Short Patches of Outliers, ARCH and Volatility Modeling*

Philip Hans Franses†
Econometric Institute, Erasmus University Rotterdam

Dick van Dijk‡
Tinbergen Institute, Erasmus University Rotterdam

André Lucas§
Financial Sector Management, Vrije Universiteit Amsterdam and
Tinbergen Institute

May 18, 1998

Abstract

In this paper we test for (Generalized) AutoRegressive Conditional Heteroskedasticity [(G)ARCH] in daily data on 22 exchange rates and 13 stock market indices using the standard Lagrange Multiplier [LM] test for GARCH and a LM test that is resistant to patches of additive outliers. The data span two samples of 5 years ranging from 1986 to 1995. Using asymptotic arguments and Monte Carlo simulations, in which we evaluate our empirical method, we show that patches of outliers can have significant effects on test outcomes. Our main empirical result is that we find spurious GARCH in about 40% of the cases, while in many other cases we find evidence of GARCH even though such sequences of extraordinary observations seem to be present.

Keywords: Generalized AutoRegressive Conditional Heteroskedasticity, Lagrange Multiplier test, Outliers, Robust testing, Exchange rates, Stock market indices.

JEL classification: C12, C22, C52, G15

---

*This paper is a substantially revised and extended version of the paper 'Do we often find ARCH because of neglected outliers?', Econometric Institute report 9706/A, by the first two authors.

†Econometric Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR, Rotterdam, The Netherlands, `email: franses@few.eur.nl` (corresponding author)

‡Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, `email: djvandijk@few.eur.nl`

§Financial Sector Management, ECO/BFS, Vrije Universiteit, De Boelelaan 1105, NL-1081 HV, Amsterdam, The Netherlands, `email: alucas@econ.vu.nl`. Financial support from the Netherlands Organization for Scientific Research (NWO) is gratefully acknowledged.

1 Introduction and motivation

Asset returns are leptokurtic, that is, the tails of the distribution contain too many extreme observations to fit the normal distribution. Moreover, these extreme observations tend to cluster together over time. Periods of large price movements alternate with periods of relative tranquillity. This clustering of extraordinary movements in financial time series is supposed to correspond with time-varying volatility. Given the importance of volatility as a measure of risk in many financial decision problems, there has been a growing interest in describing and forecasting volatility during the last, say, fifteen years. By far the most popular models which are used for this purpose are the (Generalized) AutoRegressive Conditional Heteroskedasticity [(G)ARCH] models, developed in Engle (1982) and Bollerslev (1986). Recent surveys of the GARCH (and related) literature are provided by Bollerslev, Chou and Kroner (1992), Bollerslev, Engle and Nelson (1994), Diebold and Lopez (1995), Palm (1996), Shephard (1996), and Gourieroux (1997).

GARCH models have been shown to be able to replicate the salient features of asset returns mentioned above, i.e., they are able to generate time series which have unconditional distributions with fatter tails than the normal density and which display clustering of large realizations. Furthermore,

> ...temporal clustering of outliers can be used to predict their occurrence and minimize their effects. This is exactly the approach taken by the ARCH model. [Engle, 1982, p.992]

Obviously, in order to be able to exploit the apparent clustering of outliers for forecasting, as suggested by Engle (1982), or to estimate the parameters in the (G)ARCH model with some degree of precision, the data should display at least a few periods with high volatility. The same holds, for example, for nonlinear models for business cycles, which can only usefully be considered for data with at least two (but probably more) recessions. Put differently, when financial time series display only very few short patches of extreme observations, one may wonder if the consideration of elaborate (G)ARCH models is of practical relevance. For example, if there is only one patch of two sequential outliers, there are no degrees of freedom to estimate the three parameters in the popular GARCH(1,1) model, let alone to estimate nonlinear extensions of this model which contain even more parameters. Given this potential shortage of degrees of freedom, it makes sense to test for ARCH first instead of fitting an ARCH model right-away. Since one would want to avoid being alarmed because of only a limited number of short patches of outliers, one should preferably use a test for ARCH which is robust to such circumstances. It would then be sensible to consider (a version of) an ARCH model only when this robust ARCH test indicates its potential usefulness.

In this paper we apply the robust ARCH test developed in Van Dijk, Franses and Lucas (1998) to a large set of daily observed financial time series (over two disjoint samples of 5 years). Before we turn to this application, we investigate the effect of patches of outliers on the standard (and non-robust) test for ARCH put forward by Engle (1982). Our examination uses both theoretical (asymptotic) arguments in Section 2 and Monte Carlo simulations in Section 3. Our empirical findings, presented in Section 4, show that for several daily series, ARCH effects are due to only a few short patches of outliers (out of approximately 1250 observations). For about one half of the series analyzed we still find evidence of ARCH when the robust test is applied.

2 Testing for (G)ARCH

This section is divided into two parts. First, we briefly discuss the standard and robust Lagrange Multiplier [LM] tests for (G)ARCH. Secondly, we discuss the properties of the standard LM test in the presence of short patches of outliers.

2.1 The LM tests for (G)ARCH

Consider the AR(m)-GARCH(1,1) model for a time series y_t,

$$\begin{aligned} \phi(L)(y_t-\mu) &= \varepsilon_t, \qquad t=1,\ldots,T, & (1)\\ \varepsilon_t &= \sigma_t\eta_t, & (2)\end{aligned}$$

where $\phi(L) = 1-\phi_1 L - \ldots - \phi_m L^m$ is a polynomial in the lag operator L, defined as $L^k y_t = y_{t-k}$ for all integers k, having all roots outside the unit circle, $\{\eta_t\}_{t=1}^{T}$ is a sequence of independent identically distributed [i.i.d.] random variables with $E(\eta_t)=0$ and $E(\eta_t^2)=1$, and

$$\sigma_t^2 = \omega + \alpha\varepsilon_{t-1}^2 + \beta\sigma_{t-1}^2, \tag{3}$$

with $\omega > 0$, $\alpha, \beta \geq 0$ and $\alpha+\beta < 1$. These parameter restrictions are necessary and sufficient for the conditional variance of ε_t to be always positive and the unconditional variance σ^2, given by $\sigma^2 = \omega/(1-\alpha-\beta)$, to exist.

