+/-)	0	0	0	0
# (-/+)	0	0	0	2
# (+/+)	0	0	0	1
# (-/-)	0	0	0	7
GB Spot Rates (5/4/79 to 7/6/83)				
Total #	10	12	19	69
Mean #	1.90	4.18	8.12	20.29
# Positive	5	7	11	39
Mean(σ)	4.045	3.909	3.468	2.409
Adjustment(σ)	(3.088)	(2.962)	(2.528)	(1.581)
# Negative	5	5	8	30
Mean(σ)	-4.825	-4.825	-4.010	-2.593
Adjustment(σ)	(3.882)	(3.882)	(3.101)	(1.786)
# Consecutive days	0	0	6	36
# Within 2-days	3	5	12	45
# Cluster of 2	0	0	3	9
# Cluster of 3	0	0	0	3
# Cluster > 3	0	0	0	1
# (+/-)	0	0	0	3
# (-/+)	0	0	0	3
# (+/+)	0	0	2	6
# (-/-)	0	0	1	6

[a] Mean number of observations down-weighted per estimation period, calculated over 500 windows.
[b] Clusters are defined as consecutive days with outliers. nc signifies that the method did not converge at least once.

Table 3. Mean values of 500 estimates of the parameters of the AR(1)-GARCH(1,1) model for the original and adjusted daily time series

	y		y^*							
			cv = 10		cv = 8		cv = 6		cv = 4	
	Estimate	(t-ratio)	Estimate	(t-ratio)	Estimate	(t-ratio)	Estimate	(t-ratio)	Estimate	(t-ratio)
					S&P 500 (22/3/95-1/6/99)					
μ	1.198e-3	(2.854)	1.140e-3	(2.749)	1.064e-3	(2.683)	1.084e3	(2.873)	1.182e-3	(3.543)
[std]	[2.547e-4]		[2.647e-4]		[1.484e-4]		[1.351e-4]		[9.579e-5]	
φ	0.060	(1.329)	0.043	(1.035)	0.055	(1.236)	0.053	(1.192)	0.042	(1.095)
[std]	[0.0374]		[0.060]		[0.043]		[0.041]		[0.034]	
ω	6.206e-6	(1.751)	2.589e-6	(1.295)	1.960e-6	(1.126)	2.805e-6	(1.219	3.222e-6	(1.115)
[std]	[3.491e-6]		[1.966e-6]		[1.177e-6]		[2.483e-6]		[2.229e-6]	
α	0.123	(1.998)	0.059	(2.260)	0.051	(2.164)	0.050	(2.187)	0.055	(1.866)
[std]	[0.035]		[0.014]		[0.008]		[0.011]		[0.019]	
β	0.834	(16.441)	0.919	(26.747)	0.931	(27.387)	0.922	(26.368)	0.894	(16.105)
[std]	[0.049]		[0.025]		[0.014]		[0.028]		[0.046]	
MLL	3.226		3.269		3.283		3.328		3.438	
Diagnostics:										
Mean	-0.0435		-0.0225		-0.0189		-0.0138		-0.0103	
Std	0.997		0.995		0.995		0.996		0.998	
Skewness	-0.810		-0.333		-0.332		-0.197		-0.027	
Kurtosis	5.738		3.780		3.553		3.074		2.457	
LM(N)	228.27		24.06		16.57		3.79		6.86	
Q(12)	15.40		23.80		23.28		22.60		19.93	
$Q(12)^2$	5.69		8.80		8.70		7.65		10.80	
					HSI (10/7/95-26/7/99)					
μ	8.015e-4	(1.384)	7.537e-4	(1.328)	7.567e-4	(1.338)	6.448e-4	(1.211)	5.337e-4	(1.23)
[std]	[3.56e-4]		[3.871e-4]		[3.687e-4]		[4.150e-4]		[5.766e-4]	
φ	0.113	(2.394)	0.080	(1.748)	0.069	(1.482)	0.0622	(1.429)	0.050	(1.31)
[std]	[0.013]		[0.0213]		[0.0179]		[0.028]		[0.015]	
ω	1.403e-5	(1.902)	1.257e-5	(1.601)	1.088e-5	(1.509)	9.41e-6	(1.327)	5.091e-6	(1.16)
[std]	[8.958e-6]		[1.166e-5]		[1.042e-5]		[1.119e-5]		[6.006e-6]	
α	0.167	(4.590)	0.133	(3.817)	0.120	(3.700)	0.094	(3.424)	0.054	(3.49)
[std]	[0.034]		[0.034]		[0.039]		[0.039]		[0.024]	
β	0.807	(20.803)	0.842	(19.423)	0.859	(21.762)	0.888	(26.934)	0.931	(48.63)
[std]	[0.023]		[0.037]		[0.042]		[0.050]		[0.030]	
MLL	2.657		2.701		2.714		2.749		2.884	
Diagnostics:										
Mean	-0.0584		-0.0541		-0.0503		-0.0428		-0.0248	
Std	0.996		0.996		0.995		0.994		0.994	
Skewness	-0.492		-0.405		-0.340		-0.313		-0.090	
Kurtosis	4.938		3.896		3.647		3.420		2.716	
LM(N)	144.60		39.63		20.87		16.50		4.51	
Q(12)	12.95		13.02		12.86		15.98		14.22	
$Q(12)^2$	6.19		9.52		8.37		8.16		9.33	
					Nikkei 225 (3/8/84-28/6/88)					
μ	1.349e-3	(3.287)	1.397e-3	(3.741)	1.396e-3	(3.901)	1.410e-3	(4.063)	nc	
[std]	[5.352e-4]		[2.073e-4]		[2.063e-4]		[1.878e-4]			
φ	0.107	(2.070)	0.081	(1.650)	0.074	(1.533)	0.070	(1.479)		
[std]	[0.039]		[0.058]		[0.046]		[0.054]			
ω	3.070e-5	(2.633)	6.428e-6	(1.487)	5.165e-6	(1.651)	4.290e-6	(1.599)		
[std]	[2.080e-5]		[4.411e-6]		[1.682e-6]		[2.244e-6]			
α	0.435	(2.469)	0.156	(2.801)	0.152	(3.096)	0.137	(2.928)		
[std]	[0.273]		[0.033]		[0.029]		[0.043]			
β	0.377	(4.175)	0.775	(9.775)	0.790	(10.780)	0.811	(11.748)		
[std]	[0.308]		[0.073]		[0.031]		[0.050]			
MLL	3.254		3.343		3.364		3.425			

