OUTLIERS AND CONDITIONAL AUTOREGRESSIVE HETEROSCEDASTICITY IN TIME SERIES

M. Angeles Carnero, Daniel Peña and Esther Ruiz*

Abstract

This paper reviews the literature on GARCH-type models proposed to represent the dynamic evolution of conditional variances. Effects of level outliers on the diagnostic and estimation of GARCH models are also studied. Both outliers and conditional heteroscedasticity can generate time series with excess kurtosis and autocorrelated squared observations. Consequently, both phenomena can be confused. However, since outliers are generated by unexpected events and the conditional variances are predictable, it is important to identify which one is producing the observed features in the data. We compare two alternative procedures for dealing with the simultaneous presence of outliers and conditional heteroscedasticity in time series. The first one is to clean the series of outliers before fitting a GARCH model. The second is to estimate first the GARCH model and then to clean of outliers by using the residuals adjusted by its conditional variance. It is shown that both approaches may result in different estimated conditional variances.

Keywords: GARCH; EGARCH; CHARMA; Stochastic Volatility; Asymmetry; autocorrelation of squares; kurtosis; robust procedures.

*Department of Statistics and Econometrics; Universidad Carlos III de Madrid, Getafe (Madrid), Carnero: e-mail: acarnero@est-econ.uc3m.es; Peña: e-mail: dpena@est-econ.uc3m.es; Ruiz: e-mail: ortega@est-econ.uc3m.es

Outliers and Conditional Autoregressive Heteroscedasticity in Time Series

M. Angeles Carnero, Daniel Peña and Esther Ruiz[*]

Departamento de Estadistica y Econometria

Universidad Carlos III de Madrid

C/ Madrid 126, 28903 Getafe (Spain)

Abstract

This paper reviews the literature on GARCH-type models proposed to represent the dynamic evolution of conditional variances. Effects of level outliers on the diagnostic and estimation of GARCH models are also studied. Both outliers and conditional heteroscedasticity can generate time series with excess kurtosis and autocorrelated squared observations. Consequently, both phenomena can be confused. However, since outliers are generated by unexpected events and the conditional variances are predictable, it is important to identify which one is producing the observed features in the data. We compare two alternative procedures for dealing with the simultaneous presence of outliers and conditional heteroscedasticity in time series. The first one is to clean the series of outliers before fitting a GARCH model. The second is to estimate first the GARCH model and then to clean of outliers by using the residuals adjusted by its conditional variance. It is shown that both approaches may result in different estimated conditional variances.

1 Introduction

This paper reviews the literature on Generalized Autoregressive Conditional Heteroscedasticity (GARCH)-type models proposed to represent the dynamic evolution of conditional variances often observed in real heteroscedastic time series. Since many real time series are affected by outliers, we also analyze how the presence of outliers may affect the diagnostic and modelling of conditional heteroscedasticity. The development of GARCH models has been mainly related with the empirical modelling of high frequency financial time series. Modelling volatility of returns is fundamental, for example, for option valuation and risk management; see, for example, Engle (2001). Financial series of returns are mainly characterized by having leptokurtic marginal distributions and volatility clustering. These properties have been documented as early as Mandelbrot (1963) and Fama (1965). Often, these series are not autocorrelated although they are not independent. The autocorrelation function (acf) of squared observations has a small first order autocorrelation coefficient followed by coefficients decaying very slowly towards zero; see, for example, Bollerslev and Engle (1993), Mills (1996) and Granger and Marmol (1998). Therefore, models representing the dynamic behavior of high frequency financial time series should be able to explain at least three properties: high kurtosis, small first order autocorrelation and high persistence in the autocorrelations of squares. Some review papers on these models are Bollerslev et al. (1992), Bollerslev et al. (1994), Bera and Higgins (1995), Diebold and López (1995), Pagan (1996) and Palm (1997). Engle (1995) is a survey of some of the main papers related with GARCH models and Campbell et al. (1997) provide an extensive textbook on this area. Bollerslev (2001) provides a very selective updated summary of the most influential developments in the area

Another important stylized fact of many financial series is the asymmetric response of volatility to positive and negative movements in stock prices. This is known as *leverage effect* and was originally described by Black (1976). This asymmetry has also been reported by Glosten et al. (1993), Schwert (1989), Nelson (1991), Campbell and Hentchel (1992), Engle and Ng (1993), Sentana (1995) and Shephard (1996) among others. In this paper, we will also describe models which are able to represent this asymmetry.

In order to illustrate the main empirical properties often observed in high frequency financial time series, table 1 contains descriptive statistics of twelve daily series. Denoting by p_t, the observed price at time t, the series of interest are the returns, defined as $r_t = 100(\log(p_t) - \log(p_{t-1}))$. The series described in table 1 are returns of the US Dollar against the Canadian Dollar, the Spanish Peseta, the German Mark, the Japanese Yen, the Swiss Franc,

2

the Swedish Krona and the British Pound observed from January 1993 to October 2000. Also we describe returns of five international stock market indexes, the Amsterdam E.O.E. index and the Bombay stock market index (from October 1995 to October 2000), the Dow Jones (from January 1990 to October 2000), the IBEX 35 of the Madrid Stock Exchange (from January 1992 to December 1999) and, finally, the S&P 500 index (from November 1987 to December 1998). Table 1 shows that most of these series have zero mean and all of them have excess kurtosis and negative skewness coefficients. Also, although the series are not autocorrelated, the squared observations have significant non null coefficients at low lags.

As an example, Figure 1 represents the returns of the S&P 500 index and the Dollar/Yen exchange rate. It is possible to observe volatility clustering with sequences of days of large returns in absolute value. Figure 1 also gives kernel estimates of the marginal densities of returns together with the corresponding normal density. These density plots confirm that the distributions of returns are heavy-tailed. Finally, the acf of the series y_t, y_t^2 and $|y_t|$ is also plotted in this figure. The acf of y_t does not have significant autocorrelations but the volatility clustering is reflected in the significant correlations of the transformed returns. In particular, in the acf of y_t^2 and $|y_t|$, the autocorrelations start at low values but are significant even for very large lags. This fact may suggest the presence of high persistence or long memory in the volatility process; see, for example Ding et al. (1993), Bollerslev and Mikkelsen (1996), Lobato and Savin (1998) and Lobato and Velasco (2000). Finally, Figure 1 illustrates what it is known as the "Taylor effect" that states that the absolute returns have the highest autocorrelations among all possible power transformations. High autocorrelations of absolute returns have also been found, for example, by Taylor (1986), Cao and Tsay (1992), Ding et al. (1993) and Granger and Ding (1995).

The simplest model to represent the empirical properties just described, specifies the series of interest as the product of two processes, ε_t and σ_t, that is

$$y_t = \varepsilon_t \sigma_t \qquad (1)$$

where ε_t is a serially independent and identically distributed (i.i.d.) white noise process with unit variance that is assumed to be independent of σ_t, which is known as volatility in the financial literature. GARCH models specify the volatility as a non-linear function of past returns. It is easy to show that if the conditional expectation of σ_t is finite, the process y_t in (1) is a martingale difference. Furthermore, model (1) can explain volatility clustering via autoregressive dynamics in the conditional expected value of

σ_t^2. Finally, y_t can have excess kurtosis either because ε_t has a leptokurtic distribution and/or because of the stochastic features of $\mathrm{E}(\sigma_t^2|Y_{t-1})$, where Y_{t-1} is the information set available at time $t-1$, i.e. $Y_{t-1} = \{y_1, y_2, ..., y_{t-1}\}$. Therefore, even if ε_t were a Gaussian process, the excess kurtosis observed in high frequency time series could be due to conditional heteroscedasticity.

However, it is well known that outliers may also cause excess kurtosis in time series and, when they appear in clusters, autocorrelations of squares. Thus, outliers effects can be confused with ARCH effects. Balke and Fomby (1994) analyze fifteen post World War II US macroeconomic time series and find that controlling for outliers eliminates much of the evidence of non-linearity in many of them. Once outliers are removed, there is no evidence of significant excess kurtosis or skewness in most of the series. They also test for GARCH in various series before and after controlling for outliers. They find that most of the raw series show evidence of either GARCH or non-linearity. After fitting the outlier model and controlling for the effects of outliers, the evidence of GARCH and non-linearity in many of the series is substantially weaker. The same result has been found by Fiorentini and Maravall (1996) analyzing monthly observations of the Spanish monetary aggregate known as Liquid Assets in the Hands of the Public.

On the other hand, if the series is truly heteroscedastic, the shape of the acf of squared observations can be distorted in the presence of outliers. Thus, outliers may hide genuine ARCH effects. Consequently, the presence of outliers in conditionally heteroscedastic time series may have effects on the estimates of the parameters of the equation governing the volatility dynamics. Finally, notice that conditional heteroscedasticity may generate what can be identified as outliers. Observations corresponding to periods when the conditional volatility is over the marginal standard deviation can be identified as outliers by traditional outlier detection methods. Fiorentini and Maravall (1996) also point out the possible confusion between conditional heteroscedasticity and outliers when looking at real data sets.

Since outliers are the result of non repetitive interventions, they are unpredictable given past information, while conditional heteroscedasticity generates volatility clustering and, therefore, can be predicted. Furthermore, conditional heteroscedasticity is related with uncertainty about the value of y_t and outliers are caused by unexpected events. Both phenomena have different interpretations and economic implications and, therefore, it is important to distinguish between them.

The paper is organized as follows. Section 2 reviews the growing literature on models for conditionally heteroscedastic time series. Although a wide spectrum of models has been proposed, we concentrate our attention on the GARCH class of parametric models and their ability to represent the three

main stylized facts that characterize high frequency financial time series. In this section we also describe briefly some asymmetric GARCH models and two alternatives to GARCH-type models proposed in the literature to represent the dynamic evolution of volatility. Section 3 deals with the effects of outliers on the diagnostic and estimation of GARCH models. We illustrate with real data the performance of two alternative strategies to deal with the simultaneous presence of conditional heteroscedasticity and outliers in time series. The first one consists of cleaning for outliers before fitting a GARCH model. In the second procedure, the GARCH model is estimated first and then, outliers are identified using the conditional variance. Finally, section 4 includes some concluding remarks.

2 Models for conditional heteroscedasticity

2.1 Symmetric ARCH models

The AutoRegressive Conditional Heteroscedasticity (ARCH) model was introduced by Engle (1982) to model the conditional variance of UK inflation. The ARCH(p) model allows the volatility, σ_t^2, to be a linear function of the squares of past observations. In the simplest case, the ARCH(1)[1] model, the series of interest, y_t, is given by

$$
\begin{aligned}
y_t &= \varepsilon_t \sigma_t \\
\sigma_t^2 &= \omega + \alpha y_{t-1}^2
\end{aligned}
\tag{2}
$$

where ε_t is a Gaussian white noise process with zero mean and unit variance, i.e., $\varepsilon_t \sim NID(0,1)$, and ω and α are parameters such that $\omega > 0$ and $0 \leq \alpha < 1$. The positivity conditions on the parameters ω and α are needed to guaranty the positivity of the conditional variance, and ω has to be strictly positive for the process y_t not to degenerate. Finally, $\alpha < 1$ is the covariance stationarity condition for y_t. However, Nelson (1990) shows that y_t is strictly stationary if $E\{\log(\alpha \varepsilon_t^2)\} < 0$. If ε_t is Gaussian, this condition is satisfied if $\alpha < 3.56$.

Notice that, once y_{t-1} is observed, σ_t^2 is known, and the conditional distribution of y_t is given by

$$
y_t \mid Y_{t-1} \sim N(0, \sigma_t^2).
\tag{3}
$$

---

[1]In this paper, we will focus on the simplest specification of each model considered given that they are the ones often used in practice.

It is easy to prove that y_t is a martingale difference process with marginal variance given by

$$\sigma_y^2 = \frac{\omega}{1 - \alpha}. \tag{4}$$

Assuming $3\alpha^2 < 1$, the kurtosis coefficient of y_t has the following expression

$$\kappa_y = \frac{E(y_t^4)}{[E(y_t^2)]^2} = \frac{3(1 - \alpha^2)}{1 - 3\alpha^2} \tag{5}$$

which is greater than 3. Therefore, the marginal distribution of y_t has fat tails even if its conditional distribution is normal. All the odd moments can be seen to be zero, so y_t has a symmetric marginal density.

