A discusión

SPURIOUS AND HIDDEN VOLATILITY*

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

WP-AD 2004-45

Corresponding author: M. Angeles Carnero: Dpto. Fundamentos Análisis Económico. Universidad de Alicante, Campus de San Vicente del Raspeig, 03690 Alicante, Spain. Tel. +34 965 903 400, Fax: +34 965 903 898. E-mail: acarnero@merlin.fae.ua.es

IVIE working papers offer in advance the results of economic research under way in order to encourage a discussion process before sending them to scientific journals for their final publication.

** Mª Angeles Carnero: Universidad de Alicante, Campus de San Vicente del Raspeig, 03690 Alicante, Spain.
D. Peña: Dpto Estadísitca, Universidad Carlos III, Madrid. E-mail: dpena@est-econ.uc3m.es
E. Ruiz: Dpto. Estadística, Universidad Carlos III, Madrid. E-mail: Ortega@est-econ.uc3m.es

SPURIOUS AND HIDDEN VOLATILITY

M. Angeles Carnero, Daniel Peña and Esther Ruiz

ABSTRACT

This paper analyzes the effects caused by outliers on the identification and estimation of GARCH models. We show that outliers can lead to detect spurious conditional heteroscedasticity and can also hide genuine ARCH effects. First, we derive the asymptotic biases caused by outliers on the sample autocorrelations of squared observations and their effects on some homoscedasticity tests. Then, we obtain the asymptotic biases of the OLS estimates of ARCH(p) models and analyze their finite sample behaviour by means of extensive Monte Carlo experiments. The finite sample results are extended to GLS and ML estimates ARCH(p) and GARCH(1,1) models.

1   Introduction

Generalized AutoRegressive Conditional Heteroskedasticity (GARCH) models were introduced by Engle (1982) and Bollerslev (1986) to represent the dynamic evolution of conditional variances. However, when these models are fitted to real time series, the residuals often have excess kurtosis, which could be explained, among other reasons, by the presence of outliers; see, for example, Friedman and Laibson (1989) and Franses and Ghijsels (1999).

Previous results on the effects of outliers on the identification and estimation of conditional heteroscedasticy are somehow confusing. Some authors argue that outliers generate spurious heteroscedasticity. For example, Balke and Fomby (1994) conclude that outliers in several macroeconomic series of the US economy are able to explain most of the observed non-linearities. A similar conclusion is reached by Franses and Gijsels (1999) for macroeconomic series and Aggarwal et al. (1999) and Franses et al. (2004) for financial returns. On the other hand, other authors suggest that the presence of outliers may hide genuine heteroscedasticity; see, for example, Mendes (2000) and Li and Kao (2002) for an empirical application with exchange rates returns.

We show in this paper that additive outliers in uncorrelated GARCH series may generate spurious heteroscedasticity when they appear in patches, and hide legitimate heteroscedasticity when they are isolated. Consequently, both the size and power of tests for conditional homoscedasticity can be distorted in the presence of outliers. Also, they bias the sample autocorrelations of squares and the estimators of the parameters of the conditional variance as well as their standard deviations.

The paper is organized as follows. Section 2 analyses the effects of level outliers, that do not affect the conditional variance, on the sample autocorrelations of squared observations and on several tests for conditional heteroscedasticity. Section 3 derives the asymptotic bias of the Ordinary Least Squares (OLS) estimator of the parameters of ARCH(p) models contaminated by level outliers. We also analyse their effects on the finite sample properties of the Generalized Least Squares (GLS) and the Maximum Likelihood (ML) estimators by means of extensive Monte Carlo experiments. These results are also extended to the ML estimator of the parameters of GARCH(1,1) models. Section 4 illustrates the results by analyzing real series of financial returns. Finally, Section 5 concludes the paper.

2 Effects of outliers on the identification of conditional heteroscedasticity

Suppose that the series of interest, y_t, is generated by a GARCH(1,1) model given by

$$
\begin{aligned}
y_t &= \varepsilon_t \sigma_t \\
\sigma_t^2 &= \alpha_0 + \alpha_1 y_{t-1}^2 + \beta \sigma_{t-1}^2
\end{aligned}
\tag{1}
$$

where ε_t is a Gaussian white noise with mean zero and variance one. The parameters α_0, α_1 and β are assumed to satisfy the usual restrictions to guarantee the positiveness, stationary and existence of the fourth order moment of y_t; see, for example, Bollerslev et al. (1994). It is often useful to write the GARCH(1,1) model as an ARMA(1,1) model for squares

$$y_t^2 = \alpha_0 + (\alpha_1 + \beta)y_{t-1}^2 + \nu_t - \beta\nu_{t-1} \tag{2}$$

where the noise, $\nu_t = \sigma_t^2(\varepsilon_t^2 - 1)$, is a zero mean uncorrelated sequence. However, it is conditionally heteroscedastic and, consequently, it is non-independent and non-Gaussian. The acf of y_t^2 has the shape of the acf of an ARMA(1,1) model with autoregressive parameter $\alpha_1 + \beta$ and moving average parameter β. From (2) it is also clear that when the ARCH parameter $\alpha_1 = 0$, the parameter β is not identified. In this case, the series y_t is homoscedastic.

Alternatively, the conditional variance of y_t can be specified as an ARCH(p) given by

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 \tag{3}$$

where the parameters α_i should also be restricted so that σ_t^2 is positive and y_t is stationary with finite fourth order moment. The ARCH(p) model is an AR(p) for squared observations given by

$$y_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \nu_t. \tag{4}$$

Therefore, the acf of y_t^2 has the same shape as the acf of an AR(p) model with autoregressive parameters α_i, $i = 1, ..., p$.

Given that ARCH(p) and GARCH(1,1) models are uncorrelated, the traditional distinction between additive and innovative outliers is not relevant. However, it is important to distinguish whether an outlier affects or not future conditional variances. Hotta and Tsay (1998) introduce two types of outliers in GARCH models: level (LO) and volatility (VO). In this paper

we focus on LO that affect only the level of the series and have no effect on the conditional variance; see also Sakata and White (1998). Therefore, if the series y_t is contaminated from time τ ownwards by k consecutive outliers of size ω, the observed series is given by

$$z_t = \begin{cases} y_t + \omega & \text{if} \quad t = \tau, \tau+1, \ldots, \tau+k-1 \\ y_t & \text{otherwise.} \end{cases} \tag{5}$$

but the conditional variance is like in (1) and depends on the underlying series y_t and not on the observed series z_t. Similarly, the conditional variance is given by (3) when dealing with an ARCH(p) model.

On the other hand, VO are defined in such a way that the underlying conditional variance depends on the observed series. We expect that similarly to what happens in the context of linear models, the effects of VO should be less important as they are transmitted by the same dynamics as the rest of the series; see, for example, Peña (2001).

Other alternative approaches of defining outliers in GARCH models can be found in Friedman and Laibson (1989) and Franses and van Dijk (2000). It is also interesting to point out that Li and Kao (2002) and Zhang (2004) propose to use influence measures in GARCH models to define outliers.

2.1 Effects on the correlogram of squares

The autocorrelation of order h, $h \geq 1$, of the squared observations of the contaminated series in (5) is estimated by

$$r(h) = \frac{\sum_{t=h+1}^{T} z_t^2 z_{t-h}^2 - \frac{T-h}{T^2} \left(\sum_{t=1}^{T} z_t^2 \right)^2}{\sum_{t=1}^{T} z_t^4 - T^{-1} \left(\sum_{t=1}^{T} z_t^2 \right)^2}. \tag{6}$$

If the sample size, T, is large relative to the order of the estimated auto-correlation, h, the numerator of $r(h)$ can be written as follows

$$\sum_{t \in \mathbb{T}(h)} y_t^2 y_{t-h}^2 + \sum_{i=0}^{h-1}(y_{\tau+i}+\omega)^2 y_{\tau+i-h}^2 + \sum_{i=h}^{k-1}(y_{\tau+i}+\omega)^2(y_{\tau+i-h}+\omega)^2 +$$

$$+ \sum_{i=k}^{k+h-1} y_{\tau+i}^2(y_{\tau+i-h}+\omega)^2 - T^{-1}\left[\sum_{t \in \mathbb{T}(0)} y_t^2 + \sum_{i=0}^{k-1}(y_{\tau+i}+\omega)^2\right]^2 \tag{7}$$

where $\mathbb{T}(s) = \{s+1, ..., \tau-1, \tau+k+s, ..., T\}$. Similarly, the denominator can be written as

$$\sum_{t \in \mathbb{T}(0)} y_t^4 + \sum_{i=0}^{k-1}(y_{\tau+i}+\omega)^4 - T^{-1}\left[\sum_{t \in \mathbb{T}(0)} y_t^2 + \sum_{i=0}^{k-1}(y_{\tau+i}+\omega)^2\right]^2 \tag{8}$$

If the order of the autocorrelation is smaller than the number of consecutive outliers, i.e., $h < k$, then the third summation in (7) contains $k - h$ terms which depend on ω^4. Therefore, it is easy to see that expression (7) is equal to $(k - h - \frac{k^2}{T})\omega^4 + o(\omega^4)$. However, if $h \geq k$ then the third summation in (7) disappears and the numerator of $r(h)$ is equal to $-\frac{k^2}{T}\omega^4 + o(\omega^4)$. On the other hand, expression (8) is equal to $(k - \frac{k^2}{T})\omega^4 + o(\omega^4)$. Then

$$\lim_{\omega \to \infty} r(h) = \begin{cases} 1 - \frac{h}{k(1-\frac{k}{T})} & \text{if } h < k \\ \frac{k}{k-T} & \text{if } h \geq k \end{cases} \tag{9}$$

Therefore, one single large outlier $(k = 1)$ always biases towards zero all the autocorrelations of squares while a set of k large consecutive outliers generate positive autocorrelations of squares for all the orders smaller than k and zero for the others. For example, two large consecutive outliers generate an autocorrelation of the squares of order one approximately equal to 0.5, being all the others close to zero. It is important to notice that the limits in (9) are valid for both homoscedastic and heteroscedastic series. Therefore, if a

7

heteroscedastic series is contaminated by a large single outlier, the detection of genuine heteroscedasticity is going to be difficult. On the other hand, when a homoscedastic series is contaminated by several large consecutive outliers, the positive autocorrelations of squares generated by the outliers can be confused with conditional heteroscedasticity.

As an illustration, we have simulated 1000 replicates of size $T = 1000$ of a Gaussian white noise process with zero mean and variance one and another 1000 replicates from a GARCH(1,1) model with parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.1$ and $\beta = 0.8$[1]. First, we have contaminated each series by one single LO of size 15 at time $t = 500$ and, second, by two consecutive outliers of the same size as before at times $t = 500$ and 501[2]. The top panels of Figure 1 plot the mean correlogram of squared observations through all Monte Carlo replicates corresponding to the homoscedastic Gaussian white noise. It can be seen that although the series are uncorrelated, the mean of the first estimated autocorrelation is approximately 0.5 when they are contaminated by consecutive outliers. The bottom panels of Figure 1 plot the same quantities together with the acf of squares of the GARCH(1,1) model, and we observe the same result.