Table 3 (continued)

	y		y^*							
			cv = 10		cv = 8		cv = 6		cv = 4	
	Estimate	(t-ratio)	Estimate	(t-ratio)	Estimate	(t-ratio)	Estimate	(t-ratio)	Estimate	(t-ratio)
Diagnostics:										
Mean	-0.0473		-0.0361		-0.0318		-0.0276			
Std	1.000		0.998		0.999		0.998			
Skewness	-1.435		-0.570		0.461		-0.260			
Kurtosis	13.186		4.988		4.353		3.488			
LM(N)	4858.39		120.76		66.52		12.93			
Q(12)	7.46		8.78		8.61		9.25			
$Q(12)^2$	13.24		8.45		10.98		11.77			
GBP/USD Spot Exchange Rates (1/6/88 to 13/5/92)										
μ	3.566e-4	(1.254)	3.579e-4	(1.259)	3.513e-5	(1.300)	3.400e-5	(1.319)	4.290e-4	(1.81)
[std]	[8.924e-5]		[9.031e-5]		[7.986e-5]		[9.106e-5]		[7.482e-5]	
φ	0.112	(2.315)	0.112	(2.310)	0.105	(2.292)	0.107	(2.490)	0.089	(2.33)
[std]	[0.037]		[0.037]		[0.036]		[0.034]		[0.027]	
ω	1.515e-6	(1.818)	1.517e-6	(1.818)	1.961e-6	(2.071)	1.759e-6	(2.262)	1.434e-6	(2.11)
[std]	[5.484e-7]		[5.821e-7]		[6.234e-7]		[5.102e-7]		[3.199e-7]	
α	0.082	(3.187)	0.082	(3.175)	0.086	(3.150)	0.073	(3.449)	0.049	(2.67)
[std]	[0.017]		[0.017]		[0.020]		[0.017]		[0.009]	
β	0.889	(29.397)	0.889	(29.175)	0.870	(25.795)	0.881	(32.665)	0.904	(32.54)
[std]	[0.025]		[0.025]		[0.030]		[0.025]		[0.013]	
MLL	3.611		3.612		3.657		3.700		3.804	
Diagnostics:										
Mean	-0.0321		-0.0320		-0.0260		-0.0221		-0.0133	
Std	1.000		1.000		1.000		1.000		1.000	
Skewness	-0.374		-0.371		-0.295		-0.238		-0.068	
Kurtosis	4.283		4.278		3.656		3.225		2.673	
LM(N)	47.87		47.42		19.15		6.51		2.89	
Q(12)	12.37		12.49		12.03		12.01		13.93	
$Q(12)^2$	13.35		13.50		13.31		11.39		8.97	
GB Spot Rates (5/4/79 to 7/6/83)										
μ	-6.236e-4	(-0.739)	-5.715e-4	(-0.686)	-5.525e-4	(-0.665)	-5.143e-4	(-0.614)	-4.946e-4	(-0.667)
[std]	[9.346e-4]		[9.132e-4]		[9.273e-4]		[8.820e-4]		[9.467e-4]	
φ	-0.1512	(-3.215)	-0.141	(-3.134)	-0.138	(-3.053)	-0.141	(-3.085)	-0.147	(-3.499)
[std]	[0.024]		[0.026]		[0.026]		[0.029]		[0.029]	
ω	6.051e-5	(1.968)	4.473e-5	(1.993)	3.942e-5	(2.018)	4.284e-5	(5.464)	2.383e-5	(1.929)
[std]	[3.436e-5]		[1.570e-5]		[1.269e-5]		[3.435e-5]		[7.350e-6]	
α	0.166	(2.541)	0.123	(2.632)	0.100	(2.853)	0.093	(2.610)	0.085	(3.003)
[std]	[0.051]		[0.064]		[0.042]		[0.041]		[0.030]	
β	0.706	(8.080)	0.779	(10.683)	0.807	(12.815)	0.801	(11.852)	0.848	(18.885)
[std]	[0.100]		[0.062]		[0.039]		[0.102]		[0.032]	
MLL	2.441		2.469		2.483		2.492		2.568	
Diagnostics:										
Mean	0.0082		0.0078		0.00497		0.0486		0.0080	
Std	1.000		1.000		1.000		1.000		1.000	
Skewness	-0.101		-0.040		-0.058		-0.051		-0.019	
Kurtosis	3.927		3.290		3.131		3.046		2.618	
LM(N)	28.70		2.77		1.24		0.816		3.503	
Q(12)	11.64		10.32		8.14		7.72		6.59	
$Q(12)^2$	15.23		14.10		17.11		14.05		12.41	

The robust t-ratios are those of Bollerslev and Wooldridge [1992]. MLL denotes the maximum value of the log-likelihood function. LM(N) is the Jarque-Bera LM statistic for normality of η_t, which is asymptotically χ^2 distributed with two degrees of freedom under the null hypothesis of normality. Q(12) is the Ljung-Box test statistic for serial correlation in η_t with 12 lags. $Q(12)^2$ is the Ljung-Box test statistic for an ARCH process based on η_t^2. Under the null hypotheses of uncorrelated and conditionally homoskedastic errors, respectively, the test statistics are asymptotically χ^2 distributed with 12 degrees of freedom. nc signifies that the method did not converge at least once.

Table 4. Mean values of 500 estimates of $(\alpha+\beta)$ and $(3\alpha^2+2\alpha\beta+\beta^2)$ for the AR(1)-GARCH(1,1) model for the original and adjusted time series