The dynamics of the process y_t appear in the squared observations. Notice that a large y_{t-1}^2 tends to be followed by a large y_t^2 generating volatility clustering. The acf of y_t^2 is given by

$$\rho_2(\tau) = \alpha^\tau \tag{6}$$

The shape of the acf of y_t^2 in expression (6) mimics that of an AR(1) process. Therefore, the ARCH(1) model in (2) is able to generate volatility clustering. From (5) and (6), it is possible to write down the order one autocorrelation of squares in terms of the kurtosis of y_t as follows

$$\kappa_y = 3 \frac{1 - \rho_2(1)^2}{1 - 3\rho_2(1)^2} \tag{7}$$

Figure 2 plots this relationship and the observed sample values of the kurtosis and first order autocorrelation of the squared observations for the twelve series in table 1. It can be observed that for the values of the kurtosis often observed in real time series, the implied value of $\rho_2(1)$ is extremely higher than the sample values.

The early implementation of ARCH(p) models required a large number of past values of y_t^2 in the equation of σ_t^2, making these models difficult to handle in practice. Bollerslev (1986)[2] proposed a parsimonious model able to cope with the high persistence often observed in squared observations, the Generalized ARCH, or GARCH process.

---

[2]Taylor (1986) proposed the GARCH(1,1) model simultaneously.

2.2 GARCH models

The series y_t follows a GARCH(1,1) model if

$$
\begin{aligned}
y_t &= \varepsilon_t \sigma_t \\
\sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{8}
$$

where $\varepsilon_t \sim NID(0,1)$, and ω, α and β are parameters such that $\omega > 0$, $\alpha, \beta \geq 0$ and $\alpha + \beta < 1$. Once more, the positivity conditions are needed to guaranty the positivity of the conditional variance[3] and ω has to be strictly positive for the process y_t not to degenerate. Finally, $\alpha + \beta < 1$ is the covariance stationarity condition for y_t. Nelson (1990) shows that y_t is strictly stationary if $E[\log(\beta + \alpha \varepsilon_t^2)] < 1$. This condition is satisfied even if $\alpha + \beta = 1$. Therefore, it is interesting to note that when $\alpha + \beta = 1$, the GARCH(1,1) process is strictly stationary although, as we will show later, the marginal variance is not finite.

The conditional distribution of y_t is still given by (3). All GARCH processes are martingale differences and if $\alpha + \beta < 1$, y_t has finite variance. In this case, the marginal variance of y_t is given by

$$
\sigma_y^2 = \frac{\omega}{1 - \alpha - \beta}
\tag{9}
$$

The condition for the existence of the four order moment is $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$; see Bollerslev (1986). If this condition is satisfied, the kurtosis of y_t is given by

$$
\kappa_y = \frac{E(y_t^4)}{[E(y_t^2)]^2} = 3 + \frac{6\alpha^2}{1 - 3\alpha^2 - 2\alpha\beta - \beta^2}
\tag{10}
$$

which is greater than 3.

Alternatively, the GARCH(1,1) model can be written as a non-Gaussian ARMA(1,1) model in the squared observations given by

$$
y_t^2 = \omega + (\alpha + \beta)y_{t-1}^2 + \nu_t - \beta \nu_{t-1}
\tag{11}
$$

where ν_t is an uncorrelated process defined as $\nu_t = y_t^2 - \sigma_t^2$ which has zero mean, constant variance but it is conditionally heteroscedastic. In expression (11), it is possible to observe that the dynamic behavior of the GARCH(1,1)

---

[3]The positivity conditions of σ_t^2 for the general GARCH(p,q) model have been given by Nelson and Cao (1992).

process shows up in the acf of the squared observations. Bollerslev (1988) shows that the autocorrelations of y_t^2 are given by

$$
\begin{aligned}
\rho_2(1) &= \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} \\
\rho_2(\tau) &= (\alpha + \beta)^{\tau-1}\rho_2(1), \tau > 1
\end{aligned}
\tag{12}
$$

The acf of squares has the same pattern as an ARMA(1,1) process. Notice that the persistence of the volatility process depends on the value of $\alpha + \beta$. Figure 2 plots the relationship between kurtosis and $\rho_2(1)$, as given by Teräsvirta (1996), for two normal GARCH models with different persistence measured by $\alpha + \beta$. Such relationship can be easily obtained from (10) and (12). This figure shows how large values of the kurtosis coefficient and low values of $\rho_2(1)$ cannot exist simultaneously in conditionally normal GARCH models; see Teräsvirta (1996). Carnero et al. (2001b) show that the GARCH model is very rigid, because it can only generate high kurtosis and low order one autocorrelation of squares if $\alpha + \beta$ is close to one. Therefore, it could be expected that, in empirical applications, the estimates of $\alpha + \beta$ are very close to one, even if shocks to volatility are not persistent. However, GARCH models have been successfully fitted to high frequency financial time series by a large number of authors; see, for example, the references in Palm (1997).

Table 2 reports the Maximum Likelihood (ML) estimates of the parameters of the Normal GARCH model [4] for four of the series described in section 1: the US Dollar/Spanish Peseta and US Dollar/ Japanese Yen exchange rates and the Bombay and S&P 500 indexes. In this table it is possible to observe that all the series considered have significant ARCH effects and high persistence measured by $\widehat{\alpha} + \widehat{\beta}$. Model diagnostics are based on the standardized observations defined as $\widehat{\varepsilon}_t = y_t / \widehat{\sigma}_t$, where $\widehat{\sigma}_t$ is obtained substituting the estimated parameters in the corresponding expression of the conditional variance. The plots of $\widehat{\sigma}_t$, corresponding to two of the series, appear in Figure 3. In table 2, we also report several sample moments of $\widehat{\varepsilon}_t$. Notice that the standardized observations have still heavy tails. However, the autocorrelations of squares are not any longer significant. Therefore, for these series the GARCH(1,1) model is able to represent adequately the dynamics of squares although it is not able to explain the excess kurtosis present in the data. This could be due to an inadequate assumption on the distribution of ε_t and/or to the presence of outliers in the data.

Table 3 reports analytic values of the kurtosis and acf of squares implied by the GARCH models estimated for each of the four series analyzed in

---

[4]The estimation has been carried out with EViews, version 3.1.

this section, together with their sample moments. Notice that in the four series the theoretical kurtosis implied by the model is smaller than the sample kurtosis. On the other hand, the implied first order autocorrelations of squares are below the observed ones. Therefore, the results observed in table 2 about the moments of standardized residuals are confirmed. It seems that the GARCH(1,1) model cannot represent well the observed properties of these series. This could be due to the presence of outliers that may affect both, the properties of the correlogram of squares and of the estimates of the GARCH parameters. We will consider these effects latter on.

Finally, notice that GARCH models may generate what can be identified as outliers by traditional methods. Figure 4 shows four simulated GARCH(1,1) series with parameters (ω, α, β) equal to $(0.85, 0.15, 0)$, $(0.6, 0.4, 0)$, $(0.1, 0.1, 0.8)$ and $(0.1, 0.2, 0.7)$ respectively, with Normal conditional distribution. Notice that all the models have the same marginal variance, equal to one. As we can see, several observations are, in absolute value, greater than 3.5 times the standard deviation, which means that they would be considered outliers with respect to the Normal distribution. Notice also that, as expected, the bigger is α, the bigger is the number of observations greater than 3.5 standard deviations. It is important to note that these outlying observations appear in clusters and jumping from positive big values to negative big ones. Figure 4 also represents two series generated by the first and third models described above with a conditional Student-t distribution with 7 degrees of freedom. Notice that in these two series, the number of observations greater that 3.5 standard deviations is clearly increased with respect to the corresponding conditionally normal cases.

As we have seen in table 2, in many empirical studies, the estimates of α and β are such that $\widehat{\alpha} + \widehat{\beta} \simeq 1$, suggesting high persistence of shocks to volatility. Engle and Bollerslev (1986) proposed the Integrated GARCH (IGARCH) process given by model (8) with $\alpha + \beta = 1$. Remember that although the marginal variance of an IGARCH process is not finite, the IGARCH process with a normal conditional distribution is strictly stationary. Furthermore, Kleibergen and van Dijk (1993) show that the probability of an increase in the variance is smaller than the probability of a decrease and therefore, the dynamic behavior of IGARCH series is rather regular. Alternatively, the results of Teräsvirta (1996), suggest why, in practice, one may obtain estimates of the parameters such that $\widehat{\alpha} + \widehat{\beta} \simeq 1$ even when the volatility process is not persistent. As we mentioned before, the GARCH(1,1) model with normal errors cannot adequately characterized simultaneously the high kurtosis and the small first order autocorrelation of squared observations often observed in real series. Even IGARCH models are unlikely to provide an

9

adequate characterization of both stylized facts. Teräsvirta (1996) suggests that substituting the normal distribution of ε_t by a heavy-tailed distribution as, for example, the Student-t distribution, may improve the adequacy of the GARCH model to characterize the stylized facts observed in practice. Remember that in table 3, the observed kurtosis is clearly over the implied kurtosis. In Figure 2, we also represent curves with the relationship between kurtosis and $\rho_2(1)$ for the same GARCH(1,1) models as before but with ε_t having a Student-t distribution with 7 and 10 degrees of freedom. As we can see, the GARCH-t model seems to be better at explaining the simultaneous high kurtosis and low $\rho_2(1)$ than the Normal GARCH.

The Gaussian assumption on ε_t has been relaxed by several authors. For example, Bollerslev (1987) suggests a Student-t distribution, the normal-Poisson mixture distribution is used by Jorion (1988), the power exponential distribution in Baillie and Bollerslev (1989), the normal-lognormal mixture distribution in Hsieh (1989) and the Generalized error distribution (GED) in Nelson (1991). Bollerslev et al. (1994) used the Generalized-t distribution which includes both the Student-t and the GED distributions as particular cases. Finally, Granger and Ding (1995) and González-Rivera (1998) also consider the use of the Laplace distribution in conjunction with GARCH models.

The ML estimates of the parameters of the GARCH(1,1) model with Student-t errors adjusted to the four selected financial series [5] are, with the exception of Bombay index, very similar to the ones reported in table 2. However, the estimates obtained for Bombay are dramatically different. The α parameter is estimated as 0.1127 and the estimate of β is 0.8137. Notice that the persistence of Bombay volatility is smaller when the errors have a Student-t distribution instead of being Normal. In table 3, where we report the moments implied by the estimated GARCH-t models, it is possible to observe that, with the exception of the S&P 500 index, the implied kurtosis is clearly over the observed kurtosis. Therefore, it seems that when conditional Gaussianity is assumed, the implied kurtosis is too low but when the conditional distribution is a leptokurtic Student-t distribution, the implied kurtosis is too high.

In Figure 5, we plot the News Impact Curve proposed by Engle and Ng (1993) to measure the impact of shocks on the volatility. Holding constant the information up to and including time $t - 2$ and all the lagged conditional variances evaluated at the level of the unconditional variance, the News Impact Curve measures the implied relation between y_{t-1} and σ_t^2. For

---

[5]These estimates are not reported to save space but they are available from the authors upon request.

the GARCH(1,1) model in (8), the News Impact Curve has the following expression

$$\sigma_t^2 = A + \alpha y_{t-1}^2 \tag{13}$$

where $A = \omega + \beta\sigma_y^2$. As we can see, the News Impact Curve of the GARCH model is symmetric, which means that positive shocks have the same effect on the volatility as negative ones with the same absolute value. Notice that this contradicts one of the stylized facts of many financial series.

Given that the autocorrelations of absolute returns are higher than for squared returns, several authors have also proposed to model the conditional standard deviation instead of the conditional variance[6]. Modelling the absolute returns can be traced back to Taylor (1986) and Schwert (1989) who proposed the Absolute Value GARCH (AVGARCH) model given by

$$\sigma_t = \omega + \alpha\,|y_{t-1}| + \beta\sigma_{t-1} \tag{14}$$

Nelson and Foster (1994) demonstrate that the Taylor/Schwert GARCH model is a more efficient filter of the unconditional variance in the presence of leptokurtic error distribution than the specifications based on σ_t^2. He and Teräsvirta (1999) show that the autocorrelation function of squares for the AVGARCH model is radically different from that of the GARCH model. For the latter model, the acf decays exponentially whereas for the former, the rate of decay is slower than exponential. Although the kurtosis and first order autocorrelation of squared observations of the AVGARCH model are straightforward to obtain from He and Teräsvirta (1999), they have rather complicated expressions and we remit the interested lector to their paper.

Like in the GARCH(1,1) model, the News Impact Curve of the AV-GARCH is symmetric, as we can see in Figure 5, so this model does not allow for asymmetries in the volatility .