When $\beta = 0$ in (3), the model reduces to an ARCH(1) model, introduced in Engle (1982). This model can be extended straightforwardly to a general ARCH(q) model by including additional terms $\varepsilon_{t-2}^2, \ldots, \varepsilon_{t-q}^2$ in (3). Engle (1982) also derives the LM test for the null hypothesis of conditional homoskedastic errors against the alternative of ARCH(q). Lee (1991) shows that the LM test for GARCH(p,q) is the same as the test for ARCH(q) when $p \leq q$. The GARCH(p,q) model is obtained by adding $\varepsilon_{t-2}^2, \ldots, \varepsilon_{t-q}^2$ and $\sigma_{t-2}^2, \ldots, \sigma_{t-p}^2$ as regressors to (3). The LM test against ARCH(q) is given by

$$\xi = \frac{T\hat{f}'\hat{Z}(\hat{Z}'\hat{Z})^{-1}\hat{Z}'\hat{f}}{\hat{f}'\hat{f}}, \tag{4}$$

where $\hat{Z}' = (\hat{z}_1', \ldots, \hat{z}_T')$, $\hat{z}_t = (1, \hat{\varepsilon}_{t-1}^2, \ldots, \hat{\varepsilon}_{t-q}^2)'$, and $\hat{f}' = (\hat{f}_1', \ldots, \hat{f}_T')$, $\hat{f}_t = (\hat{\varepsilon}_t^2/\hat{\sigma}^2 - 1)$, with $\hat{\sigma}^2 = \sum_{t=1}^T \hat{\varepsilon}_t^2/T$ and $\hat{\varepsilon}_t$ denoting the least squares residuals from estimating the conditional mean equation (1) under the null of no ARCH by Ordinary Least Squares [OLS]. The LM test ξ can simply be computed as TR^2, where R^2 is the coefficient of determination of an auxiliary regression of the squared residuals $\hat{\varepsilon}_t^2$ on an intercept and $\hat{\varepsilon}_{t-1}^2$ through $\hat{\varepsilon}_{t-q}^2$. Under the null hypothesis of no (G)ARCH, this standard LM test is asymptotically χ^2 distributed with q degrees of freedom.

Now assume that additive outliers occur such that, instead of the 'clean' series y_t, one observes the series x_t which is equal to y_t plus an additive outlier [AO] process,

$$x_t = y_t + \zeta\delta_t, \tag{5}$$

where $\{\delta_t\}$ is a stochastic contamination process, which takes non-zero values with positive probability, and where $\zeta > 0$ is a non-zero constant indicating the magnitude of the AO's. The model for x_t implied by (1)-(3) and (5) is similar in spirit as the model proposed by Friedman and Laibson (1989). It should be noted that this model does not imply that all variation in and clustering of volatility of x_t is explained by outliers, only that an explicit distinction is made between ordinary and extraordinary movements.

It is well-known that OLS estimates of the mean and the autoregressive parameters are severely biased when AO's are neglected. Using asymptotic arguments, Van Dijk, Franses and Lucas (1998) (VDFL hereafter) show that, in case the AO's occur in isolation, this bias adversely affects both the size and power properties of the standard LM test for ARCH. They propose a modified LM test, which can be obtained by using an outlier robust estimator for the parameters in (1) instead of OLS. In addition to estimates of the parameters in the conditional mean equation, the (iterative) estimation procedure suggested by VDFL provides weights $\hat{w}_{\varepsilon_t}$, which actually identify the observations which are to be considered as outliers. When $\hat{w}_{\varepsilon_t} = 1$, the observation at time t is perfectly regular, while a weight smaller than one indicates that the observation does not match the properties of the bulk of the data. Obviously, if $\hat{w}_{\varepsilon_t} = 0$, the corresponding data point is an extreme outlier. It should be noted that the 'observation at time t' here refers to the composite of the regressand y_t and regressors $y_{t-1}, \ldots, y_{t-m}$, as aberrant values of either of those can cause the weight to be smaller than 1. The modified LM test for ARCH, denoted as ξ_R, is given by (4), with $\hat{z}_t$ and $\hat{f}_t$ now defined as $\hat{z}_t = (1, \hat{w}^2_{\varepsilon_{t-1}}\hat{\varepsilon}^2_{t-1}, \ldots, \hat{w}^2_{\varepsilon_{t-q}}\hat{\varepsilon}^2_{t-q})'$ and $\hat{f}_t = \hat{w}^2_{\varepsilon_t}\hat{\varepsilon}^2_t$, respectively, where it should be stressed that the $\hat{\varepsilon}_t$'s are the residuals from robust estimation. The outlier robust LM test statistic can simply be computed by forming weighted regression residuals and running an auxiliary regression of the squared weighted residuals on an intercept and q lags. The LM test is again equal to TR^2, using the R^2 from this auxiliary regression. VDFL show that this robust LM test for ARCH is asymptotically distributed as χ^2 with q degrees of freedom. Simulation results in VDFL show that the robust LM test has quite satisfactory size and power properties, already in samples as small as 100 observations, even when no outliers are present.

The analysis in VDFL is confined to the case of isolated AO's, which occur when the contamination process δ_t in (5) is characterized as $P(\delta_t = 0) = 1 - \pi$, $P(\delta_t = 1) = P(\delta_t = -1) = \pi/2$, with $0 < \pi < 1$. However, additive outliers can also cluster together or, put differently, occur in patches. A patch of k outliers occurs if we allow the δ_t to be autocorrelated as follows,

$$\delta_t = \begin{cases} \tilde{\delta}_t & \text{if } v_i \neq 0 \text{ for some } i = t-k+1, \ldots, t, \\ 0 & \text{else,} \end{cases} \tag{6}$$

with $\tilde{\delta}_t$ and v_t i.i.d., $P(\tilde{\delta}_t = 1) = P(\tilde{\delta}_t = -1) = 1/2$, $P(v_t = 0) = 1 - \pi$, and $P(v_t \neq 0) = \pi$. The effect of such patches on the asymptotic distribution of the standard LM test for ARCH is considered next.

2.2 The effect of patchy additive outliers on the standard LM test

In this subsection, we use asymptotic arguments to show that the occurrence of only a few adjacent AO's may result in spurious detection of ARCH effects. We study the effect of additive outliers that occur in patches of length k in a model that contains neither AR nor GARCH behavior, i.e., (1)-(3) with (5) and (6) with $\phi(L) = 1$ and $\mu = 0$ in (1), and $\alpha = \beta = 0$ in (3).