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
S&P 500 (22/3/95-1/6/99)					
$(\alpha+\beta)$	0.956	0.977	0.982	0.972	0.950
[std]	[0.019]	[0.012]	[0.008]	[0.020]	[0.035]
$\# > 1$	0	0	0	0	0
$(3\alpha^2+2\alpha\beta+\beta^2)$	0.948	0.963	0.969	0.950	0.910
[std]	[0.029]	[0.022]	[0.015]	[0.037]	[0.064]
$\# > 1$	13	0	0	0	0
HSI (10/7/95-26/7/99)					
$(\alpha+\beta)$	0.975	0.975	0.979	0.965	0.985
[std]	[0.024]	[0.018]	[0.016]	[0.017]	[0.014]
$\# > 1$	29	0	0	14	69
$(3\alpha^2+2\alpha\beta+\beta^2)$	1.009	0.989	0.991	0.939	0.977
[std]	[0.061]	[0.040]	[0.033]	[0.031]	[0.027]
$\# > 1$	332	226	219	189	76
Nikkei 225 (3/8/84-28/6/88)					
$(\alpha+\beta)$	0.812	0.930	0.942	0.948	nc
[std]	[0.247]	[0.056]	[0.022]	[0.014]	
$\# > 1$	22	0	0	0	
$(3\alpha^2+2\alpha\beta+\beta^2)$	1.249	0.919	0.935	0.940	
[std]	[0.745]	[0.093]	[0.050]	[0.028]	
$\# > 1$	270	45	24	4	
GBP/USD Spot Exchange Rates (1/6/88 to 13/5/92)					
$(\alpha+\beta)$	0.971	0.970	0.955	0.955	0.952
[std]	[0.010]	[0.010]	[0.012]	[0.010]	[0.009]
$\# > 1$	0	0	0	0	0
$(3\alpha^2+2\alpha\beta+\beta^2)$	0.956	0.956	0.928	0.923	0.911
[std]	[0.017]	[0.017]	[0.019]	[0.016]	[0.016]
$\# > 1$	0	0	0	0	0
GB Spot Rates (5/4/79 to 7/6/83)					
$(\alpha+\beta)$	0.872	0.902	0.907	0.894	0.933
[std]	[0.106]	[0.039]	[0.039]	[0.105]	[0.020]
$\# > 1$	0	0	0	0	0
$(3\alpha^2+2\alpha\beta+\beta^2)$	0.832	0.853	0.847	0.830	0.887
[std]	[0.177]	[0.086]	[0.080]	[0.133]	[0.042]
$\# > 1$	447	13	0	0	0

nc signifies that the method did not converge at least once.

Table 5. Forecast errors of the AR(1)-GARCH(1,1) model for the original and adjusted daily time series

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
		S&P 500 (22/3/95-1/6/99)			
ME	2.486e-3	1.506e-3	1.314e-3	6.424e-4	-7.226e-4
S	-1.72	-2.42	-2.48	-2.53	-2.55
K	12.71	15.16	15.09	15.46	15.33
MAE	6.790e-3	6.478e-3	6.456e-3	6.165e-3	5.890e-3