2.2 Testing for conditional heteroscedasticity

Many popular tests for conditional homoscedasticity are based on autocorrelations of squares. Therefore, if these autocorrelations are biased, the properties of the tests will be affected. In this subsection we analyze the behavior

---

[1]Similar results have been obtained generating series by alternative conditional heteroscedastic models like EGARCH and Stochastic Volatility.

[2]Similar results have been obtained when outliers appear in other positions.

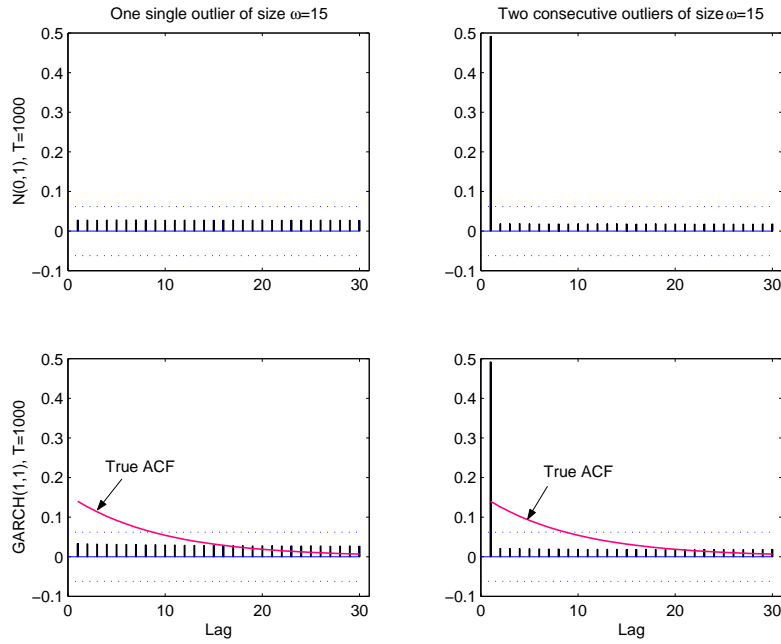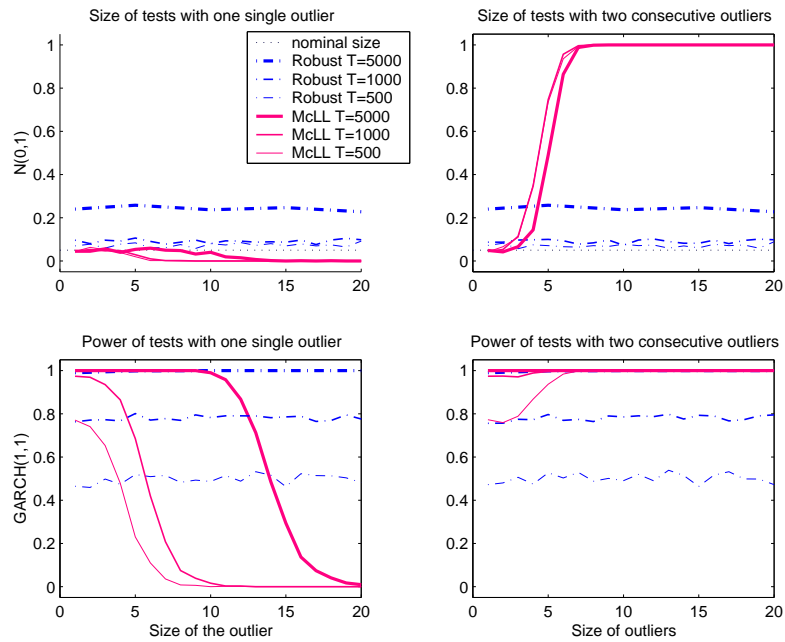Figure 1: Biases caused by outliers on the correlogram of squared observations

Figure 2: Effects caused by outliers on the size and power of conditional homoscedasticity tests

9

of two tests for conditional homoscedasticity, namely, the McLeod and Li
(1983) and the robust version of the Lagrange Multiplier (LM) test proposed
by Van Dijk et al. (1999)[3].

McLeod and Li (1983) proposed to test for conditional heteroscedasticity
using the Box-Ljung statistic for squared observations given by

$$Q(m) = T(T+2) \sum_{j=1}^{m} \frac{r^2(j)}{(T-j)}. \tag{10}$$

Under the null hypothesis of conditional homoscedasticity, if the eighth
order moment of y_t exists, $Q(m)$ is approximately distributed as a χ^2 with
m degrees of freedom.

Later, Van Dijk et al. (1999) show that, in the presence of consecutive
additive outliers, the LM test rejects the null hypothesis too often. Further-
more, large isolated outliers lead to an asymptotic power loss of the LM test;
see also Lee and King (1993). They propose an alternative robust statistic
(RLM) with better size and power properties; see also Franses et al. (2004)
for an empirical illustration with series of financial returns.

We consider first the properties of the McLeod-Li test in (10) when the
series y_t is affected by an isolated large outlier. In this case, from (9), the
limit of the estimated autocorrelations of any order is zero, so that the null
is never rejected. Thus, if the series is homoscedastic the size is zero while if
the series is heteroscedastic, the power is also zero.

When the series is affected by k consecutive outliers, from (9) we know

---

[3]Results for the LM test of Engle (1982) and the test proposed by Peña and Rodriguez
(2002) are similar to the ones obtained for the McLeod-Li test. They are not reported to
save space but are available from the authors upon request.

that the limit of the order one autocorrelation is $1 - \frac{T}{k(T-k)}$. Then,

$$\lim_{\omega \to \infty} Q(1) = \frac{T(T+2)}{(T-1)} \left(1 - \frac{T}{k(T-k)}\right)^2 \xrightarrow[T \to \infty]{} \infty,$$

and the null will always be rejected. Thus, if the series is truly homoscedastic, the asymptotic size is, in this case, one. On the other hand, if the series is heteroscedastic, the power is also one.

To analyze the finite sample effects of moderate outliers on these tests, we have simulated 1000 Gaussian white noise series of sizes $T = 500$, 1000 and 5000 that have been contaminated first, by one single outlier and then, by two consecutive outliers of the same size. For each simulated series, we test the null hypothesis of conditional homoscedasticity using the $Q(20)$ and the RLM(1) test. The top panel on the left of Figure 2 plots the empirical sizes of both tests as a function of the outlier size when it is isolated and the nominal size is 5%. This plot shows that, for $T = 500$ or 1000, the size of $Q(20)$ is zero for outliers larger than 7 standard deviations while the size of RLM(1) is around 9%, i.e. nearly double the nominal, independently of the outlier size. However, when $T = 5000$, the size of $Q(20)$ only goes to zero if the outlier is larger than 10 standard deviations while the size of RLM(1) is around 25%. The robust test is clearly oversized in large samples. Lee and King (1993) find similar size distortions in the robust test proposed by Wooldridge (1990).

The right panel on top of Figure 2 plots the empirical sizes of both tests when the Gaussian series are contaminated by two consecutive outliers. In this case, the behavior of the robust test is similar to the one observed when there is just one outlier. However, for relatively small outliers sizes, like for

11

example, 5 standard deviations, the size of the non-robust tests is almost 1 for any of the three sample sizes considered. Therefore, rather small consecutive outliers in homoscedastic series make the tests to detect conditional heteroscedasticity even for relatively large samples.

The power of the tests for isolated outliers is shown in the left bottom panel of Figure 2. Now the series are generated by the same GARCH$(1, 1)$ model considered above. This figure shows that if the outlier size is smaller than 4 or 5 standard deviations, the power of the portmanteau test is larger than the power of the robust test when the sample size is $T = 500$ or 1000. For these sample sizes, the power of the $Q(20)$ test decreases very rapidly with the size of the outlier. If this size is larger than approximately 7 standard deviations, the power is negligible. However, if $T = 5000$, then a very large outlier is needed for the RLM(1) test to have more power than the $Q(20)$ test. In our experiments, the power of the $Q(20)$ test is affected only if the outlier is larger than 10 standard deviations. We have also contaminated the GARCH series with two consecutive outliers. The empirical powers have been plotted in the right bottom panel of Figure 2. As we can see, for all sample sizes and outlier sizes chosen, the power of the robust test is clearly lower than the power of the non-robust test considered. A similar result is obtained by Lee and King (1993) comparing the power of the robust test proposed by Wooldridge (1990) with the LM test.

Summarizing, relatively small consecutive outliers are able to generate spurious heteroscedasticity while large isolated outliers are required to hide genuine heteroscedasticity when standard tests are used for testing for conditional homoscedasticity. On the other hand, the available robust LM test

seems to be of little help because it suffers from important size distortions that get worse with the sample size.

3 Effects of outliers on the estimation of ARCH and GARCH models

The ARCH(p) model often requires a large number of lags, p, to adequately represent the dynamic evolution of the conditional variances. However, this model is attractive because it is possible to obtain a closed-form expression for the OLS estimator of its parameters. In the following subsection, we quantify the effects of level outliers on the OLS estimates of ARCH(p) models. In subsection 3.2, we also analyze the effects of outliers on the GLS estimator. Finally, the results are extended in the next subsections to ML estimators for ARCH and GARCH models.

3.1 OLS estimator

The OLS estimator of the parameters of the ARCH(p) model defined in (4) is given by

$$\widehat{\alpha}^{OLS} = (X'X)^{-1}(X'Y) \tag{11}$$

where $\alpha = (\alpha_0 \quad \alpha_1 \quad \ldots \quad \alpha_p)'$, $Y = (y_{p+1}^2 \quad y_{p+2}^2 \quad \ldots \quad y_T^2)'$ and

$$X = \begin{pmatrix} 1 & y_p^2 & y_{p-1}^2 & \cdots & y_1^2 \\ 1 & y_{p+1}^2 & y_p^2 & \cdots & y_2^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & y_{T-1}^2 & y_{T-2}^2 & \cdots & y_{T-p}^2 \end{pmatrix}$$

Weiss (1986) shows that if the 4th order moment of y_t exists, $\widehat{\alpha}^{OLS}$ is consistent. Furthermore, if the 8th order moment is finite, the asymptotic distri-

bution of $\widehat{\alpha}^{OLS}$ is given by

$$\sqrt{T}(\widehat{\alpha}^{OLS} - \alpha) \xrightarrow{d} N(0, \Sigma_{XX}^{-1} \Sigma_{X\Omega X} \Sigma_{XX}^{-1})$$

where $Plim\frac{X'X}{T} = \Sigma_{XX}$ and $Plim\frac{X'VV'X}{T} = \Sigma_{X\Omega X}$ with $V = (\nu_{p+1}^2 \quad \nu_{p+2}^2 \quad \cdots \quad \nu_T^2)'$;
see Engle (1982) for sufficient conditions for the existence of higher moments
of y_t when ε_t is Gaussian.