Deriving the effect of outliers on the asymptotic distribution of the ARCH test ξ in (4) is nontrivial, because 8-th order (cross-)moments of the different stochastic variables in the patchy AO model are involved. Therefore, instead of deriving the exact effects, we follow VDFL and concentrate on the effect of outliers on the noncentrality parameter of the asymptotic χ^2 distribution of the ARCH(1) test. To be more precise, we only look at the effect of outliers on the main determinant of this noncentrality parameter, namely the expectation of

$$\left(\frac{x_t^2}{\tilde{\sigma}^2} - 1\right) x_{t-1}^2, \tag{7}$$

where $\tilde{\sigma}^2$ is the probability limit of the OLS estimator of the variance of the regression errors, i.e., $\tilde{\sigma}^2 \equiv E(\hat{\varepsilon}_t^2)$. The noncentrality parameter is given by the squared expectation of (7) divided by the variance of (7). For the effect of patchy outliers on the estimate of the variance of the regression errors, $\tilde{\sigma}^2$, observe that

$$\begin{aligned}
\tilde{\sigma}^2 &= E(\hat{\varepsilon}_t^2) = E(x_t^2) \\
&= E(y_t^2) + \zeta^2 E(\delta_t^2) \\
&= \omega + \zeta^2(1 - P(v_{t-k+1} = 0, \ldots, v_t = 0)) \\
&= \omega + \zeta^2(1 - (1-\pi)^k). \qquad (8)
\end{aligned}$$

Hence, on average the variance is overestimated by $\zeta^2(1 - (1-\pi)^k)$. Furthermore, we obtain that

$$\begin{aligned}
E(x_t^2 x_{t-1}^2) &= E((y_t + \zeta\delta_t)^2 (y_{t-1} + \zeta\delta_{t-1})^2) \\
&= E(y_t^2 y_{t-1}^2 + y_t^2 \zeta^2 \delta_{t-1}^2 + y_{t-1}^2 \zeta^2 \delta_t^2 + \zeta^4 \delta_t^2 \delta_{t-1}^2) \\
&= \omega^2 + 2\omega\zeta^2(1 - (1-\pi)^k) + \zeta^4 \Big(P(\exists_{i \in \{t-k+1,\ldots,t-1\}} :\ v_i \neq 0) + \\
&\quad P(v_t \neq 0, v_{t-1} = 0, \ldots, v_{t-k+1} = 0, v_{t-k} \neq 0)\Big) \\
&= \omega^2 + 2\omega\zeta^2(1 - (1-\pi)^k) + \zeta^4 \left(1 - (1-\pi)^{k-1} + (1-\pi)^{k-1}\pi^2\right) \\
&= \omega^2 + 2\omega\zeta^2(1 - (1-\pi)^k) + \zeta^4 \left(1 - (1-\pi)^k (1+\pi)\right). \qquad (9)
\end{aligned}$$

As a result, the expectation of (7) can be written as

$$\frac{E(x_t^2 x_{t-1}^2) - \tilde{\sigma}^4}{\tilde{\sigma}^2} = \frac{\zeta^4 (1-\pi)^{k+1} (1 - (1-\pi)^{k-1})}{\omega + \zeta^2 (1 - (1-\pi)^k)}. \tag{10}$$

Expression (10) clearly demonstrates that unless there are no outliers, i.e., $\zeta = 0$ or $\pi = 0$, or only outliers, i.e., $\pi = 1$, the noncentrality parameter of the asymptotic distribution of the ARCH test statistic is nonzero. This results in a rejection frequency of the standard test above the nominal level, despite the absence of ARCH effects. Hence, AO's occurring in patches can result in a spurious detection of ARCH effects. This is intuitively clear, as additive outliers result in large values of the innovations. If several of such values occur in a row, the ARCH test is biased towards the detection of volatility clustering, i.e., large innovations following large innovations. If patches become very long ($k \to \infty$) the noncentrality parameter tends to zero again. Long patches of dominant outliers result in a distribution of the ARCH test close to its null distribution. Put differently, long patches lead to small size distortions. It can be shown, however, that the same phenomenon for the noncentrality parameter holds under the alternative of genuine ARCH effects, such that long patches of outliers asymptotically lead to a power loss of the ARCH test. This is again intuitively clear, because in such cases the homoskedastic white noise contamination will dominate the original ARCH signal, such that volatility clustering will go unnoticed. Notice that such a power loss might also occur even if the outliers do not dominate the ARCH signal. From (8) it is seen that outliers tend to inflate the estimate of the residual variance, which reduces the value of the test statistic ξ and, hence, reduces the power of the test.

The results which have been presented above can be generalized to models in which the model for the conditional mean contains autoregressive components, i.e., $m > 0$ in (1), and to tests against higher order ARCH alternatives. In the next section we investigate the performance of the standard and robust ARCH tests in small samples using Monte Carlo simulations.

3 Small sample properties of ARCH tests

In this section we add to the Monte Carlo experiments presented in VDFL some new simulation evidence on the small sample properties of the standard and robust ARCH tests. We examine the effects of both isolated and patchy AO's for a sample size which matches the length of the empirical time series which we analyze in Section 4. To investigate the effects of outliers, we employ two different data generating processes [DGP's]: first, a zero mean white noise process with homoskedastic errors [DGP I] and, second, a zero mean white noise process with GARCH(1,1) errors [DGP II]. For both DGP's, we generate 100 series of length 1250, which approximately corresponds with 5 years of daily data. The first 100 observations of each series are discarded to avoid dependence of our results on starting values. For each replication we record whether we find ARCH with both tests, denoted as (Y,Y), ARCH with the standard test but not with the robust test [(Y,N)], or one of the other combinations [(N,Y) or (N,N)]. These simulations will guide the interpretation of the empirical findings to be presented in the next section.