	(100)	(95.4)	(95.1)	(90.8)	(86.7)
RMSE	9.035e-3	8.703e-3	8.718e-3	8.550e-3	8.538e-3
	(100)	(96.3)	(96.5)	(94.6)	(94.5)
RMedSE	5.923e-3	5.878e-3	5.813e-3	5.336e-3	4.806e-3
	(100)	(99.2)	(98.1)	(90.1)	(81.1)
RMSE(+)	8.153e-3	6.866e-3	6.644e-3	6.076e-3	5.164e-3
	(100)	(84.2)	(81.5)	(74.5)	(63.3)
RMSE(-)	1.071e-2	1.157e-2	1.186e-2	1.182e-2	1.167e-2
	(100)	(108.0)	(110.7)	(110.4)	(109.0)
SMAPE	72.82	71.11	71.13	69.47	68.36
	(100)	(97.7)	(97.7)	(95.4)	(93.9)
SMedAPE	60.19	60.82	61.90	56.94	61.10
	(100)	(101.0)	(102.8)	(94.6)	(101.5)
SMWAPE	52.97	55.94	57.09	57.93	63.50
	(100)	(105.6)	(107.8)	(109.4)	(119.9)
SMedWAPE	32.90	32.37	32.32	29.44	28.48
	100	(98.4)	(98.2)	(89.5)	(86.6)
Under(%)	31.4	33.0	33.0	35.2	42.2
PTTEST	-10.12	-9.76	-10.12	-10.10	-9.88
R^2(%)	5.55	2.38	1.28	2.22	2.34
R_e^2(%)	7.37	1.68	1.65	0.48	0.17
		HSI (10/7/95-26/7/99)			
ME	5.456e-3	3.611e-3	2.956e-3	1.415e-3	-2.111e-3
S	-1.11	-2.19	-2.24	-2.68	-2.77
K	8.15	12.89	13.13	16.40	16.43
MAE	1.515e-2	1.440e-2	1.433e-2	1.414e-2	1.339e-2
	(100)	(95.0)	(94.6)	(93.3)	(88.4)
RMSE	2.021e-2	1.946e-2	1.961e-2	1.996e-2	2.026e-2
	(100)	(96.3)	(97.0)	(98.8)	(100.2)
RMedSE	1.259e-2	1.219e-2	1.187e-2	1.171e-2	1.002e-2
	(100)	(96.8)	(94.3)	(93.0)	(79.6)
RMSE(+)	1.830e-2	1.537e-2	1.493e-2	1.3710e-2	1.095e-2
	(100)	(84.0)	(81.6)	(74.9)	(59.8)
RMSE(-)	2.420e-2	2.642e-2	2.717e-2	2.828e-2	2.903e-2
	(100)	(109.2)	(112.3)	(116.9)	(120.0)
SMAPE	78.14	77.23	77.24	77.33	75.98
	(100)	(98.8)	(98.8)	(99.0)	(97.2)
SMedAPE	70.52	67.73	67.67	68.18	66.91
	(100)	(96.0)	(96.0)	(96.7)	(94.9)
SMWAPE	55.87	59.14	61.05	65.56	73.96
	(100)	(105.9)	(109.3)	(117.3)	(132.4)
SMedWAPE	31.85	31.82	31.67	31.70	27.52
	(100)	(99.9)	(99.4)	(99.5)	(86.4)
Under(%)	29.2	30.8	31.4	34.4	40.2
PTTEST	-8.38	-8.56	-8.57	-7.77	-7.54
R^2(%)	10.46	7.00	4.90	0.95	0.05
R_e^2(%)	8.36	1.42	1.96	3.20	4.67