2.3   EGARCH models

Nelson (1991) points out three important limitations of GARCH processes. First of all, the non negativity constrains on the parameters are sometimes violated in empirical applications. Secondly, GARCH models are not able to represent the asymmetry of volatility responses to positive and negative shocks often observed in real time series. Finally, the interpretation of persistence in GARCH processes is not clear. To overcome these problems, Nelson (1991) proposes the Exponential GARCH (EGARCH) model. If $\varepsilon_t \sim NID(0,1)$, the simplest EGARCH(1,1) model is given by

---

[6]However, notice that He and Teräsvirta (1999) suggest that the Taylor effect may be due to the severe bias in the sample autocorrelations of squares.

$$y_t = \varepsilon_t \sigma_t \tag{15}$$
$$\log \sigma_t^2 = \omega + \beta \log \sigma_{t-1}^2 + \alpha[|y_{t-1}| - (2/\pi)^{1/2}] + \gamma y_{t-1}$$

where there is no need to restrict the parameters to guaranty the positivity of the conditional variance given that the model is formulated for the $\log \sigma_t^2$ process. The stationarity condition is $|\beta| < 1$. There is an asymmetric response of volatility to negative and positive returns. When $y_{t-1} > 0$, then $\partial \log \sigma_t^2 / \partial y_{t-1} = \gamma + \alpha$ while the derivative is $\gamma - \alpha$ when $y_{t-1} < 0$. In Figure 5 we can see the News Impact Curve of the EGARCH(1,1) model, which is given by

$$\sigma_t^2 = \begin{cases} A \exp \left(\frac{\gamma+\alpha}{\sigma_y} y_{t-1} \right) & \text{if } y_{t-1} > 0 \\ A \exp \left(\frac{\gamma-\alpha}{\sigma_y} y_{t-1} \right) & \text{if } y_{t-1} < 0 \end{cases} \tag{16}$$

where $A = \sigma_y^{2\beta} \exp \left(\omega - \alpha \sqrt{(2/\pi)} \right)$. Notice that in this case, where we have considered $\omega > 0$, $0 < \alpha < 1$, $0 < \beta < 1$ and, importantly, $\gamma < 0$, negative shocks have bigger effect on the volatility than positive ones.

The marginal variance, kurtosis and acf of squared observations of the EGARCH process in (15) were derived by He et al. (1999) and they are given by

$$\sigma_y^2 = \exp \left(\frac{\omega}{1-\beta} \right) \prod_{i=1}^{\infty} E \left(\exp \left(\beta^{i-1} g \right) \right) \tag{17}$$

$$\kappa_y = 3 \prod_{i=1}^{\infty} \frac{E \left(\exp \left(2\beta^{i-1} g \right) \right)}{[E \left(\exp \left(\beta^{i-1} g \right) \right)]^2} \tag{18}$$

and

$$\rho_2(\tau) = \frac{E \left(\varepsilon_{t-1}^2 \exp \left(\beta^{\tau-1} g \right) \right) P_1 P_2 - P_3}{3 P_4 - P_3} \tag{19}$$

where $g = g(y_{t-1}) = \alpha \left(|y_{t-1}| - \sqrt{2/\pi} \right) + \gamma y_{t-1}$, $P_1 = \prod_{i=1}^{\tau} E \left(\exp \left(\beta^{i-1} g \right) \right)$, $P_2 = \prod_{i=1}^{\infty} E \left(\exp \left((1 + \beta^{\tau}) \beta^{i-1} g \right) \right)$, $P_3 = \prod_{i=1}^{\infty} [E \left(\exp \left(2\beta^{i-1} g \right) \right)]^2$ and $P_4 = \prod_{i=1}^{\infty} E \left(\exp \left(2\beta^{i-1} g \right) \right)$. In particular, it is interesting to note that the acf of squared observations of EGARCH processes can be negative. Therefore, EGARCH models can produce cycles in the autocorrelation function of squares.

Figure 6 plots the relationship between kurtosis and $\rho_2(1)$ for EGARCH models with parameters $\beta = 0.99$ and 0.95, and $\gamma = -0.05$ together with the

sample values of the kurtosis and $\rho_2(1)$ for the series in table 1. It seems that the behavior of the EGARCH model in terms of the relationship between κ_y and $\rho_2(1)$ is not very different from the GARCH model. In any case, for a given value of the kurtosis, the first order autocorrelation of squares is even greater for an EGARCH than for a GARCH model with the same persistence.

EGARCH models have been fitted to real time series by Nelson (1991), Kearns and Pagan (1993), Poon and Taylor (1992), Zakoian (1994) and Chong et al. (1999) among others.

Table 4 shows the estimated EGARCH models for the four series considered in the previous subsection. Once more the persistence of shocks to volatility, measured by β, is estimated very close to one and the asymmetry parameter is significant for all the series considered, except for the US Dollar/ Spanish Peseta exchange rate. In table 3, where the moments implied by the estimated EGARCH models appear, it is possible to observe that the kurtosis and the first order autocorrelation of squares are similar to the ones implied by the corresponding GARCH models.

2.4   Other models for asymmetric conditional variances

Since the original proposal of Nelson (1991) and mainly due to the problems faced in the empirical fitting of EGARCH models, a huge number of models have been proposed to represent the asymmetric response of volatility. Among the most popular asymmetric models is the Asymmetric-Power ARCH (A-PARCH) model, proposed by Ding et al. (1993), that is able to unify seven ARCH-like models for power transformations of the conditional standard deviation. In particular, the A-PARCH model encompasses the GARCH, the AVGARCH, the GJR-GARCH of Glosten et al. (1993), the Threshold-ARCH (TARCH) of Zakoian (1994), the NARCH of Higgins and Bera (1992) and the log-ARCH of Geweke (1986) and Pantula (1986). The conditional variance in the simplest A-PARCH model is given by

$$\sigma_t^\delta = \omega + \alpha(|y_{t-1}| - \gamma y_{t-1})^\delta + \beta \sigma_{t-1}^\delta \tag{20}$$

Ding et al. (1993) and Granger and Ding (1995) argue that the parameter δ serves as a Box-Cox transformation of σ_t and it is necessary to adequately capture the dynamic characterization of volatility. The parameter γ allows σ_t to respond asymmetrically to positive and negative shocks. The statistical properties of the A-PARCH model have been addressed by He and Teräsvirta (1997) and Fornani and Mele (1997). The A-PARCH model has been fitted to returns of several Stock Markets by Brooks et al. (2000) and Paolella (2000).

Recently, there has been several new models proposed to nest most of the ARCH-type models previously described. First, Hentschel (1995) defines a parametric family of GARCH models that nets the EGARCH and A-PARCH models but not the GQARCH model of Sentana (1995). León and Mora (1999) apply the model proposed by Hentschel (1995) to daily returns of the IBEX-35 index of the Madrid Stock Exchange and conclude that models that focus on conditional standard deviation perform better than those that focus on conditional variances. They also find that the likelihood of models based on leptokurtic conditional distributions are higher than when the conditional distribution is assumed to be Normal. Finally, they show that the asymmetric behavior of the IBEX-35 returns is statistically significant. The asymmetric response of volatility and the leptokurtic conditional distribution have also been found by Blanco (2000) for the same variable.

Alternatively, Duan (1997) introduces the augmented GARCH model that is general enough to unify many of the main ARCH-like models in the literature. Loudon et al. (2000) document on an UK weighted stock index, observed daily from 1971 to 1997, the relative effectiveness of most of the major parametric ARCH models using the model proposed by Duan (1997). They find that the estimates for the ARCH parameters across all models are highly significant. They also find that volatility measures exhibit a high degree of persistence and asymmetry. However, standardized residuals are characterized by having substantial negative skewness and excess kurtosis, concluding that ARCH models with conditionally normal density functions are able to capture some, but not all, of the observed skewness and excess kurtosis, a fact already suggested by McCurdy and Morgan (1987), Milhoj (1987), Hsieh (1989) and Baillie and Bollerslev (1989).

Finally, He and Teräsvirta (1999) provide a unifying framework for considering the statistical properties of many GARCH models both symmetric and asymmetric and without making any particular assumption on the distribution of ε_t. They consider the following models: GARCH, AVARCH, GJR-GARCH, Nonlinear GARCH, which is a particular case of the A-PARCH, volatility switching GARCH of Fornani and Mele (1997), TGARCH, fourth-order nonlinear generalized moving-average conditional heteroscedasticty of Yang and Bewley (1995) and GQARCH. They do not include neither the EGARCH nor the A-PARCH models. For their family of GARCH models, they derive a general existence condition of any integer moment of absolute-valued observations as well as the moments themselves, and the acf of squared and absolute observations.

2.5 Alternative models for conditional heteroscedasticity

The literature on models for the dynamic evolution of the volatility, σ_t, is so extensive that we do not try to cover all proposed models. In this subsection, we briefly describe two models that can be interesting alternatives to the GARCH-type models.

The Conditional Heteroscedastic Autoregressive Moving Average (CHARMA) process was introduced by Tsay (1987) . A simple CHARMA model is given by,

$$\begin{aligned}
\phi(L)y_t &= \theta(L)a_t \\
\delta_t(L)a_t &= \eta_t
\end{aligned} \qquad (21)$$

where $\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \ldots - \phi_p L^p$ and $\theta(L) = 1 - \theta_1 L - \theta_2 L^2 - \ldots - \theta_q L^q$ are constant coefficient polynomials in L of degrees p and q respectively, L is the lag operator such that $L^j y_t = y_{t-j}$, η_t is a Gaussian white noise with variance σ_η^2 and $\delta_t(L) = 1 - \delta_{1t}L - \delta_{2t}L^2 - \ldots - \delta_{rt}L^r$ is a purely random coefficient polynomial in L of degree r. The random coefficient vector $\delta_t = (\delta_{1t}, \delta_{2t}, ..., \delta_{rt})'$ is a sequence of iid random vectors with zero mean and nonnegative definite covariance matrix Σ. In addition, δ_t is independent of η_t. The CHARMA model uses random coefficients to produce conditional heteroscedasticity. If, for example, p=q=0 and r=1, then the series y_t is given by

$$y_t = \delta_{1t}y_{t-1} + \eta_t$$

It is easy to prove that $\{y_t\}$ is uncorrelated, conditional heteroscedastic and with leptokurtic unconditional distribution. Tsay (1987) considers an application of the CHARMA models using their fatter-tailed property and employs the heteroscedastic structure as an alternative approach for handling outliers in time series analysis.

Another important alternative models to represent the dynamic evolution of volatilities are the Stochastic Volatility (SV) models, originally proposed by Taylor (1986). SV models assume that σ_t is a latent variable that usually follows an autoregressive process after being transformed into logarithms. Surveys on the properties of SV models are given by Taylor (1994), Ghysels et al. (1996) and Shephard (1996).

The simplest Autoregressive SV model of order 1, ARSV(1), is given by:

$$
\begin{aligned}
y_t &= \sigma_\star \varepsilon_t \sigma_t \\
\ln \sigma_t^2 &= \phi \ln \sigma_{t-1}^2 + \eta_t
\end{aligned}
\tag{22}
$$

where ε_t and η_t are assumed to be white noise processes mutually independent and normally distributed with zero mean and variances 1 and σ_η^2 respectively, $\sigma_\star$ is a scale factor that removes the necessity of including a constant term in the equation of $\ln \sigma_t^2$ and the restriction $|\phi| < 1$ guarantees the stationarity of y_t. Although the assumption of Gaussianity of η_t can seem *ad hoc* at first sight, Andersen et al. (1999) show that the daily log-volatility distribution of real financial series may be well approximated by a normal distribution. Notice that σ_η^2 is the variance of the volatility disturbance. When σ_η^2 is zero, the model in (22) is no longer identified. The ARSV model generates series with excess kurtosis and autocorrelated squared observations. The shape of the acf of squared observations is similar to that of an ARMA(1,1) model. Although the properties of the ARSV(1) and GARCH(1,1) models may seem very similar, Carnero et al. (2001b) show that SV models are more flexible than GARCH models to represent simultaneously the three properties characteristic of high frequency financial time series: high kurtosis, small order one autocorrelation and slow decay of the autocorrelation coefficients of squared observations. However, the estimation of SV models is not straightforward since they are not conditionally Gaussian even if ε_t is assumed to be Gaussian. Inference of ARSV models is usually based either on approximations or on numerically intensive methods.

3   Modelling conditional heteroscedastic time series in the presence of outliers

None of the empirical studies previously mentioned take into account that long real time series usually have outliers and these observations may affect both the correlogram of squares and the estimated model for the conditional variance. In this section, we deal with the simultaneous presence of outliers and conditional heteroscedasticity.