A consistent estimator of the asymptotic covariance matrix of $\widehat{\alpha}^{OLS}$ is
given by

$$(X'X)^{-1}S(X'X)^{-1} \tag{12}$$

where

$$S = \begin{pmatrix} \sum_{t=p+1}^{T} \widehat{\nu}_t^2 & \sum_{t=p+1}^{T} \widehat{\nu}_t^2 y_{t-1}^2 & \cdots & \sum_{t=p+1}^{T} \widehat{\nu}_t^2 y_{t-p}^2 \\ & \sum_{t=p+1}^{T} \widehat{\nu}_t^2 y_{t-1}^4 & \cdots & \sum_{t=p+1}^{T} \widehat{\nu}_t^2 y_{t-1}^2 y_{t-p}^2 \\ & & \ddots & \vdots \\ & & & \sum_{t=p+1}^{T} \widehat{\nu}_t^2 y_{t-p}^4 \end{pmatrix}$$

and $\widehat{\nu}_t$ are the residuals from the OLS regression in (4).

Next, we analyze how a single outlier affects the asymptotic properties of
$\widehat{\alpha}^{OLS}$. We then consider the effects of patches of outliers.

3.1.1 Isolated outliers

Consider a series generated by an ARCH(p) model which is contaminated at
time τ by a single level outlier of size ω, as in (5) with $k = 1$. Then, $\widehat{\alpha}^{OLS}$ in
(11) will be computed using the contaminated observations z_t^2 instead of y_t^2.
It is shown in the Appendix that

$$\lim_{\omega \to \infty} \widehat{\alpha}_i^{OLS} = \begin{cases} \infty & \text{for } i = 0 \\ -\frac{1}{T-2p} & \text{for } i = 1, \ldots, p. \end{cases} \tag{13}$$

14

The limit in (13) shows that if the sample size is large enough: (i) the estimated unconditional variance, given by $\widehat{\alpha}_0/(1 - \sum_{i=1}^p \widehat{\alpha}_i)$, tends to infinity when the outlier size tends to infinity; (ii) all the estimated ARCH parameters tend to zero and, consequently, the dynamic dependence in the conditional variance disappears. Notice that the persistence of the volatility in an ARCH(p) model, measured by $\sum_{i=1}^p \alpha_i$, also decreases as the size of the outlier increases. Finally, it is also important to notice that if the sample size is not very large, it is possible to obtain estimates that do not satisfy the usual non-negativity restrictions.

3.1.2 Patches of outliers

When the original series, y_t, is contaminated by k consecutive outliers as in (5), the effects on the OLS estimator depend on the relationship between the number of outliers and the order of the ARCH model. First, let us consider $k \geq p$, i.e., there are at least as many outliers as the number of lags in the ARCH model. In this case, it is necessary to consider separately the cases where $p = 1$ and $p > 1$. This is because in the first case, the parameter α_1 receives the whole effect of the outliers while in the latter, this effect is shared by all the parameters.

We consider first the effect of k consecutive outliers on the estimates of the parameters of an ARCH(1) model. In this case, it is shown in the Appendix that

$$\lim_{\omega \to \infty} \widehat{\alpha}_i^{OLS} = \begin{cases} \infty & \text{for } i = 0 \\ \frac{(T-1)(k-1)-k^2}{(T-1)k-k^2} & \text{for } i = 1 \end{cases} \tag{14}$$

Notice that if $k = 1$, we obtain the same result as in (13). If the number of consecutive outliers is large, the estimated ARCH parameter, $\widehat{\alpha}_1$, tends to

15

one when the outliers size tends to infinity. Therefore, given that in ARCH(1) models, the persistence to shocks to volatility is measured by α_1, the presence of long patches of large outliers can lead to infer that the volatility is characterized by a unit root and, consequently, that y_t is not stationary. Notice that patches of large outliers can overestimate or underestimate the ARCH parameter depending on its original value. For example, if the sample size is moderate and there are two large consecutive outliers, $\widehat{\alpha}_1$ tends to 0.5. Therefore, if $\alpha_1 < 0.5$, the OLS estimator will have a positive bias while if $\alpha_1 > 0.5$, the bias will be negative. However, notice that in cases of empirical interest in the context of financial time series, the ARCH parameter is usually rather small, never over 0.3, and then with patches of consecutive outliers, the OLS estimator will overestimate the ARCH parameter. In particular, if the series is truly homoscedastic, i.e. $\alpha_1 = 0$, then the estimated ARCH parameter will be close to 0.5 and can lead to conclude that the series is conditionally heteroscedastic. Finally, it is also important to point out that the limit in (14) increases very quickly with the number of consecutive outliers. For example, if $k = 3$, $\widehat{\alpha}_1$ tends to 0.66 while if $k = 4$ the limit is 0.75.

Next, we consider the effect of $k \geq p$ consecutive outliers in an ARCH(p) model with $p > 1$. It is shown in the Appendix that the limit of the estimates when the size of the outliers tend to infinity is given by

$$
\lim_{\omega \to \infty} \widehat{\alpha}_i^{OLS} = \begin{cases} \infty & \text{for } i = 0 \\ \frac{-2k^2 + (2k-p)(T-p)}{-2k^2 + (2k-p+1)(T-p)} & \text{for } i = 1 \\ 0 & \text{for } i = 2, \ldots, p-1 \\ \frac{-(T-p)}{-2k^2 + (2k-p+1)(T-p)} & \text{for } i = p \end{cases} \tag{15}
$$

The estimated parameters, $\widehat{\alpha}_i$, tend to zero, except $\widehat{\alpha}_1$ and $\widehat{\alpha}_p$. If the number of consecutive outliers is large relative to the order of the model, then $\widehat{\alpha}_1$

tends to a quantity close to one and $\widehat{\alpha}_p$ tends to zero. Consequently, the estimated persistence, given by $\sum_{i=1}^{p} \widehat{\alpha}_i$, tends to $\frac{-2k^2+(2k-p-1)(T-p)}{-2k^2+(2k-p+1)(T-p)}$ which is close to one. Notice that if $p = 1$, the limit of the persistence coincides with the limit of $\widehat{\alpha}_1$ given in (14). Consider, for example, an ARCH(2) series contaminated by 2 large consecutive outliers. In this case, if the sample size is moderately large, $\widehat{\alpha}_1$ tends approximately to 0.66 and $\widehat{\alpha}_2$ to -0.34 and, consequently, the persistence tends to 0.32. However, if the number of consecutive outliers is 5, $\widehat{\alpha}_1$ tends to 0.89 and $\widehat{\alpha}_2$ to -0.11 and the persistence tends to 0.78. On the other hand, if there are 5 consecutive outliers in an ARCH(4) series, $\widehat{\alpha}_1$ tends to 0.86 and $\widehat{\alpha}_4$ to -0.15 and the persistence to 0.71. It is also important to notice that in the presence of patches of outliers, the estimates may easily violate the non-negativity restrictions.

3.2   Generalized Least Squares estimator

The OLS estimator is not efficient because the noise ν_t is conditionally heteroscedastic. Taking into account this heterogeneity, it is possible to obtain a better estimator. Model (4) can be expressed in matrix form as follows,

$$Y = X\alpha + V$$

and pre-multiplying by P, where $P'P = \Omega^{-1}$ and $\Omega = diag(\sigma_{p+1}^4, \cdots, \sigma_T^4)$ the following expression is obtained

$$PY = PX\alpha + PV \tag{16}$$

The GLS estimator is obtained by estimating by OLS the parameters α in equation (16). In practice given that the matrix Ω is unknown, it can be

substituted by $\widehat{\Omega} = diag(\widehat{\sigma}_{p+1}^4, \cdots, \widehat{\sigma}_T^4)$, where $\widehat{\sigma}_t^2 = \widehat{\alpha}_0^{OLS} + \widehat{\alpha}_1^{OLS} y_{t-1}^2 + ... +$ $\widehat{\alpha}_p^{OLS} y_{t-p}^2$. Therefore, the GLS estimator is given by

$$\widehat{\alpha}^{GLS} = (X'\widehat{\Omega}^{-1}X)^{-1}(X'\widehat{\Omega}^{-1}Y) \qquad (17)$$

The GLS estimator is very easy to obtain and its asymptotic efficiency is equivalent to the ML estimator. Bose and Mukherjee (2003) derive the asymptotic distribution of $\widehat{\alpha}^{GLS}$ and show that if the sixth order moment of y_t is finite, then

$$\sqrt{T}\left(\widehat{\alpha}^{GLS} - \alpha\right) \xrightarrow{d} N(0, \Sigma_\Omega^{-1}) \qquad (18)$$

where $Plim\frac{X'\Omega^{-1}X}{T} = \Sigma_\Omega$.

Suppose that there is a single outlier in an ARCH(1) series. Then, we have seen before that $\widehat{\alpha}_1^{OLS}$ will be close to zero. Consequently, the weights $\widehat{\sigma}_t^{-4}$ for the GLS estimator will be almost constant and therefore the GLS and OLS estimates will be very similar. To illustrate this behaviour we have generated 1000 series of sizes $T = 500$, 1000 and 5000 by an ARCH(1) model with parameters $\alpha_0 = 0.8$ and $\alpha_1 = 0.2$[4]. All the series have been contaminated with a single LO of size ω with $\omega = 5, 10$ and 15 standard deviations. Figure 3 plots kernel estimates of the density of the OLS and GLS estimators of α_0 and α_1 obtained through all Monte Carlo replicates. Comparing the kernel densities of the estimates of α_0, we can observe that in small or moderate samples both estimators have similar sample distributions. The estimates of α_0 are positively biased. However, when $T = 5000$, the bias of the GLS estimator is negligible even if $\omega = 15$, while the OLS estimator

---

[4]Similar results have been obtained when the series are generated by ARCH(p) models with $p > 1$.