For both DGP's, we set $\mu = 0$ and $\phi(L) = 1$ in (1). In the GARCH(1,1) case, the parameters in the conditional variance equation (3) are set equal to values which are typically found for financial time series, $\alpha = 0.15, \beta = 0.80$. The intercept in this equation, ω, is set equal to 0.05, such that the unconditional variance of ε_t is equal to the variance of η_t. In the simulations, we investigate the effects of the magnitude and frequency of AO's as in (5). We examine isolated outliers of size $\zeta = 3$, 5, and, 7 which occur with probability $\pi = 0.01$ and 0.05, and in addition, patchy outliers of size $\zeta = 3$ and 5. Instead of using (6) to generate the patches, we opt for a more controlled experiment and add one or two patches of length $k = 2, 3$ or 5 to each series (occurring at random places in the series). All possible combinations of the characteristics of the contamination process render 18 different experiments per DGP. The effect of the distribution of the innovations is investigated by considering a Student t_ν distribution with degrees of freedom equal to $\nu = 5$ and ∞, the latter of course corresponding to normal errors. The t_5 errors are rescaled such that they have variance equal to 1. The assumption of t_ν distributed errors is quite common in applications of GARCH models to financial time series, where y_t typically is the return on a financial asset. Finally, the tests are computed for the uncontaminated series as well, to obtain estimates of their size and power. We evaluate all tests at the 5% significance level and use the asymptotic χ^2 critical values. To investigate the effect of lag length selection in the auxiliary ARCH test regressions, we set q equal to 1, 5 and 10. Another reason for considering tests against ARCH(q) errors with q larger than the true GARCH order is that a GARCH(1,1) process can be approximated by an ARCH(q) process with q sufficiently large. Throughout, the ARCH tests are applied to series from which the mean has been removed, either by estimating it with OLS or the robust estimator.

The results for DGP I with isolated outliers are reported in Table 1. From this table, several conclusions emerge. First of all, when applied to the clean series with normal errors, the empirical size of our robust test, which can be obtained by adding up the entries in the columns headed (Y,Y) and (N,Y), is quite satisfactory. It is comparable with the size of the standard test, which is given by the sum of the columns (Y,Y) and (Y,N). Second, the occurrence of AO's has quite different effects on the standard and robust tests. For the standard tests, the size increases if large outliers ($\zeta = 5, 7$) occur very rarely ($\pi = 0.01$). If outliers occur more frequently, the size returns to the nominal level. Note that this effect is most pronounced for $q > 1$. On the other hand, the size of the robust test is hardly affected. The results obtained when the errors are Student t_5 distributed are roughly comparable with the results for normally distributed innovations.

- insert Table 1 about here -

Results for the case of patchy outliers are set out in Table 2. In contrast to the limited impact of neglecting isolated AO's in white noise series on the standard ARCH test, it is seen that in case of clustering of AO's the standard LM test is affected to a much larger extent. For almost all combinations of n, k and ζ considered here, the test statistic is severely oversized. In fact, the empirical rejection frequency gets close to 100%, when there is, for example, only a single patch of 3 or 5 outliers (of magnitude 5) out of the 1250 observations. In general, the size distortion is smaller in case the errors are distributed as Student $t(5)$ and the outliers are not very large. In sharp contrast with these findings for the standard test, the empirical size of our robust test is usually close to the nominal 5% significance level.

- insert Table 2 about here -

The results for DGP II, which concern the empirical power of the tests, are shown in Tables 3 and 4, for isolated and patchy outliers, respectively. The entries for $\pi = 0$ and $\zeta = 0$ in Table 3 reveal that a disadvantage of the robust test is slight decrease in power when no outliers are present for the Student t_5 distribution. This illustrates that protection against aberrant observations sometimes comes at a cost. Notice however that this effect only occurs for the test with $q = 1$. If isolated outliers occur, the situation is completely reversed. In general, the power of the standard test decreases quite dramatically, while the power of the robust test remains high. As mentioned in Section 2.2, this is due to the fact that outliers inflate the estimate of the residual variance, which decreases the power of the standard test.

- insert Table 3 bout here -

In the presence of patchy outliers, the power of the robust test against ARCH(1) again suffers from a power loss for Student t_5 distributed errors, as can be seen from the relevant entries in Table 4. In all other cases, the power of both the standard and robust tests is very high.

- insert Table 4 about here -

The above simulation results clearly suggest how the outcomes of the standard and robust ARCH tests can be interpreted in practice. First of all, it seems best to focus on the outcomes for $q = 5$ or 10 since this minimizes the power loss from using the robust test. Given this focus, it is seen that if the robust test finds no ARCH, there likely is no ARCH, and when it finds ARCH, there likely is. Furthermore, the result (Y,N), meaning finding ARCH with the standard test but not with the robust test, can most likely be seen as evidence against ARCH in favor of one or a few short sequences of extraordinary observations. The opposite result (N,Y), meaning finding ARCH with the robust test but not with the standard test, can be interpreted as evidence of ARCH, possibly contaminated with a few isolated outliers. In both cases we recommend to have a closer look at the weights from the robust regression and the corresponding observations in the original time series, before carrying on with any subsequent analyses. A possible way to proceed is to downweight those observations which are found to be extreme outliers, and to estimate the GARCH model for the cleaned series. Alternatively, one might use an iterative outlier-detection method, in which one alternates between removing outliers and estimating GARCH (or other) models, see Franses and Ghijsels (1998) and Hotta and Tsay (1998). The forecasting results in Franses and Ghijsels (1998) show that volatility forecasts from GARCH models of series which have been cleaned in this manner can dramatically improve upon those from models for contaminated data. Yet another approach would be to use robust estimation methods to estimate GARCH models.

4 ARCH in exchange rate and stock market returns

In this section we illustrate the use of our robust ARCH test in practice by examining 35 daily financial time series. In Section 4.1 we discuss the data and our empirical research methodology, and in Section 4.2 we report our empirical findings.

4.1 Data and methodology

We consider 35 financial time series, which are sampled daily over a ten-year period ranging from 1986 to 1995. The 35 series can be grouped into 22 exchange rates (versus the US dollar) and 13 stock market indices. The exchange rates concern the Australian dollar, Austrian shilling, Belgian franc, British pound, Canadian dollar, Danish kroner, Dutch guilder, ECU, Finnish markka, French franc, German DMark, Greek drachme, Irish pound, Japanese yen, Malaysian ringgit, New Zealand dollar, Norwegian kroner, Singapore dollar, South African rand, Spanish peseta, Swedish kroner, and the Swiss franc. The stock markets concern those in Amsterdam (CBS), Brussels (BSE), Frankfurt (DAX), Hong Kong (Hang Seng), London (FTSE), Madrid (MSE), Milan (MC), New York (Dow Jones and S&P500), Singapore, Stockholm (VEC), Taipei, and Tokyo (Nikkei). We compute our tests for two subsamples, 1986-1990 and 1991-1995, each of which contain 5 years of data. In terms of daily data, this means that we have samples of about 1250 observations. Given this choice to split the sample, our first sample contains the 1987 stock market crash. Results for other samples, which turn out to be qualitatively similar, are available upon request from the corresponding author. We apply the LM tests for ARCH with q equal to 1, 5 and 10, although we will focus the discussion of the results on the latter two values. The tests are applied to both the raw and 'prewhitened' series. In both cases the series are demeaned ('demedianed') first by subtracting the mean (median) before applying the standard (robust) test to the series. Daily means and medians are considered in order to allow for possible day-of-the-week effects, cf. Baillie and Bollerslev (1989). The prewhitened series are obtained as the residuals from fitting an AR model of order 5. We also compute the tests when the AR order in (1) is selected by the Akaike and Schwarz Information Criteria. This does not yield qualitatively different results, and in order to save space we do not report them. All tests are evaluated at the 5% significance level. Similar results are obtained using 1% and 10% significance levels, which therefore are not displayed. We summarize our empirical findings by recording the number of times both tests find ARCH [(Y,Y)], the standard test finds ARCH while the robust test does not [(Y,N)] and vice versa [(N,Y)], and when both tests do not find ARCH [(N,N)].