3.1   Types of Outliers

The study of outliers in time series has been mainly done in the context of linear ARMA models, where two main types of outliers can be considered: the Additive (AO) and the Innovative outlier (IO). These types of observations

16

were introduced by Fox (1972) and generalized later by Tsay (1988). Reviews on outliers in ARMA models can be found in Tolvi (2000) and Peña (2001). A linear ARMA(p,q) model is given by

$$\phi(L)y_t = \theta(L)a_t \tag{23}$$

where $\phi(L)$ and $\theta(L)$ are defined as in (21) with all their roots outside the unit circle and a_t is assumed to be NID$(0,\sigma_a^2)$. Alternatively, y_t may be expressed as the AR(∞) process $\pi(L)y_t = a_t$ where $\pi(L) = \phi(L)\theta(L)^{-1}$, or the MA($\infty$) process $y_t = \psi(L)a_t$ where $\psi(L) = \theta(L)\phi(L)^{-1}$.

In this context, an AO is related to an exogenous change that directly affects the series y_t. That is, instead of y_t, we observe a series z_t, which is contaminated at time τ by an outlier of size w_A, i.e.

$$z_t = y_t + w_A \mathrm{I}(t = \tau) = \begin{cases} y_t & \text{if} \quad t \neq \tau \\ y_t + w_A & \text{if} \quad t = \tau \end{cases} \tag{24}$$

An additive outlier only affects the level of the given observation at time τ and therefore, the model for the observed series is given by $z_t = w_A \mathrm{I}(t = \tau) + \psi(L)a_t$ or, equivalently,

$$\pi(L)\big(z_t - w_A \mathrm{I}(t = \tau)\big) = a_t.$$

The IO is possibly generated by an endogenous change in the time series, that is, the observed series is, in this case

$$z_t = \begin{cases} y_t & \text{if} \quad t < \tau \\ y_t + w_I \psi_j & \text{if} \quad t = \tau + j, \quad j > 0 \end{cases} \tag{25}$$

where ψ_j are the coefficients of the corresponding MA(∞) representation. An innovative outlier affects all the observations after time τ through the memory of the ARMA process. The model for the observed series is $z_t = \psi(L)\big(w_I \mathrm{I}(t = \tau) + a_t\big)$ or equivalently $\pi(L)z_t = w_I \mathrm{I}(t = \tau) + a_t$.

It is well known that outliers affect the autocorrelation structure of a time series and, therefore, they cause biases in the estimated autocorrelation coefficients depending on their number, size and position; see Chang et al. (1988) and Chan (1995). In particular, a large additive outlier will push all the autocorrelation coefficients toward zero. Since traditional ARMA model identification procedures are based on the estimated autocorrelations, outliers will have, then, important effects on identifying the corresponding ARMA(p,q) model; see, for example, Deutsch et al. (1990). Similarly, outliers bias the estimated ARMA model parameters. Least squares and maximum likelihood

methods are both sensitive to the presence of outliers, especially to AOs. It is also known that a single AO has a strong effect on the estimation of the AR(1) parameter pushing it towards zero as the size of the outlier goes to infinity. In the case of IOs, the effects are not so strong. This type of outliers produce a small effect on the autocorrelation and hence, on the parameter estimates.

With respect to outliers in nonlinear GARCH models, Hotta and Tsay (1998) introduce two types of outliers, the level outlier (LO), which affects just the level of the series and has no effect on the conditional variance and, the volatility outlier (VO), which affects both, the level and the variance of the series.

Let us consider a GARCH(1,1) uncorrelated time series, y_t. In this context, AO and IO coincide, since there is no structure in the mean, and we should only distinguish between LO and VO. The level outlier can be defined as follows,

$$
\begin{aligned}
z_t &= y_t + w_L \mathrm{I}(t = \tau) \\
\sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{26}
$$

Notice that the conditional variance depends on y_{t-1}^2, hence it is not affected by the outlier. The volatility outlier is given by

$$
\begin{aligned}
z_t &= y_t + w_V \mathrm{I}(t = \tau) \\
\sigma_t^2 &= \omega + \alpha z_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{27}
$$

Notice that in this case, the conditional variance σ_t^2 depends on z_{t-1}^2, so is affected by the outlier.

If both the conditional mean and the conditional variance evolve over time, there are three types of possible outliers: the AO as defined before, but now the IO can be a LO or a VO. In order to make this clear, let us consider, for example, the MA(1)-GARCH(1,1) model given by,

$$
\begin{aligned}
y_t &= (1 + \theta L) a_t \\
a_t &= \varepsilon_t \sigma_t \\
\sigma_t^2 &= \omega + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{28}
$$

If there is an AO, the observed series will be z_t as in (24), but if the outlier is an IO, the observed series could be

$$
\begin{aligned}
z_t &= (1 + \theta L) \tilde{a}_t \\
\tilde{a}_t &= \varepsilon_t \sigma_t + w_L \mathrm{I}(t = \tau) \\
\sigma_t^2 &= \omega + \alpha a_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{29}
$$

18

or

$$
\begin{aligned}
z_t &= (1 + \theta L)\tilde{a}_t \\
\tilde{a}_t &= \varepsilon_t \sigma_t + w_V \mathrm{I}(t = \tau) \\
\sigma_t^2 &= \omega + \alpha \tilde{a}_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{30}
$$

Since in this paper we focus on the analysis of financial time series, we will assume that y_t is an uncorrelated process, so additive and innovative outliers coincide. Furthermore, we will focus only on the effects of level outliers on the identification and estimation of GARCH models. Like innovative outliers in linear models, we expect that the effects of volatility outliers are not so strong as for level outliers.

3.2 Effects of outliers on identification of ARCH effects

The Lagrange Multiplier (LM) test statistic for ARCH effects proposed by Engle (1982) is given by TR^2, where T is the sample size and R^2 is the determination coefficient computed from the regression of the squared observations y_t^2 on a constant and p lagged values, $y_{t-1}^2, ..., y_{t-p}^2$. Under the null hypothesis of homoscedasticity, the test statistic is asymptotically distributed as a χ^2 variable with p degrees of freedom. It is also quite common to use the asymptotically equivalent portmanteau test proposed by McLeod and Li (1983) based on the analogue of the Box-Pierce statistic that uses autocorrelation coefficients of squared observations; see Granger and Teräsvirta (1993). The finite sample properties of this stastistic have been studied in Engle et al. (1985). Therefore, the correlogram of squared observations is one of the main tools used in practice to test for conditional heteroscedasticity in time series.

Van Dijk et al. (1999) show how the presence of level outliers can produce both spurious ARCH effects and hide true conditional heteroscedasticity. They propose a robust test which seems to work well in discriminating spurious ARCH effects due to consecutive additive outliers from true conditional heteroscedasticity. However, the power is smaller than the corresponding to the LM test.

Lumsdaine and Ng (1999) analyze the effects of a possibly misspecified conditional mean on the LM test for ARCH. They show that misspecification will lead to overrejection of the null hypothesis of conditional homoscedasticity and propose a robust test based on adding additional terms in the estimated model for the mean, in particular, additional lags of the variable being analyzed and functions of lagged recursive residuals. Analyzing

19

by means of Monte Carlo experiments the performance of the LM test for ARCH in the presence of three consecutive outliers, they conclude that the null hypothesis is rejected too frequently. They also find evidence that this effect is exacerbated by higher levels of persistence. The robust test they propose does not work properly in this case.

Ruiz et al. (2001) analyzing monthly series of inflation of the G7 countries and by means of simulations illustrate the problems raised by the simultaneous presence of outliers and conditional heteroscedasticity in time series. They show that the presence of outliers in conditional heteroscedastic series, generates a big first order autocorrelation of squares. The same result was found by Deutsch et al. (1990) in relation to the identification of linear ARMA models. They conclude that the presence of a single outlier in an AR model leads to the identification of a MA or ARMA model.

To illustrate the potential effects of outliers on the correlogram of squared observations, in table 5 we report sample moments of the twelve financial series described before, corrected by outliers. In this table, all observations bigger than 4 standard deviations have been substituted by the sample mean. Notice that the magnitude of the autocorrelations of squares and the Box-Ljung statistic for y_t^2 are reduced for most of the series and the reduction of the autocorrelations is specially remarkable for the order one autocorrelation. See, for example, the correlograms of squared observations for the US Dollar/ German Mark and the US Dollar/Japanese Yen exchange rates. However, there are three series, the Bombay, Dow-Jones and S&P 500 indexes, where potential outliers are hiding the dynamic structure in the squares. Consequently, the results in this table point out the necessity of dealing properly with the presence of outliers. They can hide dynamic structure of squares or imply autocorrelations of squares not due to conditional heteroscedasticity. However, notice that the series have been corrected by outliers defined with respect to the marginal variance. If these series are conditionally heteroscedastic, it is not clear that the corrected observations are truly outliers. Granger and Orr (1972), in an early paper, also pointed out the danger involved in correcting too many outliers.

In order to illustrate these effects, we have simulated three series of size $T = 500$. The first one is a Gaussian zero mean white noise with variance one, denoted by a_t. The second series, y_t, is generated by a GARCH(1,1) process with parameters $\omega = 0.1$, $\alpha = 0.1$ and $\beta = 0.8$ and the third one, x_t, is an EGARCH(1,1) with parameters $\omega = -0.001$, $\alpha = 0.07$, $\beta = 0.98$ and $\gamma = -0.0456$. We have contaminated the series a_t first with three consecutive outliers at observations $t = 200, 201$ and 202 and second, with three isolated outliers, at observations $t = 100, 200$ and 300, obtaining the contaminated series $a_t^\star$ and a'_t respectively. Series $y_t^\star$, y_t', $x_t^\star$ and x_t' have been

20

generated from y_t and x_t in the same way. All outliers have size w equal to five standard deviations. Table 6 reports the Monte Carlo results on several descriptive statistics based on 1000 replicates generated by each of the previously described processes. The effect of outliers on skewness and kurtosis is, as expected, the same regardless of whether the outliers are isolated or consecutive and if the original series is white noise, GARCH or EGARCH. Both coefficients are bigger in the contaminated series and the magnitude of the effect is also similar. However, the effect on the order one autocorrelation of squared observations depends on whether the outliers are consecutive and on whether the original series is white noise, GARCH or EGARCH. If the series is white noise, isolated outliers do not generate autocorrelations of squares and the LM test for heteroscedasticity has lower size than nominal. However, the presence of consecutive outliers generates a significant order one autocorrelation of squared observations. In this case, the LM test rejects the null hypothesis of homoscedasticity and, therefore, consecutive outliers can be confused with heteroscedasticity. On the other hand, when the series is conditionally heteroscedastic, consecutive outliers increment the order one autocorrelation while isolated outliers can hide conditional heteroscedasticity.

Notice that for EGARCH models, the LM test has very low power. When the series has no outliers, the test rejects the null hypothesis of homoscedasticity just 27.10% when the alternative is in fact true. When outliers are present, they affect the size and power of the test in the same way as before.

It is also important to notice that in the case of series with some structure in the mean, what we usually do is (i) to model the mean and (ii) to check homoscedasticity in the residuals. For example, in the case of a simple AR(2) model, if the series is contaminated with just one outlier, the residuals appear contaminated with 3 consecutive outliers, which leads us to the first case considered, the white noise series with 3 consecutive outliers.

3.3   Effects of outliers on estimation of ARCH models

At the moment, there are very few articles analyzing how the presence of outliers in time series with ARCH effects, affects the estimation of the parameters of the conditional variance equation. There are two main procedures to estimate these parameters. The simplest one is to estimate by Ordinary Least Squares (OLS) the parameters of ARCH(p) models, expressed as AR(p) models for y_t^2. The OLS estimator is not efficient and cannot be applied when the conditional variance is modelled as a GARCH process. On the other hand, the estimation of the parameters of GARCH models can be carried out by Quasi-Maximum Likelihood (QML) by maximizing the

Gaussian log-likelihood; see, for example, Bollerslev and Wooldridge (1992).

Carnero et al. (2001a) study the effects of level outliers on the OLS estimation of ARCH models, finding that a single outlier biases the estimation of the parameter ω of the ARCH(p) model toward ∞ and the α_i toward $-\frac{1}{T-2p}$, $\quad \forall i = 1, \ldots, p$ as the size of the outlier goes to ∞. When there are k consecutive outliers of the same size, ω is biased toward ∞ and the bias of the estimate of α_i depends on the number of outliers k. When k is big enough, persistence in the variance is estimated very close to one. These results are extended to the QML estimates of GARCH(1,1) models by means of simulations.