Figure 3: Kernel estimates of the density of estimators of an ARCH(1) model with a single outlier
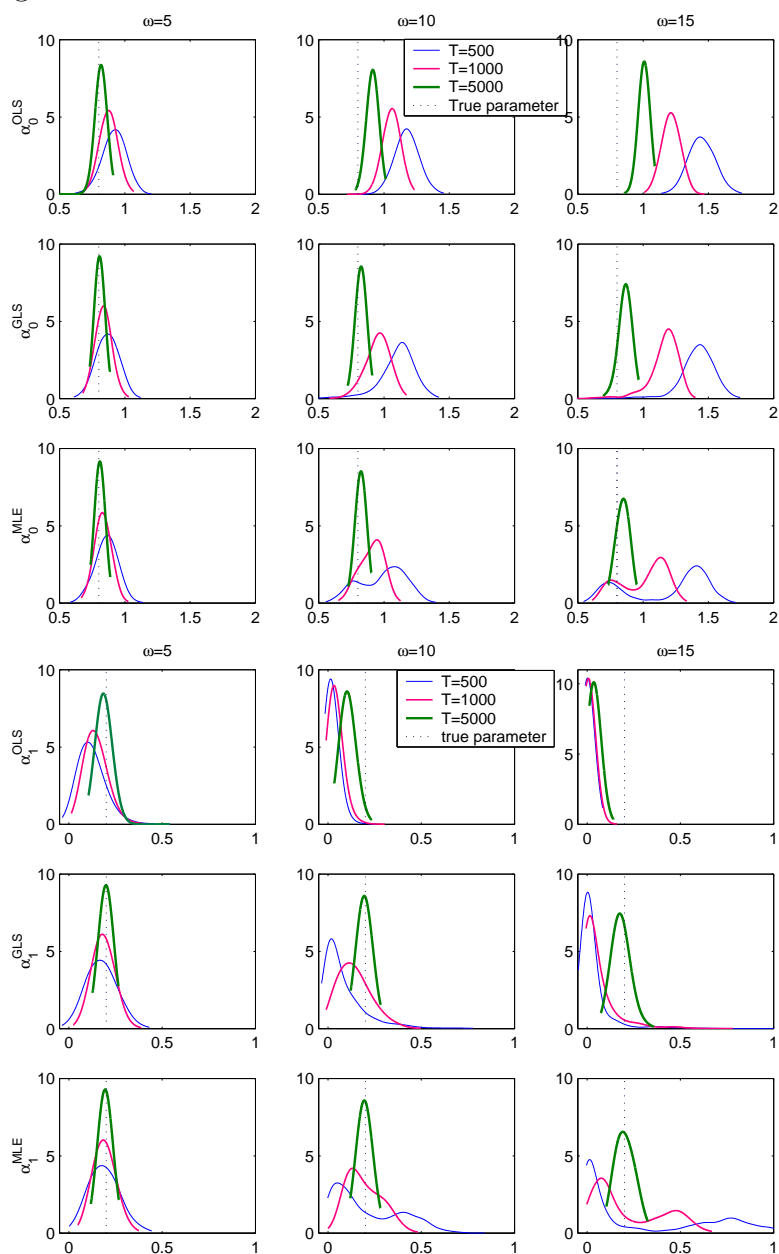
has large biases for rather small outliers. For example, when the outlier size is 10, the means of the OLS estimator of α_0 for $T = 500, 1000$ and 5000 are 1.18, 1.06 and 0.91 respectively.

Looking at the results for the OLS and GLS estimators of α_1, we can observe that the negative bias of the OLS estimator is large for moderate outliers even if the sample size is large. However, the GLS estimator of α_1 is more robust against outliers than the OLS estimator. For example, when $T = 5000$, the GLS estimator is unbiased in the presence of outliers as large as 15 standard deviations. However, the means of the OLS estimator when $\omega = 10$ are 0.02, 0.04 and 0.11 when $T = 500$, 1000 and 5000 respectively.

We have also analysed how an isolated outlier affects the small sample estimates of the asymptotic variances of the OLS and GLS estimators. Figure 4 plots the ratio between the empirical variances, and the estimated asymptotic variances of both estimators of α_0 and α_1 averaged through all Monte Carlo replicates. With respect to the variance of $\widehat{\alpha}_0^{OLS}$, it is possible to observe that the White variances in (12) overestimate the empirical variances. The bias is larger, the larger the outlier size. When estimating the variance of $\widehat{\alpha}_1^{OLS}$, the results in Figure 4 suggest that biases are rather small for moderate sample sizes. However, if the sample size is large, the empirical variance tends to zero and it is clearly overestimated using the White estimator. With respect to the variances of the GLS estimator, the ratio is larger than one, meaning that the asymptotic variance underestimates the empirical variance.

If there are consecutive outliers, we have seen that $\widehat{\alpha}_1^{OLS}$ will be overestimated and therefore the weights $\widehat{\sigma}_t^{-4}$ will down-weight the outliers. Thus the GLS estimator will be more robust than the OLS because of this down-

Figure 4: Ratio of variances of estimators of an ARCH(1) model with a single outlier

21
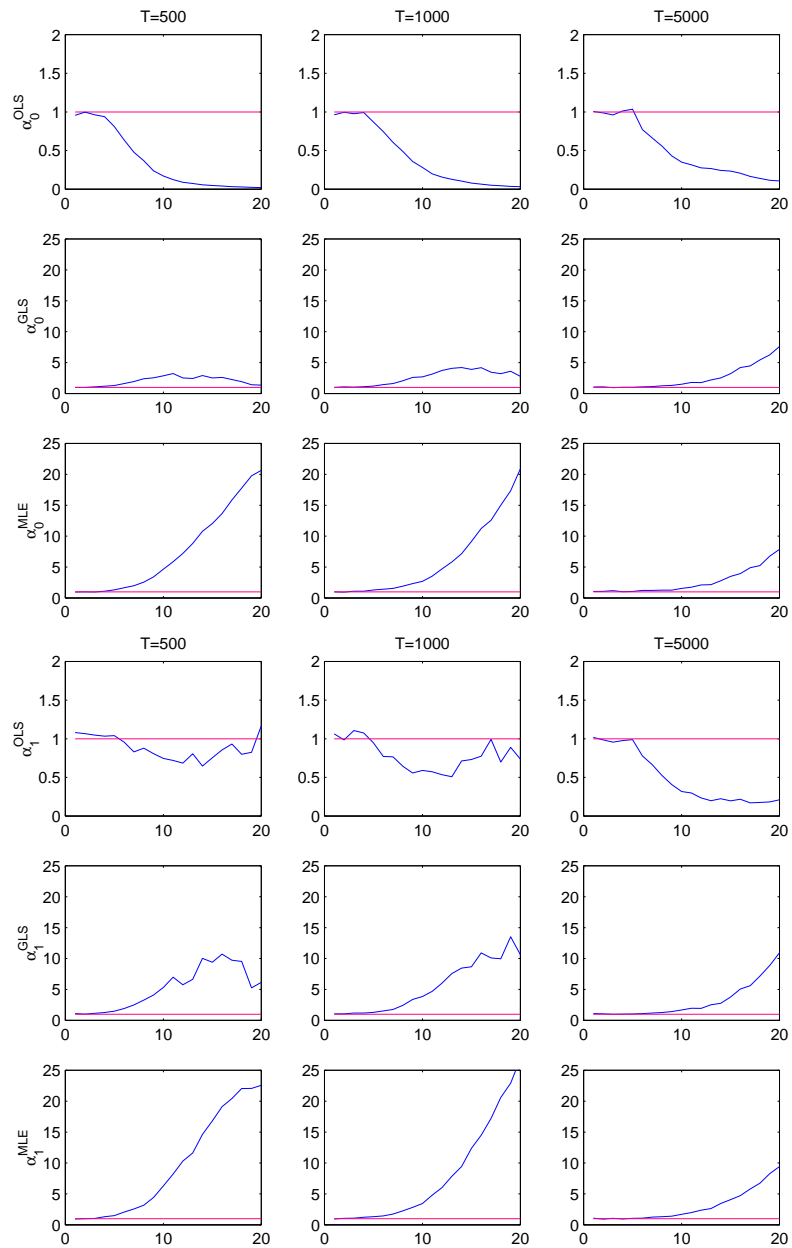
weighting. To illustrate these results, we have generated 1000 series by the same ARCH(1) model as before. Each series has been contaminated by 2 consecutive outliers. Figure 5 plots kernel estimates of the densities of the OLS and GLS estimators of α_0 and α_1 respectively. Although in the limit, $\widehat{\alpha}_0^{OLS}$ increases with ω, notice in this figure that for small outliers, α_0 can be underestimated. For example, consider $T = 500$ or 1000, then if the outlier size is 5 standard deviations, the mean of the estimates $\widehat{\alpha}_0^{OLS}$ is 0.75, below the true value of 0.8. However, if the size of the outlier is 15, the mean is 0.98. Consequently, for the outlier sizes typically encountered in empirical applications, the constant can be underestimated in the presence of patches of outliers. Remember that in the presence of a single outlier, the OLS estimates of α_0 tend monotonically to infinity. Therefore, although the effect in the limit is the same, in practice, isolated outliers overestimate the constant while consecutive outliers underestimate the constant. However, $\widehat{\alpha}_0^{GLS}$ is unbiased for all the outliers sizes considered in this paper as far as the sample size is large enough. When the sample size is small or moderate, large consecutive outliers increase the dispersion of the $\widehat{\alpha}_0^{GLS}$ estimates in such a way that the inference is useless.

Looking at the results for $\widehat{\alpha}_1^{OLS}$, observe that in concordance with the limit in (14), they tend to 0.5 when $k = 2$. Furthermore, for all the sample sizes considered, the limit is reached for sizes of the outliers relatively small. For example, for $T = 500$, the mean of the estimates of α_1 is 0.31 when $\omega = 5$, 0.46 when $\omega = 10$ and 0.49 when $\omega = 15$. Once more, the GLS estimates of α_1 are unbiased for $T = 5000$. However, if the sample size is moderate and the outliers are large, then $\widehat{\alpha}_1^{GLS}$ can take any value between 0 and 1, being

Figure 5: Kernel estimation of the density of estimators of an ARCH(1) model with two consecutive outliers
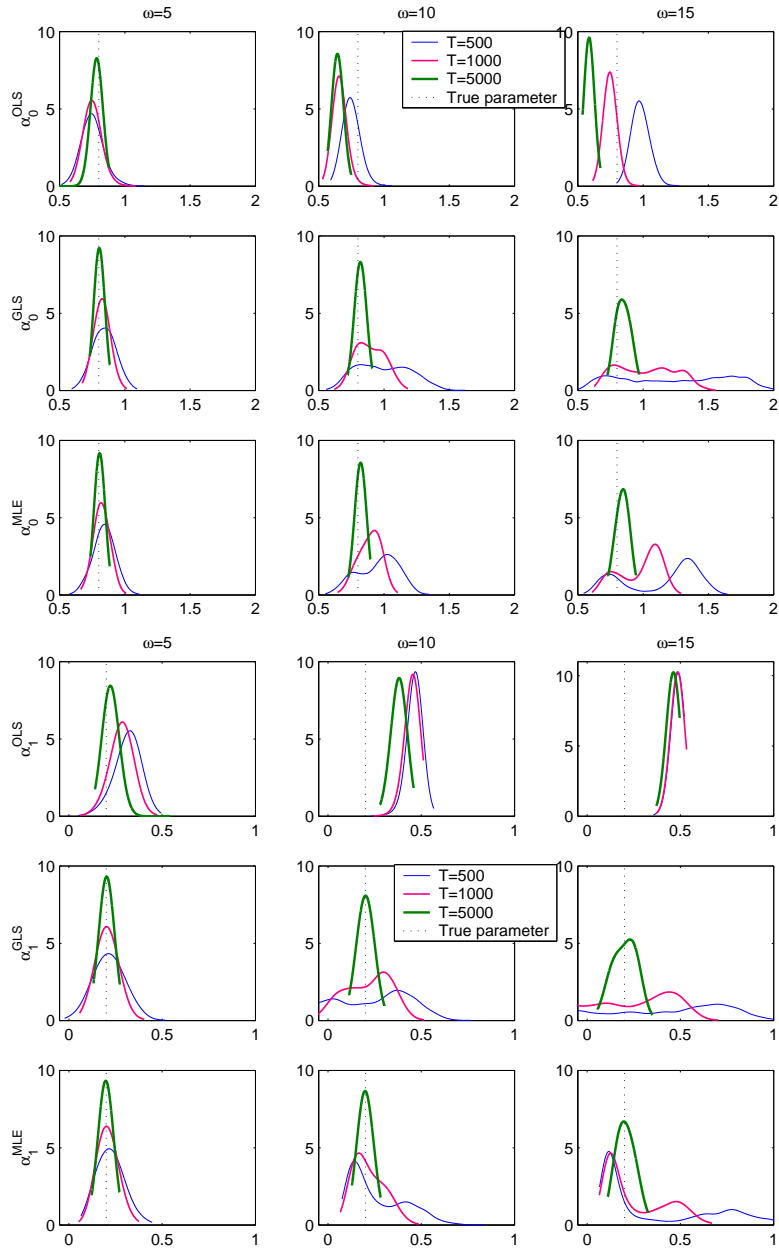
possible that α_1 is underestimated or overestimated.