4.2 Results

Table 5 presents some general results for the raw and prewhitened series. Using the guidelines obtained from the Monte Carlo experiments in Section 3, interpreting the outcomes is straightforward when we focus on $q = 5$ and $q = 10$. Based on the robust ARCH test, we find evidence in favor of ARCH in little over 50% of the cases, which can be seen by summing the entries in the columns headed (Y,Y) and (N,Y). Additionally, the robust test does not find ARCH while the standard test does in about 40% of the cases (columns (Y,N) and (Y,N)/((Y,Y)+Y,N))). This suggests that ARCH may quite often be caused by the occurrence of only one or a few short clusters of outliers. For the time series which belong to this last group, it may not be possible to exploit this clustering for forecasting or there may not even be enough degrees of freedom to estimate the parameters in (versions of) a GARCH(1,1) model.

- insert Table 5 about here -

In Table 6, the results for the tests applied to the raw series are displayed at a more disaggregated level, by focusing on the two subsamples and exchange rates and stock market returns individually. Only results for $q = 5$ are displayed, the corresponding findings for $q = 10$ look very much alike. It appears that the evidence of ARCH is a bit more convincing in the 1986-1990 sample, as the entries in the (Y,Y) column in general are larger than those for the subsample comprising 1991-1995.

- insert Table 6 about here -

To illustrate that exceptional outliers are important, consider the final two columns of this table, which contain percentages of observations which receive weights $\hat{w}_{\varepsilon_t}$ equal to zero and equal to one in the robust estimation procedure which is used to compute the robust LM statistic. It can be seen from the last column that between 4.4% and 7.7% of the observations are downweighted, i.e., obtain a weight smaller than one, of which between 2.6% and 4.9% of the observations are discarded completely (which is what effectively is done when an observation receives weight zero). In the other columns of the table, these percentages are shown based on subsets of the series for which a particular is obtained. test outcome. For example, the entry 3.9 in the block on exchange rates for the sample period 1986-1990, in the column (Y,N) and row $\hat{w}_{\varepsilon_t} = 0$, means that for the 10 series for which the standard test indicates ARCH and the robust test does not, on average 3.9% of the observations receives a weight smaller than 0.05 in the robust estimation procedure. These percentages not only show that indeed aberrant observations which may cause spurious evidence of GARCH are found by the robust method, but also that outliers are found in case there is GARCH. Recall that the model for the series x_t as specified in (1)-(3) and (5) does not imply that all apparent ARCH effects can be explained by outliers, but rather that the observed time series consists of a regular component y_t, which may very well exhibit GARCH, and an irregular component $\zeta\delta_t$. When estimating the parameters of the regular GARCH component one should take the irregular component into account, using, for example, the methods outlined in Friedman and Laibson (1989), Franses and Ghijsels (1998) and Hotta and Tsay (1998).

Table 7 presents our empirical results in even more detail, by showing the outcomes of the standard and robust tests against ARCH(q) with $q = 5$ for the individual series. In addition, the table contains the number of patches of outliers of length $k = 2, \ldots, 5$ which are present in the series, where as before an observation is considered to be an outlier if it receives a weight $\hat{w}_{\varepsilon_t}$ smaller than 0.05.

- insert Table 7 about here -

Certain series, such as the Malaysian ringgit and the South African rand, contain quite a few patches of 2 extraordinary observations. For all series, sequences of more than two outliers are very rare. This substantiates our claim that the evidence of ARCH which is obtained from applying standard techniques may often be due to only very few aberrant observations which happen to be clustered over time.

Detailed inspection of the dates of occurrence of the indicated patches reveals several interesting observations. For the stock market series, most of the patches in the first subsample are concentrated around the turmoil in October 1987 and October 1989. In fact, many series seem to have 'common patches' during these periods. The Asian stock indices (Hong Kong, Tokyo, Singapore and Taipei) also contain common patches in August 1990. Concerning the exchange rate series, the Austrian shilling, Belgian franc, Danish kroner, Greek drachme, Dutch guilder, Finnish markka, Spanish peseta, South African rand, Swedish kroner and Irish pound have a common patch at the end of March 1986. The British pound, French franc, Dutch guilder, Belgian franc, Finnish markka, Spanish peseta have a common patch in the beginning of June 1989. In the second subsample, common patches for many of the European currencies occur in the second half of September 1992, during the speculative attack on the French franc and the subsequent demise of the EMS exchange rate system. Finally, the German DMark, British pound, Japanese yen, Swiss franc, Dutch guilder, Irish pound, Singapore dollar, Malaysian ringgit, South African rand have sequence of extraordinary observations in common in the beginning of March 1995. In our subsequent work we will study the construction of multivariate GARCH models while taking care of the possible occurrence of common patches of outliers.

5 Conclusions

In this paper we have explored the possibility that apparent ARCH in daily financial time series is caused by only a few sequences of aberrant observations. By comparing the outcomes of standard and robust tests for ARCH for a large number of empirical time series with results from extensive Monte Carlo simulations, we conclude that isolated outliers are unlikely to suggest spurious ARCH but that short patches of outliers do. We therefore recommend to apply both the standard and robust tests prior to the estimation of GARCH models, as the joint outcome of these tests might be informative with regard to the appropriateness of this class of models and the presence of outliers. In case only the standard test finds ARCH, the possibility of one or a few patches of AO's instead of ARCH has to be given some serious thought. If the robust test rejects the null hypothesis and the estimated weights indicate a few patches of outliers, it might be worthwhile to use a robust estimation method for ARCH models in order to reduce their potential impact on parameter estimates.