Verhoeven and McAleer (2000) study, empirically, the effects of outliers on the AR(1)-GARCH(1,1) process by analyzing 1000 trading days of five financial time series: S&P 00, Nikkei 225, HSI, British Pound - US Dollar spot exchange rate and the Gold Bullion spot rate. They find that outliers tend to dominate the QML estimates resulting in larger ARCH and smaller GARCH estimates and may give rise to spurious AR(1) and ARCH effects. They also find that outliers are frequently clustered and do not appear to be i.i.d. This fact could explain the biases found for the ARCH and GARCH parameters, α and β, toward one and zero respectively, since it could be due to consecutive outliers, as Carnero et al. (2001a) point out. Another possible explanation to this founding is that the outliers detected correspond to periods of high volatility, and considering those observations as outliers may bias the estimates of the conditional variance.

In order to illustrate the kind of biases outliers cause on the estimation of GARCH models, Figure 7 plots the results of a simple Monte Carlo study in which we have simulated 100 replicates of GARCH(1,1) series of sample size T=500, with parameters $\omega = 0.1$, $\alpha = 0.1$ and $\beta = 0.8$. In the first row of Figure 7, we can see estimates of the parameters for the original series and after correcting for observations bigger than 4 standard deviations. Most of the times, there are not such observations and then, the original and corrected series are the same, but when some observation is bigger than 4 standard deviations and the series is corrected, we can see that the estimates are different for the original and for the corrected series, resulting that in most of the corrected series, ω and α are estimated smaller and β is estimated bigger than in the original series. So, we have to be careful about correcting for outliers, because we can introduce important biases in the estimates. This could be the case in Verhoeven and McAleer (2000) and in the US Dollar/Japanese Yen exchange rate.

The second and third rows of Figure 7 show estimates for the original series and the contaminated ones. It is important to notice that for all the contaminated series, ω is overestimated without depending on whether the

22

series is contaminated with consecutive or isolated outliers. But in the case of estimates of α and β, the biases depend on the nature of outliers, if they are consecutive, α is overestimated and β is underestimated, while in the case of isolated outliers, the biases are not very clear, both parameters α and β can be overestimated or underestimated.

3.4 Alternative modelling strategies

In this section, we compare two alternative procedures for dealing with the simultaneous presence of outliers and conditional heteroscedasticity in time series. First, it is possible to correct the series for outliers using standard criteria and then estimate the conditional variance. Alternatively, it is possible to estimate initially a model for the conditional variance, and then obtain the 'conditional' outliers using the resulting estimated conditional standard deviations. There are two procedures to detect outliers in GARCH models based on these 'conditional' outliers. Hotta and Tsay (1998) propose two test statistics to detect outliers in ARCH and GARCH processes. They applied the proposed tests to simulated and real examples and conclude that the tests work well in both applications. Franses and Ghijsels (1999) and Franses and van Dijk (1999) proposed to apply the Chen and Liu (1993) method to correct for additive outliers in stock market returns, when GARCH models for these returns are used for forecasting volatility. Notice that, a third approach to deal with this problem is to estimate the conditional variance parameters using robust estimation methods as proposed by Sakata and White (1998).

We will compare the two alternative strategies by applying them to the four selected series. Using the first strategy Table 7 shows the estimated GARCH(1,1) parameters for the series corrected by outliers bigger than 4 standard deviations. These estimates are for the Dollar/Peseta exchange rate and the S&P 500 index similar to the ones previously obtained (see table 2). However, those corresponding to the Dollar/Yen exchange rate and the Bombay index are quite different. For the Dollar/Yen exchange rate, the α parameter is estimated smaller and the estimate of β bigger, after correcting by outliers. For the Bombay index the result is the opposite. This could indicate that in the case of the Dollar/Yen exchange rate we are correcting observations which are not outliers and, consequently, pushing the estimates of the conditional variance towards the homoscedastic case, which is $\alpha = 0$ and $\beta = 1$. However, the Bombay index seems to have outliers and after controlling for them the dynamics on the squares appears more clear. It is also important to note that the estimated parameters obtained after correcting for outliers are similar to the ones obtained when fitting the GARCH model with a conditional Student-t distribution. Therefore, this

result suggests that the lack of fit when conditional normality is assumed could be due to the presence of outliers. Figure 3 plots estimates of the volatility after correcting for outliers bigger than 4 standard deviations. As we can see in this plot, there are important differences between estimates of the volatility before and after correcting for outliers. For the Dollar/Yen exchange rate, correcting for outliers makes the estimated volatility smoother while for the Bombay index, the estimated volatility without correcting for outliers is smoother.

Now we apply the second strategy, that is, we estimate first the conditional variance and then using the standardized observations look for outliers. The series is corrected by these 'conditional' outliers and then new estimates are computed. In this way we obtain the GARCH estimates in Table 8. Estimates of the volatility based on the new estimations are also plotted in Figure 3. For the Bombay index, as we saw before, there are differences between estimated volatilities after and before correcting for outliers but it seems that correcting for marginal or conditional outliers lead us to similar estimates. For the Dollar/Yen series the differences appear when we correct for marginal outliers. Table 9 indicates which observations are detected as outliers. As we can see, observations detected as outliers using the the marginal variance are not the same as the ones detected using the conditional variance.

Finally, in order to illustrate the effects that outliers may have on the estimation of the asymmetry response of volatility to negative and positive shocks, we fit EGARCH models to the series corrected by outliers bigger than 4 standard deviations[7] and we observe that the estimated asymmetric parameter, γ, is smaller in absolute value, for the four series considered, although it is still significant for all the series, except the US Dollar/Spanish Peseta exchange rate. Notice that the effect of outliers on the estimated parameter of asymmetry may depend on the sign of the outlier.

4   Conclusions

We have seen that financial time series have high kurtosis and correlations in the squared observations. This features can be explained by ARCH and GARCH models, although we have seen that although these models are able to capture some of these features they do not represent well many observed time series. This result may be due to the presence of outliers that can produce also high kurtosis and correlations in the squared observations. A key problem is to distinguish both effects. We have compared two alternative

---

[7]These estimations are available from authors upon request.

strategies for dealing with the simultaneous presence of outliers and conditional heteroscedasticity in time series. The first one is to correct the series for outliers using standard criteria and then estimate the conditional variance. The second one is to estimate the conditional variance and then correct the series for 'conditional' outliers. We have shown that both approaches may result in different estimated conditional variances. An important area of research is to compare the relative advantages of both procedures in practical problems.

It would be important to analyze the effects of other types of outliers, in particular level shifts and variance changes. For example, Tsay (1988) shows, analyzing a real time series, that if a variance change is ignored, more than 15 outliers are identified. However, when the variance change is taken into account, the series seems to have only two outliers. On the other hand, Lamoureux and Lastrapes (1990) show that variance changes can also be confused with highly persistent conditional heteroscedasticity. Thus deriving procedures to deal with these problems seems to be a promising line of future research.

References

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (1999). The Distribution of Exchange Rate Volatility. W.P. 99-08 and 6961, Wharton Financial Institutions Center and NBER.

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

Balke, N.S. and T.B. Fomby (1994). Large Shocks, Small Shocks and Economic Fluctuations: Outliers in Macroeconomic Time Series. *Journal of Applied Econometrics*, **9**, 181–200.

Bera, A.K. and M.L. Higgins (1995). On ARCH Models: Properties, Estimation and Testing. In: *Surveys in Econometrics*, L. Oxley, D.A.R. George, C.J. Roberts and S. Sayer (eds). Oxford:Blackwell.

Black, F. (1976). Studies of Stock Price Volatility Changes. *Proceedings of the 1976 Meetings of the Business and Economic Statistics, American Statistical Society*, 177–181.

Blanco, R. (2000). Efectos sobre la Volatilidad del Mercado Bursátil de la Introducción de los Contratos de Futuros y Opciones sobre el ibex35. *Investigaciones Económicas*, **24**, 139–175.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Bollerslev, T. (1987). A Conditionally Heteroskedasticity Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, **69**, 542–547.

Bollerslev, T. (1988). On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process. *Journal of Time Series Analysis*, **9**, 121–131.

Bollerslev, T. (2001). Financial Econometrics: Past Developments and Future Challenges. *Journal of Econometrics*, **100**, 41–51.

Bollerslev, T., R.Y. Chou, and K.F. Kroner (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 498–505.

Bollerslev, T. and R.F. Engle (1993). Common Persistence in Conditional Variances. *Econometrica*, **61**, 166–187.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. *The Handbook of Econometrics*, **4**, 2959–3038.

Bollerslev, T. and H.O. Mikkelsen (1996). Modeling and Pricing Long-Memory in Stock Market Volatility. *Journal of Econometrics*, **73**, 151–184.

Bollerslev, T. and Wooldridge (1992). Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances. *Econometric Reviews*, **11**, 143–172.

Brooks, R.D., R.W. Faff, M.D. McKenzie, and H. Mitchell (2000). A Multi-Country Study of Power ARCH Models and National Stock Market Returns. *Journal of International Money and Finance*, **19**, 377–397.

Campbell, J.Y. and L. Hentchel (1992). No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics*, **31**, 281–318.

Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997). *The Economics of Finantial Markets*. Princeton: Princeton University Press.

Cao, C.Q. and R.S. Tsay (1992). Nonlinear Time Series Analysis of Stock Volatilities. *Journal of Applied Econometrics*, **7**, 165–185.

Carnero, M.A., D. Peña, and E. Ruiz (2001a). Effects of Level Outliers on the Estimation of ARCH Models. Manuscript.

Carnero, M.A., D. Peña, and E. Ruiz (2001b). Is Stochastic Volatility more flexible than GARCH? W.P. 01-08, Universidad Carlos III de Madrid.

Chan, W. (1995). Understanding the Effect of Time Series Outliers on Sample Autocorrelations. *Test*, **4**, 179–186.

Chang, I., G.C. Tiao, and C. Chen (1988). Estimation of Time Series Parameters in the Presence of Outliers. *Technometrics*, **30**, 193–204.

Chen, C. and J. Liu (1993). Joint Estimation of Model Parameters and Outlier Effects in Time Series. *Journal of the American Statistical Association*, **88**, 284–296.

Chong, C.W., M.I. Ahmad, and M.Y. Abdullah (1999). Performace of GARCH Models in Forecasting Stock Market Volatility. *Journal of Forecasting*, **18**, 333–343.

Deutsch, S.J., J.E. Richards, and J.J. Savin (1990). Effects of a Single Outlier on ARMA Identification. *Communications in Statistics, Theory and Methods*, **19**, 2207–2227.

Diebold, F.X. and J.A. López (1995). ARCH Models. In: *Macroeconomics: Developments, Tensions and Prospects* , K. Hoover (eds). Oxford:Blackwell.

Ding, Z., C.W.J. Granger, and R.F. Engle (1993). A Long Memory Property of Stock Market Returns and a New Model. *Journal of Empirical Finance*, **1**, 83–106.

Duan, J.C. (1997). Augmented GARCH(p,q) Process and its Diffusion Limit. *Journal of Econometrics*, **79**, 97–127.

Engle, R.F. (1995). *ARCH*. Oxford Universtity Press.

Engle, R.F., D.F. Hendry, and D. Trumble (1985). Small Sample Properties of ARCH Estimators and Tests. *Canadian Journal of Economics*, **18**, 66–93.

Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**, 987–1007.

Engle, R. F. (2001). Financial Econometrics - A New Discipline with New Methods. *Journal of Econometrics*, **100**, 53–56.

Engle, R. F. and T. Bollerslev (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 1–50.

Engle, R. F. and V. K. Ng (1993). Measuring and Testing the Impact of News on Volatility. *Journal of Finance*, **48**, 1749–1778.

Fama, E. (1965). The Behaviour of Stock Prices. *Journal of Business*, **38**, 34–105.

Fiorentini, G. and A. Maravall (1996). Unobserved Components in ARCH Models: An Application to Seasonal Adjustment. *Journal of Forecasting*, **15**, 175–201.

Fornani, F. and A. Mele (1997). Weak Convergence and Distributional Assumptions for a General Class of Nonlinear ARCH models. *Econometric Reviews*, **16**, 205–227.

Franses, P.H. and H. Ghijsels (1999). Additive Outliers, GARCH and Forecasting Volatility. *International Journal of Forecasting*, **15**, 1–9.

Franses, P.H. and D. van Dijk (1999). Outlier Detection in the GARCH(1,1) Model. W.P. EI-9926/A, Econometric Institute, Erasmus University Rotterdam.

Geweke, J. (1986). Comment: Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 57–61.