Table 5 (continued)

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
		Nikkei 225 (3/8/84-28/6/88)			
ME	2.694e-3	1.460e-3	1.004e-3	2.881e-4	nc
S	-3.61	-6.85	-7.00	-7.89	
K	69.65	82.49	83.85	87.74	
MAE	6.617e-3	6.127e-3	5.951e-3	5.867e-3	
	(100)	(92.6)	(89.9)	(88.7)	
RMSE	1.189e-2	1.120e-2	1.101e-2	1.117e-2	
	(100)	(94.2)	(92.6)	(93.9)	
RMedSE	4.715e-3	4.500e-3	4.385e-3	4.319e-3	
	(100)	(95.4)	(93.0)	(91.6)	
RMSE(+)	9.744e-3	6.788e-3	6.337e-3	5.895e-3	
	(100)	(69.7)	(65.0)	(60.5)	
RMSE(-)	1.654e-2	1.748e-2	1.683e-2	1.662e-2	
	(100)	(105.7)	(101.8)	(100.5)	
SMAPE	74.42	72.39	71.79	71.90	
	(100)	(97.3)	(96.5)	(96.6)	
SMedAPE	60.81	57.24	57.09	57.00	
	(100)	(94.1)	(93.9)	(93.7)	
SMWAPE	56.44	59.51	60.39	64.62	
	(100)	(105.4)	(107.0)	(114.5)	
SMedWAPE	29.66	27.34	26.97	26.69	
	(100)	(92.2)	(90.9)	(90.0)	
PTTEST	-7.12	-7.68	-7.97	-8.00	
Under(%)	26.0	30.6	33.4	37.2	
R^2(%)	17.43	7.10	8.99	5.96	
R_e^2(%)	18.64	0.46	0.00	0.15	
		GBP/USD Spot Exchange Rates (1/6/88 to 13/5/92)			
ME	1.599e-3	1.598e-3	1.320e-3	8.937e-4	1.906e-4
S	-1.33	-1.33	-1.39	-1.39	-1.52
K	5.59	5.59	5.77	5.85	5.98
MAE	4.117e-3	4.117e-3	4.010e-3	3.874e-3	3.673e-3
	(100)	(100)	(97.4)	(94.1)	(89.2)
RMSE	5.056e-3	5.057e-3	4.984e-3	4.909e-3	4.819e-3
	(100)	(100)	(98.6)	(97.1)	(95.3)
RMedSE	3.750e-3	3.750e-3	3.557e-3	3.404e-3	3.262e-3
	(100)	(100)	(94.9)	(90.8)	(87.0)
RMSE(+)	4.539e-3	4.538e-3	4.306e-3	4.067e-3	3.581e-3
	(100)	(100)	(94.9)	(89.6)	(78.9)
RMSE(-)	6.222e-3	6.225e-3	6.395e-3	6.294e-3	6.284e-3
	(100)	(100)	(102.8)	(101.2)	(101.0)
SMAPE	76.98	76.98	76.19	75.20	73.84
	(100)	(100)	(99.0)	(97.7)	(95.9)
SMedAPE	66.69	66.69	66.08	66.44	65.51
	(100)	(100)	(99.1)	(99.6)	(98.2)
SMWAPE	53.12	53.14	54.12	55.94	59.74
	(100)	(100.3)	(101.8)	(105.3)	(112.5)
SMedWAPE	34.08	34.08	33.28	32.36	29.92
	(100)	(100)	(97.7)	(95.0)	(87.8)
PTTEST	-8.63	-8.63	-8.91	-9.00	-8.90
Under(%)	27.4	27.4	28.2	32.8	39.0
R^2(%)	3.61	2.59	2.85	1.38	0.68
R_e^2(%)	2.49	2.51	2.12	1.51	0.29