References

Baillie, R.T. and T. Bollerslev, 1989, The message in daily exchange rates: a conditional variance tale, *Journal of Business and Economic Statistics* **7**, 297–305.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics* **31**, 307–327.

Bollerslev, T., R.F. Engle and D.B. Nelson, 1994, ARCH models, in R.F. Engle and D.L. McFadden (editors), *Handbook of Econometrics IV*, Elsevier Science, Amsterdam, pp. 2961–3038.

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**, 5–59.

Diebold, F.X. and J.A. Lopez, 1995, Modelling volatility dynamics, in K. Hoover (editor), *Macroeconometrics - Developments, Tensions and Prospects*, Kluwer, Boston, pp. 181–198.

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

Franses, P.H. and H. Ghijsels, 1998, Additive outliers, GARCH and forecasting volatility, *International Journal of Forecasting* (forthcoming).

Friedman, B.M. and D.I. Laibson, 1989, Economic implications of extarordinary movements in stock prices (with comments and discussion), *Brookings Pappers on Economic Activity* **20**, 137–189.

Gourieroux, C., 1997, *ARCH Models and Financial Applications*, Springer-Verlag, Berlin.

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

Lee, J.H.H., 1991, A Lagrange multiplier test for GARCH models, *Economics Letters* **37**, 265–271.

Palm, F.C., 1996, GARCH models of volatility, in G.S. Maddala and C.R. Rao (editors), *Handbook of Statistics, vol. 14*, Elsevier, Amsterdam, pp. 209–240.

Shephard, N., 1996, Statistical aspects of ARCH and stochastic volatility, in O.E. Barndorff-Nielsen D.R. Cox and D.V. Hinkley (editors), *Statistical models in econometrics, finance and other fields*, Chapman and Hall, London, pp. 1–67.

Van Dijk, D., P.H. Franses and A. Lucas, 1998, Testing for ARCH in the presence of additive outliers, Econometric Institute Report 9659/rv, Erasmus University Rotterdam (revised version, submitted for publication).

Table 1: Evidence of ARCH - Empirical size in the presence of isolated outliers[1]

			$\nu=\infty$				$\nu=5$			
π	ζ	q	(Y,Y)	(Y,N)	(N,Y)	(N,N)	(Y,Y)	(Y,N)	(N,Y)	(N,N)
.0	0	1	2	6	5	87	0	2	4	94
		5	1	3	2	94	0	4	6	90
		10	2	3	2	93	0	5	6	89
.01	3	1	1	2	6	91	0	1	6	93
		5	1	3	3	93	0	9	5	86
		10	0	1	5	94	0	10	7	83
	5	1	0	3	7	90	0	1	5	94
		5	0	13	3	86	0	9	5	86
		10	0	9	4	87	0	9	7	84
	7	1	0	5	8	87	0	4	5	91
		5	0	14	3	83	0	9	5	86
		10	0	13	4	83	0	8	7	85
.05	3	1	0	5	3	92	0	0	8	92
		5	0	3	6	91	0	2	6	92
		10	1	2	5	92	0	8	6	86
	5	1	0	3	6	91	0	5	5	90
		5	0	3	2	95	0	4	5	91
		10	0	5	3	92	0	8	4	88
	7	1	1	4	7	88	0	6	5	89
		5	0	5	4	91	0	4	4	92
		10	0	3	6	91	0	9	5	86

[1] Evidence of ARCH using the standard LM test and using the robust LM test, developed by VDFL. The table is based on 100 replications for sample size $T = 1250$, which roughly corresponds with 5 years of daily data. The cells report the number of times that a certain outcome occurs when the test statistics are evaluated at the 5% significance level. For example, (Y,N) means that the standard LM test detects ARCH (Y) while the robust test does not (N). The series are generated according a homoskedastic white noise process drawn from a Student t distribution with ν degrees of freedom. Isolated outliers of absolute magnitude ζ occur with probability π. The q is the number of lags included in the auxiliary ARCH test regressions, presented in Section 2.1.

Table 2: Evidence of ARCH - Empirical size in the presence of patchy outliers[1]

				$\nu=\infty$				$\nu=5$			
n	k	ζ	q	(Y,Y)	(Y,N)	(N,Y)	(N,N)	(Y,Y)	(Y,N)	(N,Y)	(N,N)
1	2	3	1	4	14	3	79	0	9	4	87
			5	3	12	3	82	0	12	4	84
			10	2	11	1	86	1	16	1	82
		5	1	7	88	0	5	3	44	1	52
			5	4	81	1	14	1	32	4	63
			10	4	76	0	20	1	33	1	65
	3	3	1	9	38	1	53	0	15	3	82
			5	5	31	2	64	1	14	5	80
			10	6	26	3	68	1	17	1	81
		5	1	7	93	0	0	4	80	0	16
			5	5	94	0	1	1	80	2	17
			10	3	94	1	2	1	74	1	24
	5	3	1	8	70	0	22	2	29	1	68
			5	5	73	2	20	2	28	4	66
			10	5	67	0	28	1	31	0	68
		5	1	7	93	0	0	3	94	1	3
			5	4	96	0	0	6	91	0	3
			10	5	95	0	0	1	94	0	5
2	2	3	1	6	34	0	60	1	14	2	83
			5	1	19	4	76	1	15	6	78
			10	1	14	5	80	1	14	0	85
		5	1	7	93	0	0	3	82	1	14
			5	4	95	0	1	5	66	1	28
			10	6	92	0	2	1	67	0	32
	3	3	1	6	66	1	27	1	22	3	74
			5	4	66	2	28	0	24	4	72
			10	3	59	2	36	2	24	1	73
		5	1	8	92	0	0	4	93	0	3
			5	6	94	0	0	4	90	1	5
			10	4	96	0	0	1	90	0	9
	5	3	1	5	93	0	2	1	61	0	38
			5	4	94	0	2	6	52	1	41
			10	5	90	0	5	1	54	2	43
		5	1	6	94	0	0	2	97	0	1
			5	5	95	0	0	6	93	0	1
			10	6	94	0	0	2	97	0	1

[1] Evidence of ARCH using the standard LM test and using the robust LM test, developed in VDFL. The table is based on 100 replications for sample size $T = 1250$, which roughly corresponds with 5 years of daily data. The cells report the number of times that a certain outcome occurs when the test statistics are evaluated at the 5% significance level. For example, (Y,N) means that the standard LM test detects ARCH (Y) while the robust test does not (N). The series are generated according to a homoskedastic white noise process drawn from a Student t distribution with ν degrees of freedom. n patches of k outliers of magnitude ζ are added at random places in the series. The q is the number of lags included in the auxiliary ARCH test regressions, presented in Section 2.1.