Ghysels, E., A. Harvey, and E. Renault (1996). Stochastic Volatility. *Handbook of Statistics*, **14**.

Glosten, L.R., R. Jagannathan, and D.E. Runkle (1993). On the Relation between Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, **48**, 1779–1801.

González-Rivera, G. (1998). Smooth Transition GARCH Models. *Studies in Nonlinear Dynamics and Econometrics*, **3**, 61–78.

Granger, C.W.J. and Z. Ding (1995). Some Properties of Absolute Return, an Alternative Measure of Risk. *Annales d'Economie et de Statistique*, **40**, 67–90.

Granger, C.W.J. and F. Marmol (1998). The Correlogram of a Long Memory Process Plus a Simple Noise. W.P. 98-14, Universidad Carlos III de Madrid.

Granger, C.W.J. and D. Orr (1972). Infinite Variance and Research Strategy in Time Series Analysis. *Journal of the American Statistical Association*, **67**, 275–285.

Granger, C.W.J. and T. Teräsvirta (1993). *Modelling Nonlinear Economic Relationships*. New York: Oxford University Press.

He, C. and T. Teräsvirta (1997). Statistical Properties of the Asymmetric Power ARCH Process. W.P. 199, Stockholm School of Economics.

He, C. and T. Teräsvirta (1999). Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics*, **92**, 173–192.

He, C., T. Teräsvirta, and H. Malmsten (1999). Fourth Moment Structure of a Family of First-Order Exponential GARCH Models. W.P. 345, Stockholm School of Economics.

Hentschel, L. (1995). All in the Family: Nesting Symmetric and Asymmetric GARCH Models. *Journal of Financial Economics*, **39**, 71–104.

Higgins, M.L. and A.K. Bera (1992). A Class of Nonlinear ARCH Models. *International Economic Review*, **33**, 137–158.

Hotta, L. K. and R. S. Tsay (1998). Outliers in GARCH Processes. Manuscript.

Hsieh, D.A. (1989). Modelling Heteroskedasticity in Daily Foreign Exchange Rates. *Journal of Business and Economic Statistics*, **7**, 307–317.

Jorion, P. (1988). On Jump Processes in the Foreign Exchange and Stock Markets. *Review of Financial Studies*, **1**, 427–445.

Kearns, P. and A.R. Pagan (1993). Australian Stock Market Volatility. *The Economic Record*, **69**, 163–178.

Kleibergen, F. and H.K. van Dijk (1993). Non-stationarity in GARCH Models: A Bayesian Analysis. *Journal of Applied Econometrics*, **8**, 41–61.

Lamoureux, C.G. and W.D. Lastrapes (1990). Persistence in Variance, Structural Change, and the GARCH Model. *Journal of Business and Economic Statistics*, **8**, 225–234.

León, A. and J. Mora (1999). Modelling Conditional Heteroskedasticity: Application to the IBEX-35 Stock-Return Index. *Spanish Economic Review*, **1**, 215–238.

Lobato, I. and N.E. Savin (1998). Real and Spurious Long-Memory Properties of Stock-Market Data. *Journal of Business and Economic Statistics*, **16**, 475–495.

Lobato, I. and C. Velasco (2000). Long Memory in Stock-Market Trading Volume. *Journal of Business and Economic Statistics*, **18**, 410–427.

Loudon, G.F., W.H. Watt, and P.K. Yadav (2000). An Empirical Analysis of Alternative Parametric ARCH Models. *Journal of Applied Econometrics*, **15**, 117–136.

Lumsdaine, R.L. and S. Ng (1999). Testing for ARCH in the Presence of a Possibly Misspecified Conditional Mean. *Journal of Econometrics*, **93**, 257–279.

Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, **36**, 394–419.

McCurdy, T.H. and I. Morgan (1987). Tests of the Martingale Hypothesis for Foreign Currency Futures with Time Varying Volatility. *International Journal of Forescasting*, **3**, 131–148.

McLeod, A.J. and W.K. Li (1983). Diagnostic Checking ARMA Time Series Models Using Squared-Residual Correlations. *Journal of Time Series Analysis*, **4**, 269–273.

Milhoj, A. (1987). A Conditional Variance Model for Daily Observations of an Exchange Rate. *Journal of Business and Economic Statistics*, **5**, 99–103.

Mills, T.C. (1996). Non-Linear Forecasting of Financial Time Series: An Overview and some New Models. *Journal of Forecasting*, **15**, 127–135.

Nelson, D. B. (1990). Stationary and Persistence in the GARCH(1,1) Model. *Econometric Theory*, **6**, 318–344.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347–370.

Nelson, D. B. and C.Q. Cao (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics*, **10**, 229–235.

Nelson, D. B. and D.P. Foster (1994). Asymptotic Filtering Theory for Univariate ARCH Models. *Econometrica*, **62**, 1–41.

Pagan, A. (1996). The Econometrics of Financial Markets. *Journal of Empirical Finance*, **3**, 15–102.

Palm, F.C. (1997). GARCH Models of Volatility. *Handbook of Statistics*, **14**.

Pantula, S.G. (1986). Comment: Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 71–73.

Paolella, M.S. (2000). Generalized GARCH Models with Nonnormal Distributions. Manuscript, Institute of Statistics and Econometrics, Kiel University.

Peña, D. (2001). Outliers, Influential Observations and Missing data. In: *A course in time series*, Peña, D., G.C. Tiao and R.S. Tsay (eds). John Wiley, New York.

Poon, S.H. and S.J. Taylor (1992). Stock Returns and Volatility: An Empirical Study of the UK Stock Market. *Journal of Banking and Finance*, **16**, 37–59.

Ruiz, E., F. Lorenzo, and M.A. Carnero (2001). Heteroscedasticity and Outliers in Time Series. Manuscript.

Sakata, S. and H. White (1998). High Breakdown Point Conditional Dispersion Estimation with Application to S&P 500 Daily Returns Volatility. *Econometrica*, **66**, 529–567.

Schwert, G.W. (1989). Why does Stock Market Volatility change over Time? *Journal of Finance*, **45**, 1129–1155.

Sentana, E. (1995). Quadratic ARCH Models. *Review of Economic Studies*, **62**, 639–661.

Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility. In: *Time Series Models in Econometrics, Finance and Other Fields*, D.R. Cox, D.V. Hinkley and O.E. Bardorff-Nielsen (eds). Chapman & Hall, London.

Taylor, S.J. (1986). *Modelling Financial Time Series*. John Wiley.

Taylor, S.J. (1994). Modelling Stochastic Volatility: A Review and Comparative Study. *Mathematical Finance*, **4**, 183–204.

Teräsvirta, T. (1996). Two Stylized Facts and the GARCH(1,1) Model. W.P. Series in Finance and Economics 96, Stockholm School of Economics.

Tolvi, J. (2000). Outliers in Time Series: A Review. Manuscript.

Tsay, R. S. (1987). Conditional Heteroscedastic Time Series Models. *Journal of the American Statistical Association*, **82**, 590–604.

Tsay, R. S. (1988). Outliers, Level Shifts and Variance Changes in Time Series. *Journal of Forecasting*, **7**, 1–20.

Van Dijk, D., P. H. Frances, and A. Lucas (1999). Testing for ARCH in the Presence of Additive Outliers. *Journal of Applied Econometrics*, **14**, 539–562.

Verhoeven, P. and M. McAleer (2000). Modelling Outliers and Extreme Observatios for ARMA-GARCH Processes. Manuscript.

Yang, M. and R. Bewley (1995). Moving Average Conditional Heteroskedastic Processes. *Economics Letters*, **49**, 367–372.

Zakoian, J.M. (1994). Threshold Heteroskedastic Models. *Journal of Economic Dynamics and Control*, **18**, 931–955.

Figures and Tables

Figure 1: S&P 500 and US Dollar/Japanese Yen exchange rate
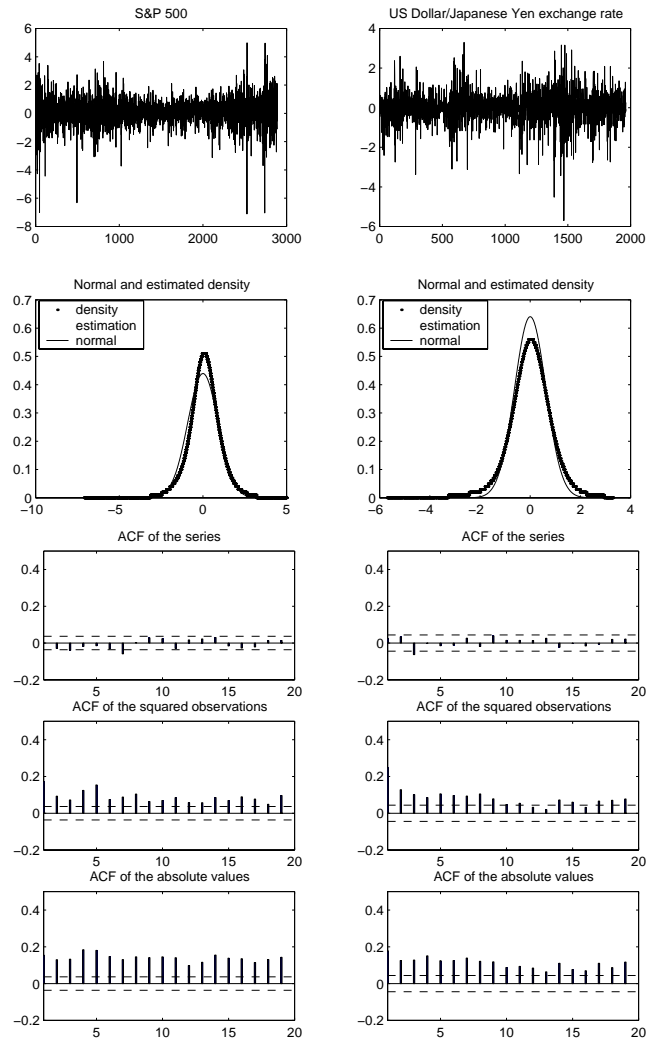
Figure 2: Relationship between κ_y and $\rho_2(1)$ for symmetric GARCH models
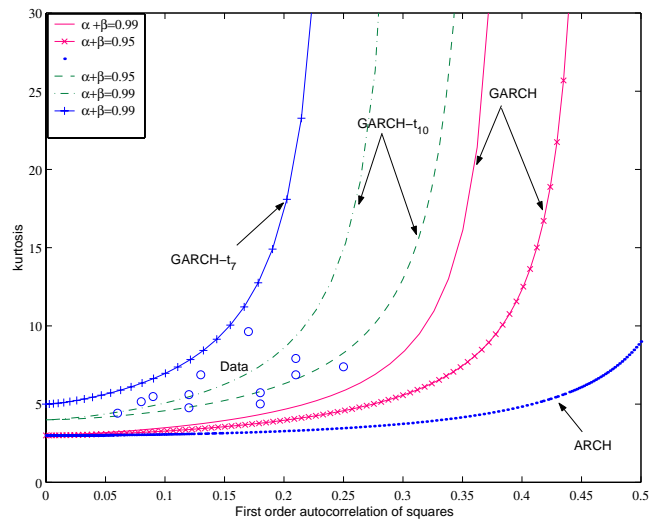
Figure 3: Estimated volatilities for daily returns