Table 5 (continued)

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
		GB Spot Rates (5/4/79 to 7/6/83)			
ME	4.955e-3	4.179e-3	3.952e-3	3.797e-3	2.577e-3
S	-2.09	-2.14	-2.16	-2.16	-2.17
K	12.79	12.60	12.62	12.57	12.54
MAE	1.133e-2	1.102e-2	1.099e-2	1.096e-2	1.053e-2
	(100)	(97.3)	(97.0)	(96.7)	(92.9)
RMSE	1.424e-2	1.391e-2	1.388e-2	1.387e-2	1.368e-2
	(100)	(97.7)	(97.5)	(97.4)	(96.1)
RMedSE	1.063e-2	1.030e-2	1.019e-2	1.008e-2	9.584e-3
	(100)	(96.9)	(95.9)	(94.8)	(90.2)
RMSE(+)	1.268e-2	1.195e-2	1.183e-2	1.171e-2	1.092e-2
	(100)	(94.2)	(93.3)	(92.4)	(86.1)
RMSE(-)	1.786e-2	1.807e-2	1.800e-2	1.813e-2	1.820e-2
	(100)	(101.2)	(100.8)	(101.5)	(101.9)
SMAPE	73.84	73.08	73.06	73.03	71.56
	(100)	(99.0)	(98.9)	(98.9)	(96.9)
SMedAPE	64.56	63.32	63.52	63.31	62.37
	(100)	(98.1)	(98.4)	(98.1)	(96.6)
SMWAPE	51.84	52.36	52.94	53.35	55.34
	(100)	(101.1)	(102.2)	(102.8)	(106.8)
SMedWAPE	35.45	35.06	35.58	35.52	34.03
	(100)	(98.9)	(100.4)	(100.2)	(98.5)
PTTEST	-7.91	-8.01	-8.11	-8.09	-8.07
Under(%)	26.6	27.6	28.6	28.8	32.0
R^2(%)	4.59	4.05	3.42	2.92	2.02
R_e^2(%)	2.09	0.21	0.11	0.12	0.66

Values in parentheses are the error measures calculated as a percentage of the values of the unadjusted series. R^2 is the coefficient of determination by regressing the ex-post volatility on the forecast volatility. R_e^2 is the coefficient of determination by regressing the forecast errors on the ex-post volatility. nc signifies that the method did not converge at least once. PTTEST is the Pesaran and Timmermann test statistic, which is asymptotically normally distributed. Under(%) is the percentage of forecasts that underpredict realised volatility. RMSE(+) and RMSE(-) are the RMSE measures for the positive and negative forecast errors, respectively. The loss functions are defined as follows:

$$\text{ME} = \frac{1}{N}\sum_{t=1}^{N}(\sqrt{h_t} - \sigma_t),\ \text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\sqrt{h_t} - \sigma_t)^2},\ \text{MAE} = \frac{1}{N}\sum_{t=1}^{N} \mid \sqrt{h_t} - \sigma_t \mid,$$

$$\text{SMAPE} = \frac{100}{N}\sum_{t=1}^{N}\left(\frac{|\sqrt{h_t}-\sigma_t|}{0.5(\sigma_t+\sqrt{h_t})}\right),\ \text{SMWAPE} = \frac{100}{N}\sum_{t=1}^{N}\left(\frac{\sigma_t}{\bar{\sigma}}\frac{|\sqrt{h_t}-\sigma_t|}{0.5(\sigma_t+\sqrt{h_t})}\right).$$