Finally, the ratios of the empirical variances and the estimated asymptotic variances of the OLS and GLS estimators are plotted Figure 6, where we can see that for both estimators, $\widehat{\alpha}_0^{OLS}$ and $\widehat{\alpha}_1^{OLS}$, this ratio tends to zero with the size of the outlier. Therefore, the asymptotic variance of the OLS estimator, estimated using (12), is overestimating the true variance, which tends to zero with the size of the outlier. Notice that in this case, the biases are larger than in the presence of a single outlier. However, the estimated asymptotic variances of the GLS estimator, strongly underestimates the empirical variances for consecutive outliers larger than 5 standard deviations.

3.3 Maximum likelihood estimator of ARCH models

Engle (1982) proposed to estimate the parameters of the ARCH(p) model by ML. The distribution of y_t conditional on $Y_{t-1} = \{y_{t-1}, y_{t-2}, \cdots, y_1\}$ is $N(0, \sigma_t^2)$ and consequently, ML estimation of their parameters is straightforward maximizing the log-likelihood function given by

$$L = -\frac{T-p}{2}\log(2\pi) - \frac{1}{2}\sum_{t=p+1}^{T}\left(\log\sigma_t^2 + \frac{y_t^2}{\sigma_t^2}\right). \tag{19}$$

If the errors are not Gaussian, the estimates obtained by maximizing (19) are Quasi-Maximum Likelihood (QML) estimates. The consistency and asymptotic normality of the QML estimator was established by Ling and McAleer (2003) assuming that the second order moment of y_t is finite. Then,

$$\sqrt{T}\left(\widehat{\alpha}^{ML} - \alpha\right) \xrightarrow{d} N(0, [I(\alpha)]^{-1}) \tag{20}$$

where $I(\alpha) = E[-\frac{\partial^2 L}{\partial\alpha\partial\alpha'}]$ is the Information matrix.

Figure 6: Ratio of variances of estimators of an ARCH(1) model with two consecutive outliers
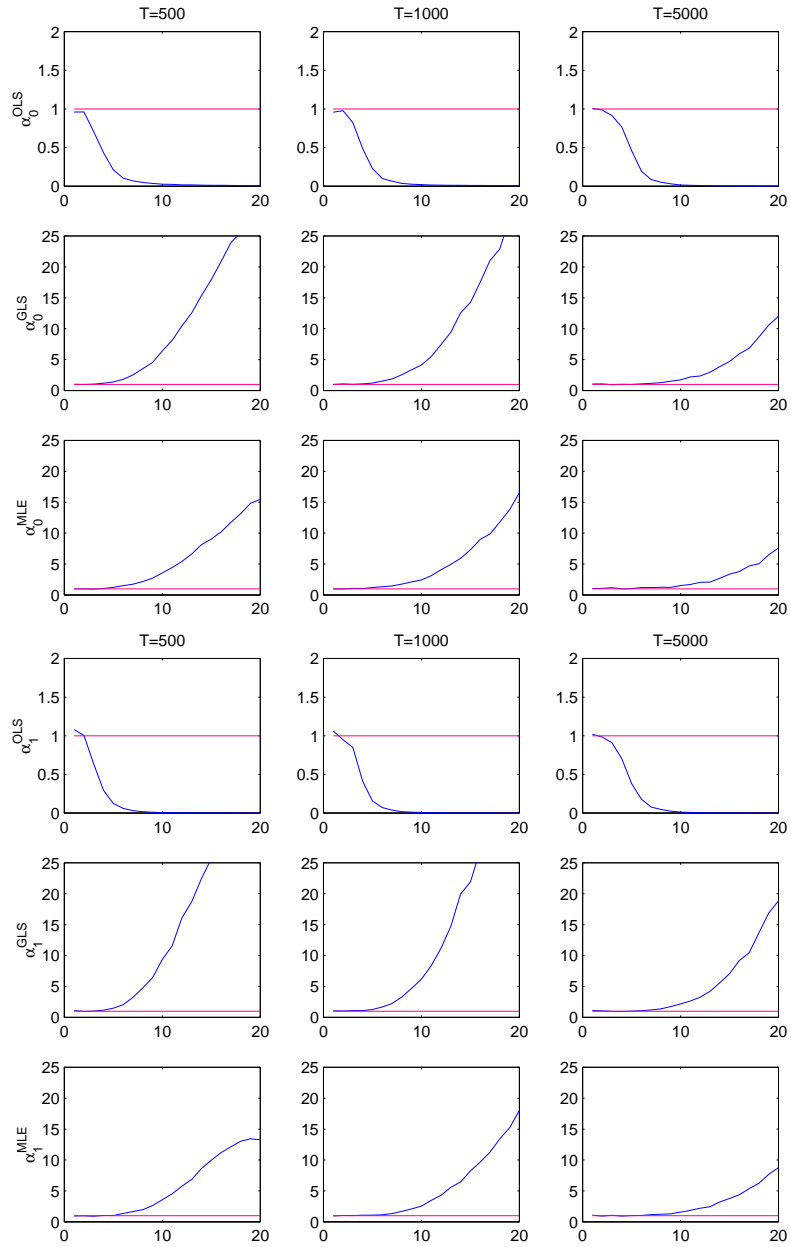
The QML is fully efficient when ε_t is Gaussian. However, there are not close form expressions of the QML estimators of the parameters α_0, $\alpha_1,...,\alpha_p$ and the numerical maximization of the Gaussian log-likelihood function is difficult because it is rather flat unless the sample size is very large; see, for example, Shephard (1996). Consequently, the analysis of the effects of outliers on the ML estimator has been carried out by simulation; see, for example, Muller and Yohai (2002) who show that the Mean Squared Error of the ML estimator of the parameters of ARCH(1) models is dramatically influenced by isolated outliers.

Consider the simplest ARCH(1) model. In this case, the log-likelihood function is given by

$$L = -\frac{T-2}{2}\log(2\pi) - \frac{1}{2}\sum_{t=2}^{T}\left(\log(\alpha_0 + \alpha_1 y_{t-1}^2) + \frac{y_t^2}{(\alpha_0 + \alpha_1 y_{t-1}^2)}\right)$$

which leads to the following ML equations to obtain the estimated parameters

$$\sum_{t=2}^{T}\frac{y_t^2}{\widehat{\sigma}_t^4} = \sum_{t=2}^{T}\frac{1}{\widehat{\sigma}_t^2}$$

$$\sum_{t=2}^{T}\frac{y_{t-1}^2 y_t^2}{\widehat{\sigma}_t^4} = \sum_{t=2}^{T}\frac{y_{t-1}^2}{\widehat{\sigma}_t^2}$$

Multiplying and dividing the right hand side by $\widehat{\sigma}_t^2 = \widehat{\alpha}_0 + \widehat{\alpha}_1 y_{t-1}^2$ we obtain

$$\widehat{\alpha}_0 \sum_{t=2}^{T}\frac{1}{\widehat{\sigma}_t^4} + \widehat{\alpha}_1 \sum_{t=2}^{T}\frac{y_{t-1}^2}{\widehat{\sigma}_t^4} = \sum_{t=2}^{T}\frac{y_t^2}{\widehat{\sigma}_t^4}$$

and

$$\widehat{\alpha}_0 \sum_{t=2}^{T}\frac{y_{t-1}^2}{\widehat{\sigma}_t^4} + \widehat{\alpha}_1 \sum_{t=2}^{T}\frac{y_{t-1}^4}{\widehat{\sigma}_t^4} = \sum_{t=2}^{T}\frac{y_{t-1}^2 y_t^2}{\widehat{\sigma}_t^4}$$

These equations represent the ML estimates as the result of solving the same system as the GLS considered in previous subsection by substituting in the

denominator the OLS estimate of σ_t^2 by the ML estimator. Therefore, as ML and GLS are asymptotically equivalent, the effects of outliers on both estimators should be similar for large samples. To illustrate this result, Figure 3 plots kernel estimates of the densities of the ML estimators of α_0 and α_1 when the ARCH(1) series are contaminated by a single outlier. Notice that, for sample sizes of $T = 500$ and 1000 and outliers of sizes 10 and 15 standard deviations, the kernel estimated density of both $\widehat{\alpha}_0^{ML}$ and $\widehat{\alpha}_1^{ML}$ are bimodal and non-symmetric. Hence, tests based on normality will be inadequate. Looking for example at the last row of Figure 3, we can see that in the presence of an outlier of size 15 standard deviations in a sample os size $T = 500$, $\widehat{\alpha}_1^{ML}$ could take any value between 0 and 1, although values close to zero seem to be more probable, like what we had for $\widehat{\alpha}_1^{OLS}$ and $\widehat{\alpha}_1^{GLS}$. Finally, if the sample size is 5000, the sample distributions of the GLS and ML are similar. Therefore, it is important to point out that the results in Figure 3 suggest that, in moderate samples, the GLS estimator has certain advantages over the ML estimator in the presence of large isolated outliers. In particular, both estimators have similar negative biases but the dispersion of the GLS is smaller.

Figure 4 plots the ratio of the empirical variance and the estimated asymptotic variance averaged through all Monte Carlo replicates for $\widehat{\alpha}_0^{ML}$ and $\widehat{\alpha}_1^{ML}$ respectively. We can see that this ratio is larger than the ratio of the GLS estimator. Therefore, the asymptotic variance of the ML estimator underestimates in the presence of large isolated outliers the empirical variance more than the GLS estimator.