Figure 4: Simulated ARCH type series

Normal ARCH(1) with α=0.15   Normal ARCH(1) with α=0.4

Normal GARCH(1,1) with α=0.1   Normal GARCH(1,1) with α=0.2

ARCH(1)–t$_7$ GARCH(1,1)–t$_7$

Figure 5: News Impact Curve of GARCH, AVGARCH and EGARCH models
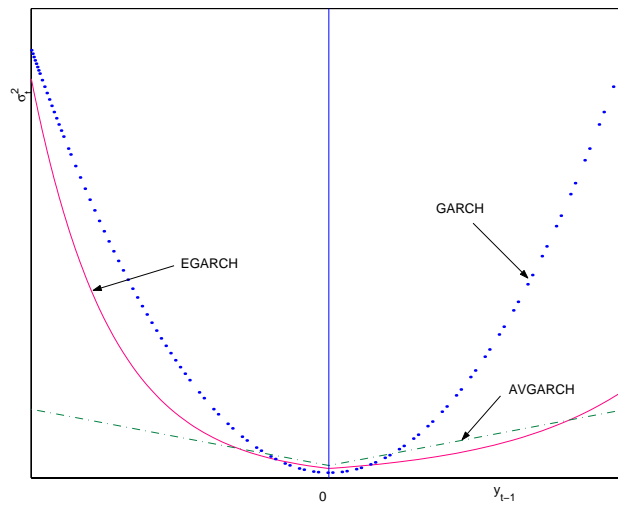
Figure 6: Relationship between κ_y and $\rho_2(1)$ for GARCH and EGARCH models
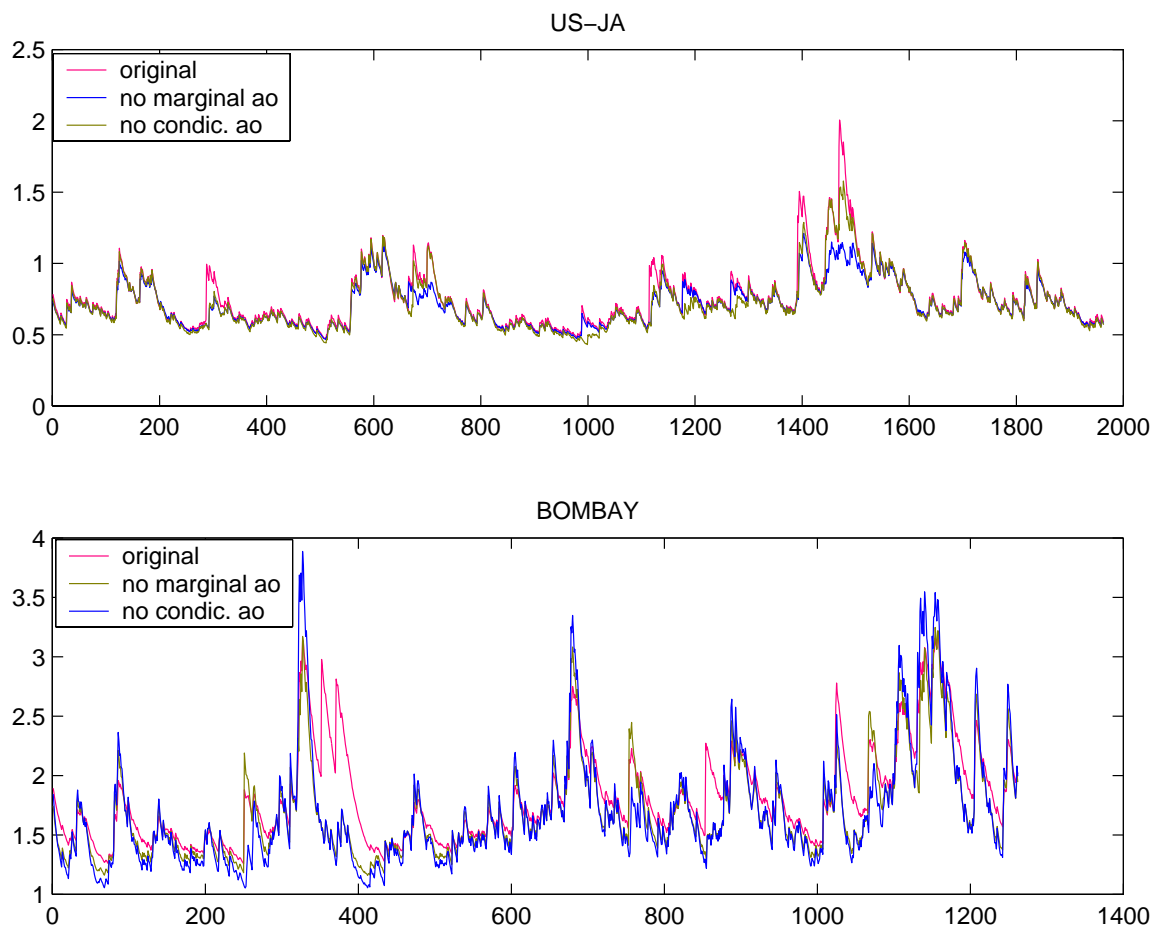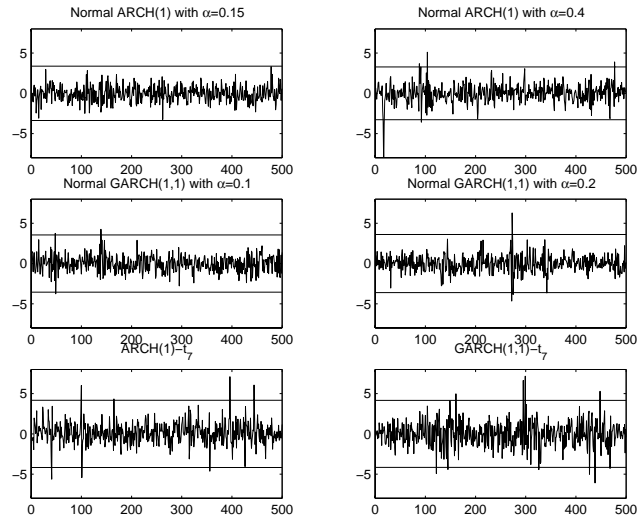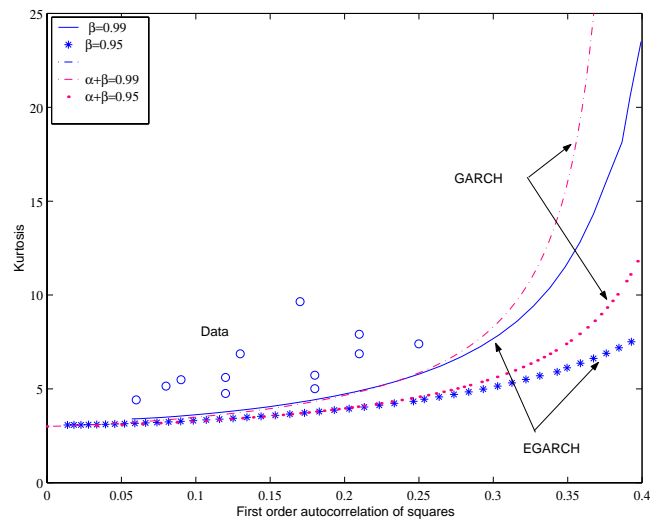
Figure 7: GARCH(1,1) estimates based on 100 simulated series
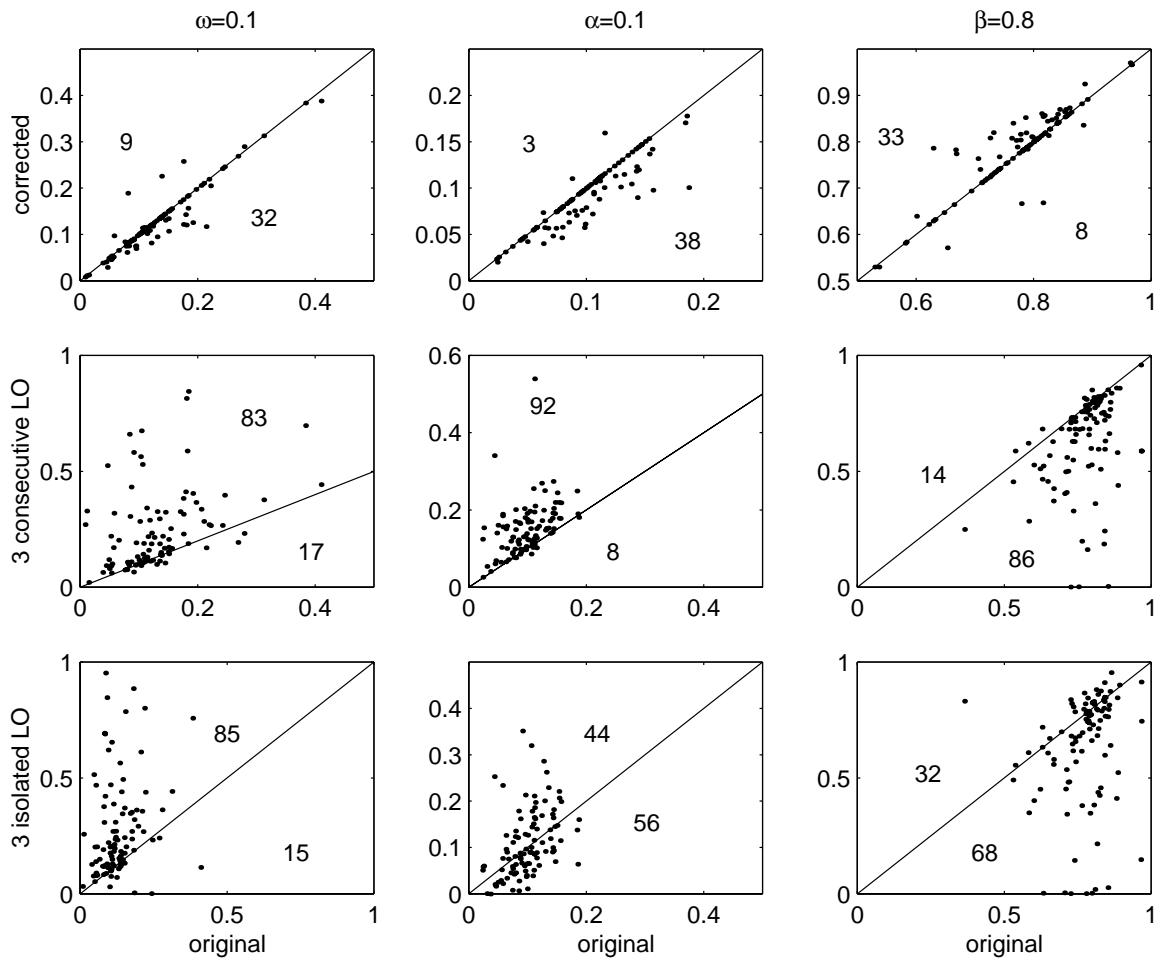
Table 1: Descriptive statistics of daily returns

	US CA	US ES	US GE	US JA	US SF	US SW	US UK	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
T	1963	1963	1963	1963	1963	1963	1963	1261	1262	2728	1982	2888
Mean	0.0084	0.0245	0.0153	-0.0077	0.0078	0.0149	0.0008	0.0818*	0.0111	0.0472*	0.0708*	0.0575*
S.D.	0.3094	0.6233	0.6129	0.7796	0.6860	0.6267	0.5090	1.3468	1.8471	0.9312	1.2495	0.9085
Skew.	-0.1449*	-0.1509*	-0.2553*	-0.5858*	-0.3539*	-0.1535*	-0.0266	-0.3101*	-0.1329*	-0.4303*	-0.3540*	-0.6108*
κ_y	5.6080*	5.1553*	4.7606*	7.3970*	5.0157*	4.4174*	5.4891*	5.7362*	6.8799*	7.9133*	6.8751*	9.6405*
$r(1)$	0.0647*	-0.0341	0.0262	0.0255	0.0278	0.0050	0.0100	0.0300	0.0700*	0.0300	0.1219*	0.0001
$Q(20)$	34.8*	24.2	20.7	23.6	18.7	20.07	32.7*	39.7*	41.3*	47.5*	72.8*	38.3*
Autocorrelations of y_t^2												
$r_2(1)$	0.1200*	0.0800*	0.1200*	0.2500*	0.1800*	0.0600*	0.0900*	0.1800*	0.1300*	0.2100*	0.2100*	0.1700*
$r_2(2)$	0.0900*	0.0500*	0.0300	0.1300*	0.0300	0.0800*	0.0800*	0.2700*	0.1300*	0.1400*	0.2000*	0.0900*
$r_2(5)$	0.1400*	0.0300	0.0600*	0.1100*	0.0600*	0.0300	0.0900*	0.1300*	0.1000*	0.1600*	0.1800*	0.1500*
$r_2(10)$	0.0800*	0.0400	0.0900*	0.0500	0.0600*	0.0400	0.0700*	0.1600*	0.0400	0.0600*	0.2300*	0.0700*
$Q_2(20)$	380*	96.7*	161*	377*	220*	141*	256*	749*	146*	511*	1414*	497*

T: Sample size.

κ_y: kurtosis of y_t.

$r(\tau)$: Autocorrelation of order τ of the original observations y_t.

$r_2(\tau)$: Autocorrelation of order τ of the squared observations y_t^2.

$Q(20)$ and $Q_2(20)$: Box-Ljung statistic for y_t and y_t^2 respectively (31.4 is the 5% critical value).

* Significant at the 5% level.