Table 6. Forecast errors for low and high volatilities of the AR(1)-GARCH(1,1) model for the original and adjusted daily time series

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
		HSI (10/7/95-26/7/99)			
# obs.	437				
	63				
ME	1.009e-2	8.689e-3	8.168e-3	6.893e-3	3.608e-3
	-2.671e-2	-3.161e-2	-3.320e-2	-3.658e-2	-4.179e-2
S	0.83	-0.06	-0.129	-0.39	-0.56
	-0.654	-2.14	-2.23	-2.60	-2.59
K	5.79	3.07	2.97	2.68	2.66
	5.38	8.66	8.96	10.64	10.47
MAE	1.313e-2	1.190e-2	1.162e-2	1.091e-2	9.296e-3
	(100)	(90.4)	(88.5)	(83.1)	(70.6)
	2.915e-2	3.177e-2	3.321e-2	3.658e-2	4.179e-2
	(100)	(109.0)	(113.9)	(125.5)	(143.4)
RMSE	1.680e-2	1.432e-2	1.395e-2	1.285e-2	1.096e-2
	(100)	(85.2)	(83.0)	(76.5)	(65.2)
	3.580e-2	3.976e-2	4.124e-2	4.492e-2	4.925e-2
	(100)	(111.1)	(115.2)	(125.5)	(137.6)
RMedSE	1.167e-2	1.124e-2	1.082e-2	1.045e-2	8.840e-3
	(100)	(96.3)	(92.7)	(89.5)	(75.7)
	2.516e-2	2.732e-2	2.882e-2	3.007e-2	3.417e-2
	(100)	(108.6)	(114.5)	(119.5)	(135.8)
RMSE(+)	1.823e-2	1.542e-2	1.495e-2	1.371e-2	1.095e-2
	(100)	(84.6)	(82.0)	(75.2)	(60.1)
	2.286e-2	3.109e-3	6.400e-4	0.00	0.00
	(100)	(13.6)	(2.8)	(0)	(0)
RMSE(-)	9.158e-3	9.226e-3	9.563e-3	9.806e-3	1.098e-2
	(100)	(100.7)	(104.4)	(107.1)	(119.8)
	3.671e-2	4.040e-2	4.158e-2	4.492e-2	4.925e-2
	(100)	(110.1)	(113.3)	(122.4)	(134.2)
SMAPE	79.92	77.95	77.31	75.94	71.55
	(100)	(97.5)	(96.7)	(95.0)	(89.5)
	65.79	72.17	76.77	86.96	106.76
	(100)	(109.7)	(116.7)	(132.2)	(162.3)
SMedAPE	69.24	64.20	62.52	60.71	56.28
	(100)	(92.7)	(90.3)	(87.7)	(81.3)
	73.62	77.94	79.11	83.93	104.27
	(100)	(105.9)	(107.5)	(114.0)	(141.6)
SMWAPE	48.80	46.49	46.46	46.46	48.09
	(100)	(95.3)	(95.2)	(95.2)	(98.5)
	69.70	79.26	84.27	95.95	115.13
	(100)	(113.7)	(120.9)	(137.7)	(165.2)
SMedWAPE	40.92	40.40	40.30	39.85	34.97
	(100)	(98.7)	(98.5)	(97.4)	(85.5)
	58.60	68.57	69.62	72.65	87.91
	(100)	(117.0)	(118.8)	(124.0)	(150.0)
PTTEST	-5.84	-6.31	-6.31	-6.50	-6.62
	0.50	0.74	0.50	2.25	2.10
Under(%)	20.1	21.3	21.74	24.9	31.6
	92.1	96.8	98.4	100	100
R^2(%)	2.44	1.75	1.41	0.65	0.49
	17.58	10.05	7.46	0.02	0.59

Table 6 continued ...