The Monte Carlo densities when the series are contaminated by two con-

27

secutive outliers appear in Figure 5. As we can see in the plots, the effects caused by two consecutive outliers on the ML estimators are very similar to the effects caused by a single outlier. Finally, the effects of consecutive outliers on the estimated variances of the ML estimator are weaker than for the GLS estimator; see Figure 6.

3.4 Maximum Likelihood estimator of GARCH(1,1) models

The Gaussian log-likelihood of a GARCH(p, q) model is also given by (19). Ling and McAleer (2003) show that the QML estimator is consistent if the second order moment of y_t is finite and it is asymptotically normal if the sixth order moment is finite.

In this subsection, we carry out detailed Monte Carlo experiments to analyze the biases caused by isolated and consecutive LO on the QML estimates of the parameters of GARCH(1,1) models.

Figure 7 contains the kernel estimates of the density of $\widehat{\alpha}_0^{ML}$, $\widehat{\alpha}_1^{ML}$, $\widehat{\beta}^{ML}$ and $\widehat{\alpha}_1^{ML} + \widehat{\beta}^{ML}$ based on 1000 replicates, for a GARCH(1,1) model with parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.1$ and $\beta = 0.8$, contaminated with a single outlier of sizes $\omega = 5, 10$ and 15 standard deviations. As we can see in this figure, for large sample sizes, like $T = 5000$, ML estimators seem to be robust to the presence of outliers. Notice that they are unbiased even when the series is contaminated by an outlier of size 15 standard deviations. This is not true for smaller sample sizes, like $T = 500$ or 1000, where just one outlier seems to bias towards zero the estimated $\widehat{\alpha}_1^{ML}$ and towards one the estimated value $\widehat{\beta}^{ML}$. The same conclusion is obtained by Mendes (2000)

Figure 7: Kernel estimation of the density of ML estimators of GARCH models with a single outlier
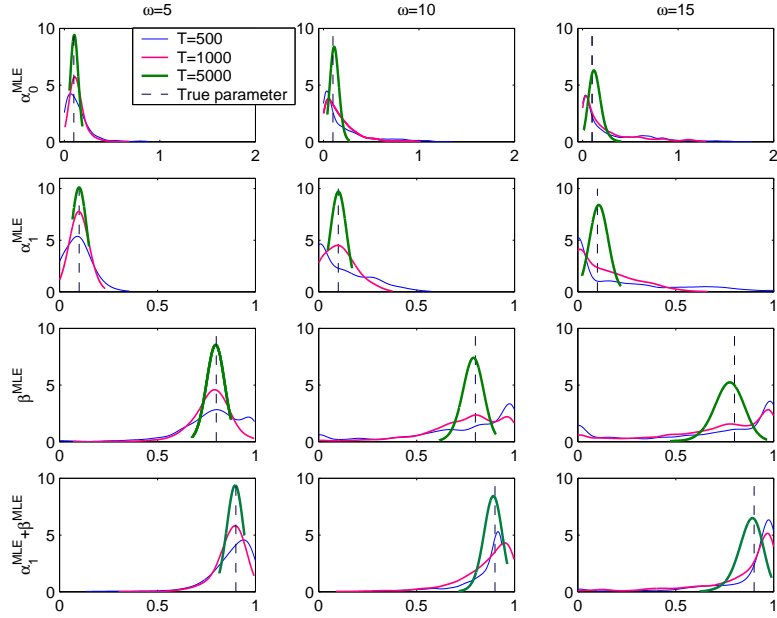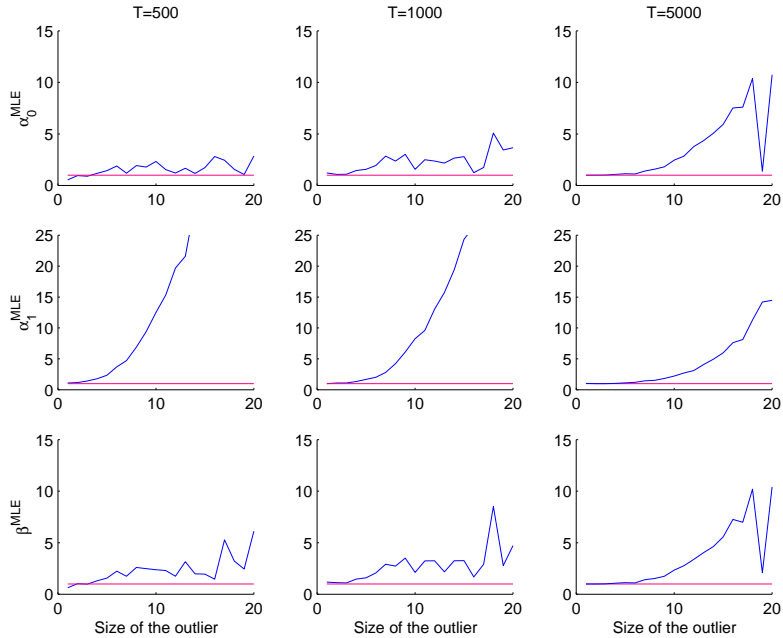

Figure 8: Ratio of variances of Maximum Likelihood estimators of a GARCH(1,1) model with a single outlier


29

and Sakata and White (1998) while it contradicts the results in Gregory and Reeves (2001) and Verhoeven and McAleer (2000). Finally, Figure 7 also plots kernel estimates of the densities of the estimated persistence. Notice that for large outliers and small sample sizes, the estimated persistence is also overestimated although there is a large tail to the left. Therefore, the distortions on the sample distribution of the ML estimates of the parameters also affect the overall persistence.

When $\widehat{\alpha}_1^{ML}$ and $\widehat{\beta}^{ML}$ take values close to zero and one respectively, there are problems in computing the asymptotic variance since the determinant of the Information matrix is very close to zero and then it is not possible to compute the asymptotic variance of the ML estimators. Figure 8 contains the ratio of the empirical variance and the estimated asymptotic variance averaged through all Monte Carlo replicates for $\widehat{\alpha}_0^{ML}$, $\widehat{\alpha}_1^{ML}$ and $\widehat{\beta}^{ML}$ for the series where this variance was finite. As we can see, like what we had found before for ARCH models, the ratio is greater than one meaning that the asymptotic variance strongly underestimates the empirical variance especially, for α_1.

Figure 9 plots kernel estimates of the density of $\widehat{\alpha}_0^{ML}$, $\widehat{\alpha}_1^{ML}$, $\widehat{\beta}^{ML}$ and $\widehat{\alpha}_1^{ML} + \widehat{\beta}^{ML}$ based on 1000 replicates, for the same GARCH(1,1) model but now contaminated with two consecutive outliers, it seems that $\widehat{\alpha}_0^{ML}$ and $\widehat{\alpha}_1^{ML}$ are overestimating the true parameters, and $\widehat{\beta}$ is underestimating the true β. However, even more important is to realize that if the outliers are large and the sample size is moderate, the sample densities of $\widehat{\alpha}_1^{ML}$ and $\widehat{\beta}^{ML}$ are such that standard inference is not reliable. Furthermore, Figure 9 shows that large consecutive outliers can have dramatic effects on the estimated persistence. For example, when $\omega = 15$ and $T = 500$ or 1000, the estimated

Figure 9: Kernel estimation of the density of ML estimators of GARCH models with two consecutive outliers
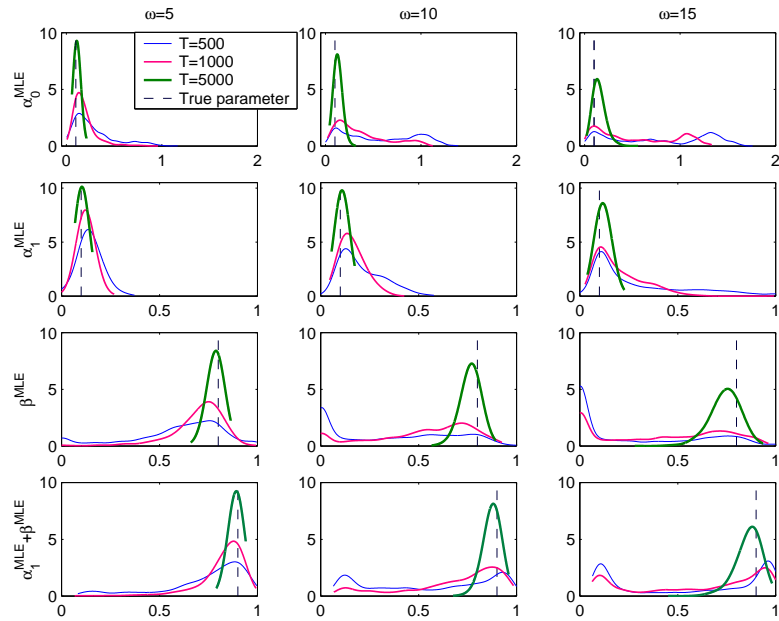
density of $\widehat{\alpha}_1^{ML} + \widehat{\beta}^{ML}$ has two modes, one around zero and another close to one. The estimates of the persistence are only reliable for very large sample sizes. Finally, the biases caused by consecutive outliers on the estimated variance of the ML estimators is similar to the biases caused by a single outlier.

4   Empirical application

This section illustrates the previous results by analyzing two world indexes. We consider daily series of returns of the S&P500 index of US and the Nikkei225 index of Japan, observed from October 20, 1982 to May 17, 2004 and from January 4, 1984 to May 19, 2004 respectively[5]. Figures 10 and 11 plot the return series and the correlogram of squared observations for the original series and for the series corrected for large outliers. Series have been corrected by substituting the corresponding outliers by the unconditional mean. In the S&P500 series there is one observation which is exactly 22 times the standard deviation. This observation corresponds to the 19th of October, 1987, also known as "October black monday", the biggest fall in all the history of Wall Street. The other two outliers correspond to the following days, October, 21 and 26. The size of these two observations is around 8.5 standard deviations. In the Nikkei225, there are two consecutive outliers corresponding to the same "October black monday", 19th and 20th of October, 1987. In this case, the corresponding return was 11 times the standard deviation. The third outlier in this series corresponds to 2nd October 1990, and this observation is 8 times the standard deviation.

---

[5]Series have been obtained in the webpage http://finance.yahoo.com/.