Table 3: Evidence of ARCH - Empirical power in the presence of isolated outliers[1]

			$\nu=\infty$				$\nu=5$			
π	ζ	q	(Y,Y)	(Y,N)	(N,Y)	(N,N)	(Y,Y)	(Y,N)	(N,Y)	(N,N)
.0	0	1	95	5	0	0	76	19	3	2
		5	100	0	0	0	93	6	1	0
		10	100	1	0	0	91	9	0	0
.01	3	1	93	3	3	1	70	18	10	2
		5	99	0	1	0	90	7	3	0
		10	98	1	1	0	89	10	1	0
	5	1	59	2	37	2	54	13	23	10
		5	78	0	22	0	72	6	21	1
		10	72	0	28	0	70	9	21	0
	7	1	18	0	78	4	32	4	45	19
		5	34	0	66	0	52	4	41	3
		10	37	0	63	0	48	6	43	3
.05	3	1	63	3	24	10	58	14	22	6
		5	83	0	16	1	79	5	15	1
		10	83	1	16	0	79	6	14	1
	5	1	18	0	77	5	30	4	52	14
		5	32	0	68	0	40	0	56	4
		10	23	0	77	0	41	1	51	7
	7	1	5	0	90	5	14	3	69	14
		5	12	0	88	0	17	0	79	4
		10	9	0	91	0	21	1	72	6

[1] Evidence of ARCH using the standard LM test and using the robust LM test, developed in VDFL. The table is based on 100 replications for sample size $T = 1250$, which roughly corresponds with 5 years of daily data. The cells report the number of times that a certain outcome occurs when the test statistics are evaluated at the 5% significance level. For example, (Y,N) means that the standard LM test detects ARCH (Y) while the robust test does not (N). The series are generated according to a GARCH(1,1) process with parameters $\alpha = 0.15$, $\beta = 0.8$, $\omega = 1/(1-\alpha-\beta)$ and t_ν distributed innovations with variance normalized to 1. Isolated additive outliers of magnitude ζ occur with probability π. The q is the number of lags included in the auxiliary ARCH test regressions used in computing the LM statistics.

Table 4: Evidence of ARCH - Empirical power in the presence of patchy outliers[1]

				$\nu = \infty$				$\nu = 5$			
n	k	ζ	q	(Y,Y)	(Y,N)	(N,Y)	(N,N)	(Y,Y)	(Y,N)	(N,Y)	(N,N)
1	2	3	1	96	4	0	0	69	28	2	1
			5	100	0	0	0	91	9	0	0
			10	100	0	0	0	93	7	0	0
		5	1	97	3	0	0	70	30	0	0
			5	100	0	0	0	91	9	0	0
			10	100	0	0	0	93	7	0	0
	3	3	1	95	5	0	0	73	26	1	0
			5	100	0	0	0	96	4	0	0
			10	100	0	0	0	94	6	0	0
		5	1	96	4	0	0	74	24	0	0
			5	100	0	0	0	96	4	0	0
			10	100	0	0	0	95	5	0	0
	5	3	1	96	4	0	0	75	24	1	0
			5	100	0	0	0	91	8	1	0
			10	100	0	0	0	91	9	0	0
		5	1	96	0	0	0	77	22	1	0
			5	100	0	0	0	91	8	1	0
			10	100	0	0	0	91	9	0	0
2	2	3	1	96	4	0	0	59	40	1	0
			5	100	0	0	0	91	9	0	0
			10	100	0	0	0	91	9	0	0
		5	1	96	4	0	0	59	40	1	0
			5	100	0	0	0	92	8	0	0
			10	100	0	0	0	91	9	0	0
	3	3	1	95	5	0	0	74	23	2	1
			5	100	0	0	0	93	7	0	0
			10	100	0	0	0	92	7	0	1
		5	1	96	4	0	0	76	24	0	0
			5	100	0	0	0	94	6	0	0
			10	100	0	0	0	92	8	0	0
	5	3	1	97	3	0	0	74	26	0	0
			5	100	0	0	0	93	7	0	0
			10	100	0	0	0	94	6	0	0
		5	1	97	3	0	0	77	23	0	0
			5	100	0	0	0	93	7	0	0
			10	100	0	0	0	94	6	0	0

[1] Evidence of ARCH using the standard LM test and using the robust LM test, developed in VDFL. The table is based on 100 replications for sample size $T = 1250$, which roughly corresponds with 5 years of daily data. The cells report the number of times that a certain outcome occurs when the test statistics are evaluated at the 5% significance level. For example, (Y,N) means that the standard LM test detects ARCH (Y) while the robust test does not (N). The series are generated according to a GARCH(1,1) process with parameters $\alpha = 0.15$, $\beta = 0.8$, $\omega = 1/(1-\alpha-\beta)$ and Student t_ν distributed innovations, with variance normalized to 1. n patches of k outliers of magnitude ζ are added. The q is the number of lags included in the auxiliary ARCH test regressions used in computing the LM statistics.

Table 5: Evidence of ARCH in exchange rate and stock market data, some overall results[1]

	q	(Y,Y)	(Y,N)	(N,Y)	(N,N)	(Y,N)/((Y,Y)+(Y,N))
Raw series						
	1	18	41	2	9	0.70
	5	36	27	3	4	0.43
	10	38	25	4	3	0.40
Prewhitened series						
	1	25	33	2	10	0.57
	5	36	28	2	4	0.44
	10	36	28	3	3	0.44

[1] Evidence of ARCH using the standard LM test and using the robust LM test, developed in VDFL. The cells report the number of times (out of 70 cases: 2 samples of 22 exchange rates and 13 stock indices) a certain outcome appears when the test statistics are evaluated at the 5% significance level. For example, (Y,N) means that the standard LM test detects ARCH (Y) while the robust test does not (N). (Y,N)/((Y,Y)+(Y,N)) denotes the frequency that the robust ARCH test does not find ARCH while the standard ARCH test does. The data and the empirical methodology are presented in Section 4.1.