	y	y^*			
		cv = 10	cv = 8	cv = 6	cv = 4
	Nikkei 225 (3/8/84-28/6/88)				
# obs.	464				
	36				
ME	4.013e-3	3.344e-3	2.870e-3	2.192e-3	nc
	-1.431e-2	-1.741e-2	-2.305e-2	-2.426e-2	
S	3.99	0.50	0.16	0.05	
	-1.42	-3.23	-3.19	-3.25	
K	40.36	4.96	3.83	3.93	
	12.05	14.37	14.12	14.57	
MAE	5.396e-3	4.826e-3	4.617e-3	4.439e-3	
	(100)	(89.4)	(85.6)	(82.3)	
	2.236e-2	2.291e-2	2.314e-2	2.426e-2	
	(100)	(102.5)	(103.5)	(108.5)	
RMSE	7.913e-3	6.188e-3	5.779e-3	5.452e-3	
	(100)	(78.2)	(73.0)	(68,9)	
	3.402e-2	3.535e-2	3.542e-2	3.670e-2	
	(100)	(104.9)	(104.1)	(107.9)	
RMedSE	4.469e-3	4.268e-3	4.174e-3	3.985e-3	
	(100)	(95.5)	(93.4)	(89.2)	
	1.592e-2	1.741e-2	1.709e-2	1.706e-2	
	(100)	(109.4)	(107.3)	(107.2)	
RMSE(+)	8.663e-3	6.797e-3	6.346e-3	5.895e-3	
	(100)	(78.5)	(73.3)	(68.0)	
	3.933e-2	1.620e-3	1.620e-3	0	
	(100)	(4.1)	(4.1)	(0)	
RMSE(-)	4.092e-3	3.886e-3	4.015e-3	2.687e-3	
	(100)	(95.0)	(98.1)	(65.6)	
	3.308e-2	3.585e-2	3.592e-2	3.670e-2	
	(100)	(108.4)	(108.6)	(110.9)	
SMAPE	74.23	71.70	70.95	70.52	
	(100)	(96.6)	(95.6)	(95.0)	
	76.78	81.20	82.61	89.76	
	(100)	(105.0)	(107.6)	(116.9)	
SMedAPE	58.75	54.80	53.90	56.05	
	(100)	(93.3)	(91.7)	(95.4)	
	87.43	85.74	84.50	92.34	
	(100)	(98.1)	(96.6)	(105.6)	
SMWAPE	42.99	41.50	42.24	45.07	
	(100)	(96.5)	(98.3)	(104.8)	
	85.68	98.64	99.85	107.15	
	(100)	(115.1)	(116.5)	(125.1)	
SMedWAPE	38.23	36.03	35.52	35.27	
	(100)	(94.2)	(92.9)	(92.3)	
	59.31	65.35	64.20	66.54	
	(100)	(110.2)	(108.2)	(112.2)	
PTTEST	-6.56	-6.31	-6.31	-6.30	
	0.18	0.74	0.50	0.28	
Under(%)	21.3	25.4	28.4	32.3	
	86.1	97.2	97.2	100	
R^2(%)	0.30	3.30	3.11	1.88	
	14.72	0.14	0.32	0.30	

Values in parentheses are the error measures calculated as a percentage of the values of the unadjusted series. R^2 is the coefficient of determination by regressing the ex-post volatility on the forecast volatility. nc signifies that the method did not converge at least once. PTTEST is the Pesaran and Timmermann test statistic, which is asymptotically normally distributed. Under(%) is the percentage of forecasts that underpredict realised volatility. RMSE(+) and RMSE(-) are the RMSE measures for the positive and negative forecast errors, respectively. The loss functions are defined as follows:

$$\text{ME} = \frac{1}{N}\sum_{t=1}^{N}(\sqrt{h_t} - \sigma_t),\ \text{RMSE} = \sqrt{\frac{1}{N}\sum_{t=1}^{N}(\sqrt{h_t} - \sigma_t)^2},\ \text{MAE} = \frac{1}{N}\sum_{t=1}^{N} \mid \sqrt{h_t} - \sigma_t \mid,$$

$$\text{SMAPE} = \frac{100}{N}\sum_{t=1}^{N}\left(\frac{|\sqrt{h_t}-\sigma_t|}{0.5(\sigma_t+\sqrt{h_t})}\right),\ \text{SMWAPE} = \frac{100}{N}\sum_{t=1}^{N}\left(\frac{\sigma_t}{\bar{\sigma}}\frac{|\sqrt{h_t}-\sigma_t|}{0.5(\sigma_t+\sqrt{h_t})}\right).$$