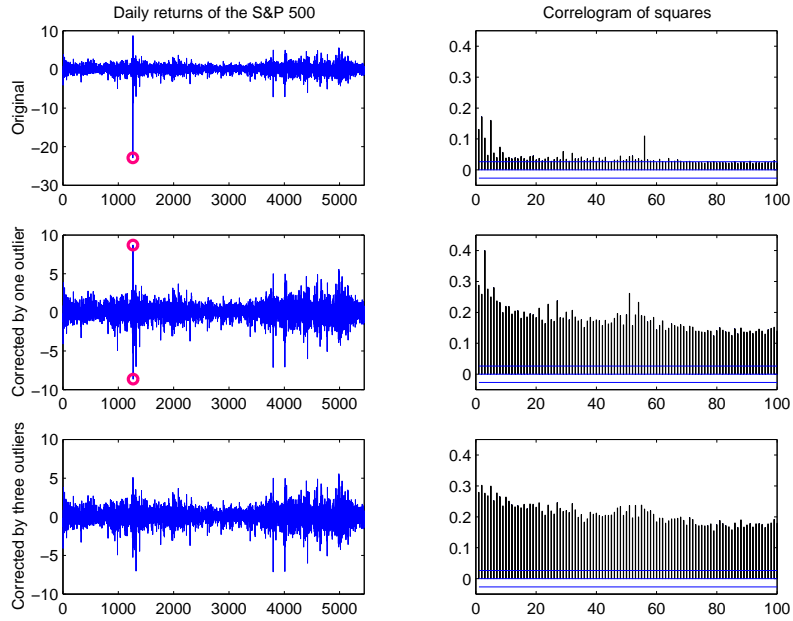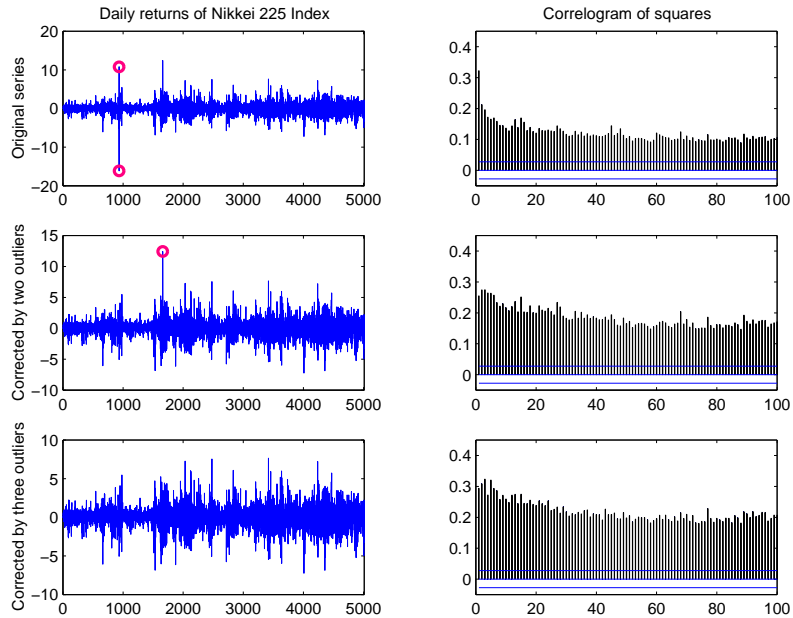Figure 10: Series and correlogram of squares of daily returns of S&P500 index

Figure 11: Series and correlogram of squares of daily returns of Nikkei 225 index

As we can observe in both figures, correcting the series by these extreme observations makes more clear the structure in the squared observations. In the case of the S&P500, we can see how just one observation biases towards zero all correlation coefficients of squared observations, and in the Nikkei 225, two consecutive outliers overestimate the first order autocorrelation and underestimate all the others. The value of the McLeod-Li, $Q(20)$, statistics are higher in the corrected series, 1847 and 1558, for the S&P500 and Nikkei225 series respectively, than in the original series of returns, 390 and 797 respectively. This is a sign of a more clear structure in the corrected squared observations.

To analyse the effects of these extreme observations on the OLS, GLS and ML estimates of the series considered in this section, Table 1 contains estimated parameters for the ARCH(9) model while Table 2 contains the ML estimates of the GARCH(1,1) model. We can see that, as expected, the constant, α_0 is overestimated in the original series. On the other parameters, α_i, for $i = 1, 2, \ldots, 9$, notice that, although the differences in the original and corrected series are not very big when we look at the point estimates, the non negativity restrictions are violated when the parameters are estimated by OLS in the contaminated series S&P 500. Also, remember that outliers affect the estimates and the variances of the estimators.

Consider now the results for the Nikkei225. The first conclusion from Table 1 is that for the three estimators, there is a positive bias in the estimated constant that decreases when the outliers are corrected. Furthermore, the OLS and ML estimates of α_1 are larger in the original series than in the corrected series while for all the other parameters, the estimates in the

34

Table 1: Estimates of the ARCH(9) Model

		S&P 500			NIKKEI 225		
		OLS	GLS	MLE	OLS	GLS	MLE
$\widehat{\alpha}_0$	original	0.71* (0.11)	0.45* (0.04)	0.27* (0.01)	1.05* (0.10)	0.52* (0.07)	0.34* (0.01)
	corrected	0.42* (0.04)	0.26* (0.04)	0.29* (0.01)	0.78* (0.07)	0.48* (0.09)	0.35* (0.02)
$\widehat{\alpha}_1$	original	0.09* (0.01)	0.16* (0.03)	0.10* (0.01)	0.22* (0.01)	0.19* (0.04)	0.23* (0.01)
	corrected	0.08* (0.01)	0.13* (0.03)	0.04* (0.01)	0.07* (0.01)	0.19* (0.04)	0.11* (0.01)
$\widehat{\alpha}_2$	original	0.13* (0.01)	0.05 (0.03)	0.11* (0.01)	0.05* (0.01)	0.10* (0.02)	0.12* (0.02)
	corrected	0.11* (0.01)	0.13* (0.03)	0.11* (0.01)	0.09* (0.01)	0.15* (0.04)	0.13* (0.02)
$\widehat{\alpha}_3$	original	0.03* (0.01)	0.06* (0.02)	0.09* (0.01)	0.05* (0.01)	0.09* (0.03)	0.11* (0.01)
	corrected	0.07* (0.01)	0.10* (0.03)	0.07* (0.01)	0.11* (0.01)	0.11* (0.04)	0.13* (0.01)
$\widehat{\alpha}_4$	original	−0.02 (0.01)	0.00 (0.01)	0.10* (0.01)	0.03 (0.01)	0.06* (0.02)	0.08* (0.01)
	corrected	0.06* (0.01)	0.11* (0.03)	0.10* (0.01)	0.03* (0.01)	0.08* (0.03)	0.08* (0.01)
$\widehat{\alpha}_5$	original	0.13* (0.01)	0.11* (0.03)	0.08* (0.02)	0.03 (0.01)	0.11* (0.02)	0.10* (0.01)
	corrected	0.10* (0.01)	0.09* (0.03)	0.09* (0.02)	0.11* (0.01)	0.09* (0.04)	0.11* (0.01)
$\widehat{\alpha}_6$	original	−0.01 (0.01)	0.03 (0.02)	0.08* (0.01)	0.04* (0.01)	0.08* (0.02)	0.06* (0.01)
	corrected	0.03* (0.01)	0.08* (0.02)	0.08* (0.01)	0.06* (0.01)	0.05 (0.04)	0.07* (0.01)
$\widehat{\alpha}_7$	original	−0.03* (0.01)	0.02* (0.01)	0.07* (0.01)	0.02 (0.01)	0.09* (0.02)	0.12* (0.02)
	corrected	0.07* (0.01)	0.07* (0.03)	0.08* (0.01)	0.06* (0.01)	0.08* (0.04)	0.13* (0.02)
$\widehat{\alpha}_8$	original	0.03* (0.01)	0.10* (0.02)	0.10* (0.01)	0.01 (0.01)	0.04 (0.02)	0.07* (0.01)
	corrected	0.06* (0.01)	0.09* (0.03)	0.11* (0.01)	0.01 (0.01)	0.05 (0.03)	0.06* (0.01)
$\widehat{\alpha}_9$	original	0.03* (0.01)	0.05* (0.02)	0.07* (0.01)	0.02 (0.01)	0.03 (0.02)	0.05* (0.01)
	corrected	0.02 (0.01)	0.04* (0.02)	0.06* (0.01)	0.04* (0.01)	0.04 (0.04)	0.06* (0.01)

Table 2: Estimates of the GARCH(1,1) Model

	Series	S&P 500	NIKKEI 225
α_0	original	0.012* (0.001)	0.027* (0.003)
	corrected	0.006* (0.001)	0.018* (0.002)
α_1	original	0.078* (0.002)	0.140* (0.003)
	corrected	0.046* (0.003)	0.101* (0.005)
β	original	0.915* (0.003)	0.862* (0.004)
	corrected	0.948* (0.004)	0.895* (0.005)
Log-Likelihood	original	−7299.6	−8145.9
	corrected	−7217.2	−8040.9

corrected series are larger. Notice that the biases are smaller for the GLS estimator. Therefore, as we have seen in the simulations, the GLS estimates are more robust in the presence of consecutive outliers than the OLS or ML estimates. Finally, it is important to notice that in Table 1, the estimated asymptotic standard deviations of the GLS estimator are larger in the corrected series. This empirical result is also in concordance with the Monte Carlo results in Figure 6 where we showed that the estimated asymptotic variances underestimate the sample variances of the GLS estimator in the presence of large consecutive outliers. With respect to the ML estimates of the GARCH(1,1) parameters in Table 2, it is possible to observe that α_0 and α_1 are also overestimated in the original series while β is underestimated.

5 Conclusions

This paper shows how in the presence of large isolated outliers, the size of the McLeod-Li test for conditional homoscedasticty is close to zero while its power is also close to zero if the sample size is relatively small. The effect of consecutive outliers is to reject the conditional homoscedasticity hypothesis even if the series is truly homoscedastic and the sample size is large. Effects of outliers on the estimation of ARCH and GARCH models have been analyzed finding that the biases caused by these observations can be very different depending on the size and position of the outliers in the series. We have also shown that outliers affect not only to estimated parameters but also the estimated variance of the estimators. For GLS and ML estimators, the asymptotic variance tends to underestimate the variance while for OLS

estimator, the asymptotic variance overestimates the empirical variance. All theoretical results have been illustrated by analyzing daily series of returns of the S&P 500 and the Nikkei 225 indexes.