Table 6: Evidence of ARCH in exchange rate and stock market data[1]

Sample		(Y,Y)	(Y,N)	(N,Y)	(N,N)	$\hat{w}_{\varepsilon_t} = 0$	$\hat{w}_{\varepsilon_t} = 1$
Exchange rates							
1986-1990	No.	10	10	1	1	4.9	92.3
	$\hat{w}_{\varepsilon_t} = 0$	6.3	3.9	2.6	2.8		
	$\hat{w}_{\varepsilon_t} = 1$	90.2	93.8	95.4	95.7		
1991-1995	No.	8	11	1	2	2.6	94.4
	$\hat{w}_{\varepsilon_t} = 0$	3.7	3.4	3.6	2.8		
	$\hat{w}_{\varepsilon_t} = 1$	94.3	94.5	94.6	94.2		
Stock markets							
1986-1990	No.	10	3	0	0	4.9	92.3
	$\hat{w}_{\varepsilon_t} = 0$	4.9	3.2	–	–		
	$\hat{w}_{\varepsilon_t} = 1$	93.3	95.1	–	–		
1991-1995	No.	8	3	1	1	2.6	94.4
	$\hat{w}_{\varepsilon_t} = 0$	3.1	2.2	2.6	1.5		
	$\hat{w}_{\varepsilon_t} = 1$	95.2	96.3	95.2	97.2		

[1] Evidence of ARCH using the standard LM test and using the robust LM test, developed in VDFL. The blocks of three cells in columns 3-6 report the number of times (out of 22 cases for the exchange rates and 13 cases of the stock market indices) a certain outcome appears when the test statistics against ARCH(q) with $q = 5$ are evaluated at the 5% significance level, the percentage of observations which receive weights smaller than 0.05 ($\hat{w}_{\varepsilon_t} = 0$) and larger than 0.95 ($\hat{w}_{\varepsilon_t} = 1$) in the robust estimation procedure, when read from top to bottom. The final two columns give those percentage averaged over all test outcomes. The data and the empirical methodology are presented in Section 4.1. Results are presented for the raw series.

Table 7: Evidence of patches of outliers in exchange rate and stock market data[1]

	1986-1990								1991-1995					
Exchange rate	(ξ, ξ_R)	k	2	3	4	5	6	7	(ξ, ξ_R)	k	2	3	4	5
Australian dollar	(Y,Y)		8	1	1	0	0	0	(Y,Y)		4	0	0	0
Austrian shilling	(Y,N)		7	0	0	0	0	0	(Y,Y)		6	0	0	0
Belgian franc	(Y,N)		6	2	0	0	0	0	(Y,N)		0	0	0	0
British pound	(N,Y)		1	0	0	0	0	0	(Y,Y)		7	0	0	0
Canadian dollar	(Y,N)		3	0	1	0	0	0	(Y,Y)		1	0	0	0
Danish kroner	(Y,Y)		3	0	0	0	0	0	(Y,N)		2	0	0	0
Dutch guilder	(Y,N)		4	0	0	0	0	0	(Y,N)		3	0	1	0
ECU	(Y,Y)		2	0	0	0	0	0	(Y,N)		5	0	0	1
Finnish markka	(Y,Y)		3	0	0	0	0	0	(N,N)		3	1	1	0
French franc	(Y,N)		6	0	0	0	0	0	(Y,N)		3	0	0	0
German DMark	(N,N)		1	0	0	0	0	0	(Y,N)		2	0	0	0
Greek drachme	(Y,Y)		15	3	0	0	0	0	(Y,N)		2	1	0	0
Irish pound	(Y,Y)		9	0	0	0	0	0	(N,N)		7	0	0	0
Japanes yen	(Y,N)		1	1	0	0	0	0	(N,Y)		3	0	0	0
Malaysian ringgit	(Y,Y)		14	2	3	2	1	0	(Y,Y)		14	3	1	0
New Zealand dollar	(Y,N)		10	0	1	0	0	0	(Y,Y)		3	1	0	0
Norwegian kroner	(Y,N)		2	0	0	0	0	0	(Y,Y)		2	0	1	0
South African rand	(Y,Y)		13	0	1	0	0	0	(Y,N)		12	1	2	0
Singapore dollar	(Y,N)		8	1	0	1	0	0	(Y,N)		7	2	1	0
Spanish peseta	(Y,Y)		23	8	1	0	0	0	(Y,N)		2	0	0	0
Swedish kroner	(Y,Y)		2	0	0	0	0	0	(Y,Y)		2	0	0	0
Swiss franc	(Y,N)		3	0	0	0	0	0	(Y,N)		2	1	0	0

	1986-1990								1991-1995					
Stock market	(ξ, ξ_R)	k	2	3	4	5	6	7	(ξ, ξ_R)	k	2	3	4	5
Amsterdam	(Y,Y)		3	1	3	0	0	0	(N,N)		0	0	0	0
Brussels	(Y,Y)		10	3	1	0	0	1	(N,Y)		3	0	0	0
Frankfurt	(Y,Y)		3	0	1	0	0	0	(Y,Y)		2	0	0	0
Hong Kong	(Y,Y)		4	5	0	0	0	0	(Y,Y)		4	2	1	0
London	(Y,N)		2	0	1	0	0	0	(Y,N)		1	0	0	0
Madrid	(Y,Y)		10	2	1	0	0	1	(Y,Y)		3	1	0	0
Milan	(Y,Y)		7	1	0	1	0	0	(Y,Y)		1	0	0	0
New York - Dow Jones	(Y,N)		6	0	0	0	0	1	(Y,N)		1	0	0	0
New York - S&P500	(Y,N)		4	0	0	0	0	1	(Y,N)		2	0	0	0
Singapore	(Y,Y)		9	3	0	0	1	0	(Y,Y)		6	0	1	0
Stockholm	(Y,Y)		2	1	2	0	1	0	(Y,Y)		2	2	0	0
Taipei	(Y,Y)		6	1	0	1	0	0	(Y,Y)		8	1	0	0
Tokyo	(Y,Y)		5	7	0	1	0	0	(Y,Y)		1	0	1	0

[1] Evidence of patches of outliers in exchange rate data and stock market data. The column headed (ξ, ξ_R) denotes the result from applying the standard LM test for ARCH and the robust LM test, developed in VDFL, when evaluated at the 5% significance level. For example, (Y,N) means that the standard LM test detects ARCH (Y) while the robust test does not (N). The remaining columns constain the number of patches of outliers of length k that occur during the sample period. An observation is considered as an outlier if it receives a weight smaller than 0.05 in the robust estimation procedure, which is used in computing the robust LM test. Results are presented for the raw series.