38

Table 2: Estimated GARCH models and diagnostics for the original series

y_t	US-ES	S&P 500	US-JA	BOMBAY
ω	0.0045	0.0040	0.0130	0.0718
	(0.0015)	(0.0008)	(0.0019)	(0.0106)
α	0.0327	0.0316	0.0559	0.0431
	(0.0058)	(0.0026)	(0.0060)	(0.0059)
β	0.9563	0.9633	0.9230	0.9369
	(0.0086)	(0.0030)	(0.0075)	(0.0063)
$\alpha + \beta$	0.9890	0.9949	0.9789	0.9800
log L	-1808.812	-3543.487	-2162.498	-2498.162
$\widehat{\varepsilon}_t = \frac{y_t}{\widehat{\sigma}_t}$				
Mean	0.0386	0.0646*	-0.0095	0.0072
S.D.	1.0002	1.0077	0.9959	0.9942
Skewness	-0.1371*	-0.7483*	-0.5439*	-0.0852
Kurtosis	5.1636*	8.1309*	6.0109*	6.8879*
$r(1)$	-0.0300	0.0300	0.0300	0.0900*
$Q(20)$	18.1	27.7	19.6	42.5*
$r_2(1)$	0.0100	0.0300	0.0300	0.0400
$r_2(2)$	-0.0200	0.0200	0.0000	0.0100
$r_2(5)$	-0.0200	0.0100	-0.0100	0.0200
$r_2(10)$	0.0000	0.0000	-0.0300	-0.0100
$Q_2(20)$	7.4	8.6	13.3	30.3

Table 3: Sample moments and implied ones by the model

	In sample	Implied by GARCH	Implied by GARCH-t	Implied by EGARCH
US-ES				
Mean	0.0245	0.0000	0.0000	0.0000
Variance	0.3885	0.4091	0.6429	0.5009
Kurtosis	5.1553	3.3250	23.9662	3.2475
y_t^2				
$\rho_2(1)$	0.0800	0.0773	0.1025	0.0645
$\rho_2(2)$	0.0500	0.0764	0.1018	0.0636
$\rho_2(5)$	0.0300	0.0739	0.0997	0.0612
$\rho_2(10)$	0.0400	0.0699	0.0962	0.0574
S&P 500				
Mean	0.0575	0.0000	0.0000	0.0000
Variance	0.8253	0.7933	0.5455	1.8048
Kurtosis	9.6405	3.7415	9.0915	4.0078
y_t^2				
$\rho_2(1)$	0.1700	0.1194	0.0934	0.1456
$\rho_2(2)$	0.0900	0.1188	0.0927	0.1438
$\rho_2(5)$	0.1500	0.1168	0.0906	0.1385
$\rho_2(10)$	0.0700	0.1124	0.0872	0.1302
US-JA				
Mean	-0.0077	0.0000	0.0000	0.0000
Variance	0.6078	0.6182	0.5143	0.7546
Kurtosis	7.3970	3.5306	22.3612	3.5132
y_t^2				
$\rho_2(1)$	0.2500	0.1204	0.1055	0.1267
$\rho_2(2)$	0.1300	0.1179	0.1044	0.1220
$\rho_2(5)$	0.1100	0.1106	0.1012	0.1092
$\rho_2(10)$	0.0500	0.0992	0.0960	0.0911
BOMBAY				
Mean	0.0111	0.0000	0.0000	0.0000
Variance	3.3948	3.6066	3.1291	4.2475
Kurtosis	6.8799	3.3226	31.7640	3.2679
y_t^2				
$\rho_2(1)$	0.1300	0.0868	0.1796	0.0934
$\rho_2(2)$	0.1300	0.0851	0.1664	0.0885
$\rho_2(5)$	0.1000	0.0802	0.1323	0.0755
$\rho_2(10)$	0.0400	0.0709	0.0903	0.0582

Table 4: Estimated EGARCH models and diagnostics for the original series

y_t	US-ES	S&P 500	US-JA	BOMBAY
ω	-0.0087	0.0045	-0.0110	0.0646
	(0.0134)	(0.0067)	(0.0116)	(0.0100)
α	0.0693	0.0886	0.1460	0.1314
	(0.0130)	(0.0086)	(0.0132)	(0.0013)
β	0.9881	0.9899	0.9692	0.9540
	(0.0046)	(0.0015)	(0.0051)	(0.0076)
γ	0.0089	-0.0495	-0.0330	-0.0301
	(0.0064)	(0.0065)	(0.0074)	(0.0099)
log L	-1800.9	-3516.3	-2160.6	-2507.9
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$				
Mean	0.0387	0.0638	-0.0062	0.0124
S.D.	0.9999	0.9935	1.0036	1.0049
Skew	-0.1557*	-0.6905*	-0.4821*	-0.0597
Kurtosis	4.9981*	7.8433*	5.8503*	7.7198*
$r(1)$	-0.0300	0.0300	0.0300	0.0900
$Q(20)$	18.51	26.65	20.53	44.44*
$r_2(1)$	0.0200	0.0200	0.0300	0.0300
$r_2(2)$	-0.0200	0.0020	0.0100	0.0200
$r_2(5)$	-0.0200	0.0030	-0.0200	0.0200
$r_2(10)$	0.0030	-0.0050	-0.0200	-0.0040
$Q_2(20)$	7.19	6.37	15.57	57.72*

Table 5: Descriptive statistics of daily returns corrected by outliers

	US CA	US ES	US GE	US JA	US SF	US SW	US UK	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
T	1963	1963	1963	1963	1963	1963	1963	1261	1262	2728	1982	2888
Mean	0.0078	0.0279	0.0190	0.0015	0.0174	0.0194	-0.0016	0.0865	0.0228	0.0478	0.0819	0.0642*
S.D.	0.2967	0.6048	0.5919	0.7307	0.6633	0.6161	0.4863	1.2748	1.7412	0.8867	1.1586	0.8255
Skew.	-0.1931*	-0.1551*	-0.2090*	-0.2354*	-0.0993	-0.0165	-0.1280*	-0.2947*	-0.0755	-0.2295*	-0.0893	-0.0944*
κ_y	4.4274*	4.0603*	3.9788*	5.0085*	4.0996*	3.9324*	4.2148*	4.6940*	5.1971*	4.7981*	4.2552*	4.3528*
$r(1)$	0.0600*	-0.0400	0.0000	-0.0200	0.0000	0.0000	0.0000	0.0200	0.0800*	0.0400	0.0900*	0.0300
$Q(20)$	23.3	20.1	23.4	20.3	21.9	21.6	25.5	37.7*	44.3*	57.9*	51.9*	51.5*
Autocorrelations of y_t^2												
$r_2(1)$	0.0800*	0.0800*	0.0400	0.1300*	0.0500*	0.0700*	0.1000*	0.1300*	0.1200*	0.0900*	0.1400*	0.1100*
$r_2(2)$	0.0600*	0.0500*	0.0200	0.0400	0.0100	0.0400	0.1200*	0.1600*	0.2100*	0.0900*	0.1600*	0.1400*
$r_2(5)$	0.0600*	0.0300	0.0400	0.0500*	0.0500*	0.0300	0.0600*	0.1400*	0.1300*	0.1000*	0.1300*	0.1300*
$r_2(10)$	0.0500*	0.0700*	0.0500*	0.0500*	0.0400	0.0400	0.0800*	0.1000*	0.0600*	0.1400*	0.1400*	0.0900*
$Q_2(20)$	166*	95.0*	74.5*	218*	123*	145*	217*	443*	246*	562*	705*	671*

T: Sample size.

κ_y: kurtosis of y_t.

$\hat{\rho}(\tau)$: Autocorrelation of order τ of the original observations y_t.

$\hat{\rho}_2(\tau)$: Autocorrelation of order τ of the squared observations y_t^2.

$Q(20)$ and $Q_2(20)$: Box-Ljung statistic for y_t and y_t^2 respectively (31.4 is the 5% critical value).

* Significant at the 5% level.

Table 6: Descriptive Statistics based on 1000 replicates for three models with isolated (IS) and consecutive (CS) outliers

Series	Mean	S.D.	Skew.	Kurt.	$r(1)$	$r_2(1)$	LM test (% rejections)
Gaussian (a_t)	0.0018	1.0003	0.0003	2.9794	-0.0019	-0.0025	4.40
CS $(a_t^\star)$	0.0318	1.0721	0.5919*	5.6537*	0.0847	0.3763*	99.50
IS $(a_t{}')$	0.0318	1.0721	0.5912*	5.6472*	-0.0021	-0.0054	1.70
GARCH (y_t)	0.0022	0.9951	0.0034	3.2619*	-0.0011	0.1167*	63.10
CS $(y_t^\star)$	0.0322	1.0675	0.6104*	5.9715*	0.0855	0.4097*	99.70
IS $(y_t{}')$	0.0322	1.0676	0.6123*	5.9894*	-0.0020	0.0397	14.30
EGARCH (x_t)	0.0015	0.9937	0.0160	3.2132*	-0.0024	0.0530*	27.10
CS $(x_t^\star)$	0.0315	1.0651	0.6029*	5.8261*	0.0837	0.3782*	99.70
IS $(x_t{}')$	0.0315	1.0648	0.5995*	5.8032*	-0.0031	0.0170	5.20

* Significant at the 5% level.

Table 7: Estimated GARCH models and diagnostics for the corrected series by outliers

y_t	US-ES	S&P 500	US-JA	BOMBAY
ω	0.0038	0.0039	0.0089	0.1298
	(0.0013)	(0.0010)	(0.0014)	(0.0259)
α	0.0299	0.0334	0.0424	0.0746
	(0.0055)	(0.0045)	(0.0055)	(0.0117)
β	0.9599	0.9602	0.9410	0.8828
	(0.0082)	(0.0052)	(0.0064)	(0.0177)
$\alpha + \beta$	0.9898	0.9936	0.9838	0.9574
log L	-1758.369	-3351.154	-2082.421	-2414.408
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$				
Mean	0.0445	0.0796*	0.0001	0.0168
S.D.	0.9915	1.0096	0.9962	1.0004
Skew	-0.1496*	-0.1853*	-0.2903*	-0.1020*
Kurtosis	4.1440*	4.0186*	4.7636*	5.4231*
$r(1)$	-0.0400	0.0400	-0.0100	0.1100*
$Q(20)$	16.9	37.6*	19.9	52.3*
$r_2(1)$	0.0200	0.0000	0.0200	-0.0100
$r_2(2)$	-0.0200	0.0100	-0.0100	0.0100
$r_2(5)$	-0.0100	0.0100	-0.0200	0.0100
$r_2(10)$	0.0100	0.0000	-0.0300	-0.0100
$Q_2(20)$	12.2	12.9	13.6	14.1

44

Table 8: Estimated GARCH models and diagnostics for the series corrected by conditional outliers

y_t	US-ES	S&P 500	US-JA	BOMBAY
ω	0.0019	0.0039	0.0091	0.1026
	(0.0009)	(0.0010)	(0.0016)	(0.0234)
α	0.0392	0.0361	0.0531	0.0963
	(0.0063)	(0.0030)	(0.0068)	(0.0137)
β	0.9569	0.9588	0.9309	0.8715
	(0.0071)	(0.0036)	(0.0079)	(0.0170)
$\alpha + \beta$	0.9961	0.9949	0.9840	0.9678
log L	-1732.074	-3479.874	-2080.440	-2386.581
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$				
Mean	0.0433	0.0729*	0.0061	0.0163
S.D.	1.0013	0.9985	1.0004	0.9999
Skew	-0.1592*	-0.6240*	-0.2104*	-0.0140
Kurtosis	3.8501*	7.8999*	4.7933*	4.1322*
$r(1)$	-0.0400	0.0370	0.0010	0.1140*
$Q(20)$	18.85	25.86	27.84	56.07*
$r_2(1)$	0.0140	0.0130	0.0140	0.0200
$r_2(2)$	-0.0170	0.0180	0.0090	-0.0120
$r_2(5)$	-0.0020	0.0120	-0.0200	0.0090
$r_2(10)$	0.0170	-0.0090	-0.0350	0.0100
$Q_2(20)$	10.66	5.96	16.67	13.66

Table 9: Observations detected as outliers

Series	Marginal outliers			Conditional outliers	
US-ES	**93**	151		**93**	**392**
	392	1876		861	909
	1958			1662	1757
S&P 500	20	47	**114**	**114**	495
	494	**812**	**1023**	**812**	**1023**
	2526	2527	2719	1338	1585
	2736	2738	2739	2112	
	2743	2770			
US-JA	**287**	673		**287**	664
	699	**1114**		988	**1114**
	1391	1443		1176	1266
	1448	**1468**		**1391**	**1468**
	1469	1476			
BOMBAY	321	**351**		250	**351**
	370	**853**		**370**	753
	1023	1130		**853**	**1023**
				1066	