Appendix: Asymptotic limits of OLS estimator of ARCH(p) models

Isolated outliers in ARCH(p) models

The OLS estimator of α in the AR(p) representation of the ARCH(p) model contaminated by one isolated outlier is given by

$$
\begin{pmatrix} \widehat{\alpha_0}^{OLS} \\ \widehat{\alpha_1}^{OLS} \\ \vdots \\ \widehat{\alpha_p}^{OLS} \end{pmatrix} = \begin{pmatrix} T-p & \sum_{t=p}^{T-1} z_t^2 & \cdots & \sum_{t=p}^{T-1} z_{t-p+1}^2 \\ \sum_{t=p}^{T-1} z_t^2 & \sum_{t=p}^{T-1} z_t^4 & \cdots & \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{t=p}^{T-1} z_{t-p+1}^2 & \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 & \cdots & \sum_{t=p}^{T-1} z_{t-p+1}^4 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=p+1}^{T} z_t^2 \\ \sum_{t=p+1}^{T} z_t^2 z_{t-1}^2 \\ \vdots \\ \sum_{t=p+1}^{T} z_t^2 z_{t-p}^2 \end{pmatrix}
$$

Taking into account that $z_\tau^2 = \omega^2 + o(\omega^2)$, $z_\tau^4 = \omega^4 + o(\omega^4)$ and $z_t^r = o(\omega)$ for $t \neq \tau$ and $\forall r \geq 0$, the matrix $X'X$ can be written as

$$
\begin{pmatrix} T-p & (\omega^2 + o(\omega^2))\,\mathbf{1}' \\ (\omega^2 + o(\omega^2))\mathbf{1} & (\omega^2 + o(\omega^2))\mathbf{F} \end{pmatrix}
$$

and consequently $(X'X)^{-1}$ can be written as

$$
\frac{1}{(T-2p)\omega^{4p}} \begin{pmatrix} \omega^{4p} + o(\omega^{4p}) & (-\omega^{4p-2} + o(\omega^{4p-2}))\mathbf{1}' \\ (-\omega^{4p-2} + o(\omega^{4p-2}))\mathbf{1} & \mathbf{V} \end{pmatrix}
$$

where $\mathbf{1}$ is a $p \times 1$ column vector of ones, $\mathbf{F}$ is a $p \times p$ symmetric matrix with $f_{ii} = \omega^2$ for $i = 1, \ldots, p$ and all other elements are equal to one. $\mathbf{V}$ is a $p \times p$ symmetric matrix with all its elements equal to $o(\omega^{4p-2})$. Finally, all

elements in $X'Y$ are equal to $\omega^2 + o(\omega^2)$. Consequently,

$$\lim_{\omega\to\infty} \widehat{\alpha}^{OLS} = \lim_{\omega\to\infty} \frac{1}{(T-2p)\omega^{4p}} \begin{pmatrix} \omega^{4p+2} + o(\omega^{4p+2}) \\ (-\omega^{4p} + o(\omega^{4p}))\mathbf{1} \end{pmatrix} = \begin{pmatrix} \infty \\ -\frac{1}{T-2p}\mathbf{1} \end{pmatrix}.$$

Consecutive outliers in ARCH(1) models

The OLS estimator of the parameters of the ARCH(1) model is given by

$$\begin{pmatrix} \widehat{\alpha_0}^{OLS} \\ \widehat{\alpha_1}^{OLS} \end{pmatrix} = \begin{pmatrix} T-1 & \sum_{t=1}^{T-1} z_t^2 \\ \sum_{t=1}^{T-1} z_t^2 & \sum_{t=1}^{T-1} z_t^4 \end{pmatrix}^{-1} \begin{pmatrix} \sum_{t=2}^{T} z_t^2 \\ \sum_{t=2}^{T} z_t^2 z_{t-1}^2 \end{pmatrix}$$

If there are k consecutive outliers of size ω, then $\sum_{t=1}^{T-1} z_t^2 = k\omega^2 + o(\omega^2)$
and $\sum_{t=1}^{T-1} z_t^4 = k\omega^4 + o(\omega^4)$ and the following result is obtained

$$\lim_{\omega\to\infty} \widehat{\alpha}^{OLS} = \lim_{\omega\to\infty} \frac{1}{((T-1)k-k^2)\omega^4} \begin{pmatrix} k\omega^4 + o(\omega^4) & -k\omega^2 + o(\omega^2) \\ -k\omega^2 + o(\omega^2) & T-1 \end{pmatrix} \begin{pmatrix} k\omega^2 + o(\omega^2) \\ (k-1)\omega^4 + o(\omega^4) \end{pmatrix}.$$

Hence,

$$\lim_{\omega\to\infty} \widehat{\alpha}^{OLS} = \begin{pmatrix} \infty \\ \frac{(T-1)(k-1)-k^2}{(T-1)k-k^2} \end{pmatrix}.$$

Consecutive outliers in ARCH(p) models with $k \geq p > 1$

Consider again the OLS estimator of the parameters of the ARCH(p) model.
If the series is contaminated by k consecutive outliers, then $\sum_{t=p}^{T-1} z_t^2, \sum_{t=p}^{T-1} z_{t-1}^2, \ldots, \sum_{t=p}^{T-1} z_{t-p+1}^2$
are equal to $k\omega^2 + o(\omega^2)$, $\sum_{t=p}^{T-1} z_t^4, \sum_{t=p}^{T-1} z_{t-1}^4, \ldots, \sum_{t=p}^{T-1} z_{t-p+1}^4$ are equal to
$k\omega^4 + o(\omega^4)$ and $\sum_{t=p}^{T-1} z_t^2 z_{t+1}^2 = (k-1)\omega^4 + o(\omega^4), \ldots, \sum_{t=p}^{T-1} z_t^2 z_{t-p+1}^2 = (k-p+1)\omega^4 + o(\omega^4)$. Therefore, the $X'X$ matrix is given by

$$\begin{pmatrix} T-p & (k\omega^2 + o(\omega^2))\,\mathbf{1}' \\ (k\omega^2 + o(\omega^2))\mathbf{1} & (\omega^4 + o(\omega^4))\mathbf{M} \end{pmatrix}$$

where $\mathbf{M}$ is a $p \times p$ symmetric matrix with $m_{ij} = k + i - j$ for $i = 1, \ldots, p$,
$j = i, \ldots, p$. Consequently, the OLS estimator is given by

$$\begin{pmatrix} \frac{2^{p-1}(k-(p-1)/2)}{2^{p-2}(-2k^2+(2k-p+1)(T-p))} & -\frac{1}{\omega^4}\frac{k}{-2k^2+(2k-p+1)(T-p)}\ell' \\ -\frac{1}{\omega^4}\frac{k}{-2k^2+(2k-p+1)(T-p)}\ell & (\omega^4+o(\omega^4))\mathbf{D} \end{pmatrix} \begin{pmatrix} k\omega^2+o(\omega^2) \\ (\omega^4+o(\omega^4))\mathbf{B} \end{pmatrix}$$

where $\mathbf{D}$ is a $p \times p$ symmetric matrix with $d_{11} = d_{pp} = \frac{1}{2\omega^4}\frac{-2k^2+(2k-p+2)(T-p)}{-2k^2+(2k-p+1)(T-p)}$, $d_{ii} = \frac{1}{\omega^4}$ for $i = 2,\ldots,p-1$, $d_{ii+1} = -\frac{1}{2\omega^4}$ for $i = 2,\ldots,p-1$, $d_{1p} = \frac{1}{2\omega^4}\frac{T-p}{-2k^2+(2k-p+1)(T-p)}$ and $d_{ij} = 0$ otherwise, $\ell = \begin{pmatrix} 1 & 0 & 0 & \ldots & 0 & 1 \end{pmatrix}'$ and $\mathbf{B}$ is a $p \times 1$ column vector such that $b_i = k - i$ for $i = 1,\ldots,p$. Then,

$$\lim_{\omega\to\infty}\widehat{\alpha}_i^{OLS} = \begin{pmatrix} \infty \\ \frac{-2k^2+(2k-p)(T-p)}{-2k^2+(2k-p+1)(T-p)} \\ 0 \\ \vdots \\ 0 \\ \frac{-(T-p)}{-2k^2+(2k-p+1)(T-p)} \end{pmatrix}.$$

References

[1] Aggarwal, R., C. Inclan and R. Leal, 1999, Volatility in Emerging Stock Markets, Journal of Financial and Quantitative Analysis 34, 33-35.

[2] Blake, 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.

[3] Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroskedasticity, Journal of Econometrics 31, 307-327.

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

[5] Bose, A. and K. Mukherjee, 2003, Estimating the ARCH Parameters by Solving Linear Equations, Journal of Time Series Analysis 24, 127-136.

[6] Engle, R.F., 1982, Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation, Econometrica 50, 987-1007.

[7] Franses, P.H. and E. Ghijsels, 1999, Additive Outliers, GARCH and Forecasting Volatility, International Journal of Forecasting 15, 1-9.

[8] Franses, P.H. and D. van Dijk, 2000, Outlier detection in GARCH models, Econometric Institute Research report, EI-9926/RV.

[9] Franses, P.H., D. van Dijk and A. Lucas, 2004, Short Patches of Outliers, ARCH and Volatility Modelling, Applied Financial Economics 14, 221-231.

[10] Friedman, B.M. and D.I. Laibson, 1989, Economic implications of extraordinary movements in stock prices (with comments and discussion), Brooking Papers on Economic Activity, 20, 137-189.

[11] Gregory, A.W. and J.J. Reeves, 2001, Estimation and Inference in ARCH Models in the Presence of Outliers, Proceedings of the 2002 North American Summer Meetings of the Econometric Society.

[12] Hotta, L.K. and R.S. Tsay, 1998, Outliers in GARCH processes, manuscript.

[13] Lee, J.H.H. and M.L. King, 1993, A Locally Most Mean Powerful Based Score Test for ARCH and GARCH Regression Disturbances, Journal of Business and Economic Statistics 11, 17-27.

[14] Li, J. and C. Kao, 2002, A Bounded Influence Estimation and Outlier Detection for ARCG/GARCH Models with an Application to Foreign Exchange Rates, manuscript.

[15] Ling, S. and M. McAleer, 2003, Asymptotic Theory for a vector ARMA-GARCH model, Econometric Theory, 19, 280-310.

[16] Mendes, B.V.M., 2000, Assessing the Bias of Maximum Likelihood Estimates of Contaminated GARCH Models, Journal of Statistical Computation and Simulation 67, 359-376.

[17] 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.

[18] Muller, N. and V.J. Yohai, 2002, Robust Estimates for ARCH Processes, Journal of Time Series Analysis 23, 341-375.

[19] Peña, D., 2001, Outliers, influential observations and missing Data, in: Peña, D., G.C. Tiao and R.S. Tsay, eds., A course in Time Series, John Wiley, New York.

[20] Peña, D. and J. Rodriguez, 2002, A Powerful Portmanteau Test of Lack of Fit in Time Series, Journal of the American Statistical Association 97, 601-610.

[21] Sakata and White, 1998, High breakdown point conditional dispersion estimation with application to S&P500 daily returns volatility, Econometrica, 66, 529-567.

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

[23] Van Dijk, D., P.H. Franses and A. Lucas, 1999, Testing for ARCH in the Presence of Additive Outliers, Journal of Applied Econometrics 14, 539-562.

[24] Verhoeven, P. and M. McAleer, 2000, Modelling outliers and extreme observations for ARMA-GARCH processes, manuscript.

[25] Weiss, A., 1986, Asymptotic Theory for ARCH Models, Econometric Theory 2, 107-131.

[26] Wooldridge, J.M., 1990, A unified approach to robust regression based specification tests, Econometric Theory, 6, 17-43.

[27] Zhang, X., 2004, Assessment of local influence in GARCH processes, Journal of Time Series Analysis, 25, 301-313.