Testing for Changes in Volatility in Heteroskedastic Time Series - A Further Examination

Michiel de Pooter*
Tinbergen Institute
Erasmus University Rotterdam

Dick van Dijk†
Econometric Institute
Erasmus University Rotterdam

Econometric Institute Report EI 2004-38

August 2004

Abstract

We consider tests for sudden changes in the unconditional volatility of conditionally heteroskedastic time series based on cumulative sums of squares. When applied to the original series these tests suffer from severe size distortions, where the correct null hypothesis of no volatility change is rejected much too frequently. Applying the tests to standardized residuals from an estimated GARCH model results in good size and reasonable power properties when testing for a single break in the variance. The tests also appear to be robust to different types of misspecification. An iterative algorithm is designed to test sequentially for the presence of multiple changes in volatility. An application to emerging markets stock returns clearly illustrates the properties of the different test statistics.

Key words: change-point tests, structural breaks, CUSUM, GARCH models, emerging markets

JEL Classification Code: C12, C22, G15

*Tinbergen Institute, Erasmus University Rotterdam, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: depooter@few.eur.nl

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

1 Introduction

Quite soon after the interest in modelling the conditional heteroskedasticity of financial time series variables developed in the early 1980s, the possibility was raised that these variables experience occasional large shifts in unconditional volatility, see Diebold (1986) and Lamoureux and Lastrapes (1990). While the issue of testing for changes in the unconditional variance of time series has received considerable attention in the literature, dating back to at least Wichern *et al.* (1976), most of the available testing procedures implicitly or explicitly assume constant conditional volatility, see Hsu (1977), Talwar and Gentle (1981), Sakata (1988), Inclán and Tiao (1994), and Chen and Gupta (1997), among others. Recently, however, Kim *et al.* (2000) and Kokoszka and Leipus (2000) have developed tests that use cumulative sums of squares (CUSUMs) to test for breaks in the unconditional variance of possibly heteroskedastic time series. The properties of the CUSUM statistic of Kokoszka and Leipus (2000) were examined by Andreou and Ghysels (2002, 2004), finding that the test has good power properties but also noting some problems, in that the test sometimes suffers from quite large size distortions. The purpose of this paper is to examine this deficiency of these CUSUM tests in more detail. In particular, we investigate whether we can fix the size properties by adopting the suggestion of Lee *et al.* (2003) to apply the CUSUM statistic to standardized residuals from an estimated GARCH model. An elaborate simulation analysis confirms that the tests have severe size distortions when applied to the original series, such that the correct null hypothesis of no change is rejected much too frequently, rendering the tests highly unreliable. However, when applied to standardized GARCH residuals, the tests are found to have only minor size distortions and reasonably good power in detection volatility changes. Furthermore, the tests appear to be be quite robust to various types of misspecification. We apply the testing procedures to examine breaks in the unconditional volatility of a set of emerging stock market returns. Doing so allows to further assess the properties of the CUSUM tests and to compare the obtained results with earlier studies such as Aggarwal *et al.* (1999).

The outline of the paper is as follows. In Section 2, we discuss the three CUSUM tests that we consider in this study. In particular we demonstrate that all three tests emanate from the same basic setup. We also pay considerable attention to testing for multiple breaks and to the use of finite sample critical values. In Section 3, we use extensive Monte Carlo experiments to assess the size of the tests and their power for detecting both single as well as multiple breaks. We find that the tests, when applied to standardized returns, work reasonably well under different data generating processes and have quite good power properties. In Section 4 we apply the tests to daily emerging stock market returns. We find that the tests are certainly equipped for detecting variance changes in these series but that the results should be interpreted carefully, as two of the tests seem to have have a tendency to be conservative and potentially underestimate the number of actual breaks. Section 5 concludes.

2 CUSUM tests for changes in volatility

The issue that we want to address in this paper concerns testing for changes in the unconditional variance of a time series variable, in particular in the presence of conditional heteroscedasticity. Let $\{y_t\}_{t=1}^T$ denote the time series of interest with T being the available sample size, and assume (for simplicity, but without loss of generality) that y_t has a constant mean equal to 0. We consider the problem of testing the null hypothesis that the unconditional variance of y_t is constant, that is $\mathsf{H}_0 : \sigma_t^2 = \sigma^2$ for all $t = 1, \ldots, T$, against the alternative hypothesis of a single structural break, that is

$$\mathsf{H}_a : \sigma_t^2 = \begin{cases} \sigma_0^2 & \text{for } t = 1, \ldots, \kappa, \\ \sigma_1^2 & \text{for } t = \kappa + 1, \ldots, T, \end{cases} \tag{1}$$

where the change-point κ is assumed unknown. Many different approaches for tackling this testing problem have been developed, see the references in the Introduction. Here we limit ourselves to test statistics based on cumulative sums of squares (CUSUMs), as first proposed by Inclán and Tiao (1994) and subsequently further developed by Kim *et al.* (2000), Kokoszka and Leipus (2000), and Lee and Park (2001). In this section we first discuss the design of the CUSUM statistics for the above single break testing problem. In particular, we demonstrate that all tests, which might appear to be quite different at first sight, nevertheless fit into a single framework. Next, we address the problem of testing for multiple breaks in volatility. We conclude this section with a discussion on the use of finite sample critical values.

2.1 Testing for a single structural change

Our starting point is the cumulative sum of squares (CUSUM) process $C_y(k) = \sum_{t=1}^k y_t^2$. The centered and normalized CUSUM process is then defined as

$$D_y(k) \equiv \frac{1}{\sqrt{T}} \sum_{t=1}^{k} y_t^2 - \frac{k}{T\sqrt{T}} \sum_{t=1}^{T} y_t^2, \tag{2}$$

such that $D_y(0) = D_y(T) = 0$. When y_t satisfies the null hypothesis of constant unconditional variance, a plot of $D_y(k)$ against k will be a horizontal line oscillating around zero. However, under the alternative of a sudden change in the variance occurring at a certain point κ during the sample, the value of $D_y(k)$ will move away from zero in either the positive or negative direction for values of $k < \kappa$. Theoretically, the absolute value of $D_y(k)$ will achieve its maximum at $k = \kappa$, after which it will return towards zero. For this reason $|D_y(k)|$ provides a natural test for a volatility change, as well as an estimate of the change-point.

Suppose that the maximum of $|D_y(k)|$ is attained at $k = k^*$, that is

$$|D_y(k^*)| = \max_{1\leq k\leq T} |D_y(k)|. \tag{3}$$

We then identify a breakpoint at k^* if $|D_y(k^*)|$ is larger then some predetermined critical value, which can be obtained from the asymptotic distribution of $D_y(k)$. It can be shown that under fairly mild regularity conditions, see Boswijk (2004) among others, $D_y(k)$ weakly converges to a (scaled) Brownian bridge, such that

$$\frac{1}{\varsigma}|D_y(k^*)| \xrightarrow{d} \sup_{0\leq s\leq 1} |B(s)|, \tag{4}$$

where ς^2 is the long-run variance of the squared series y_t^2, that is $\varsigma^2 = \sum_{j=-\infty}^{\infty} \gamma_j$ with γ_j the j-th order autocovariance of y_t^2, and where $B(s)$ is a standard Brownian bridge, defined as $B(s) = W(s) - sW(1)$ with $W(\cdot)$ a standard Wiener process and $0 \leq s = k/T \leq 1$. It follows that an appropriate CUSUM test statistic is given by

$$U_y(k^*) = \frac{1}{\hat{\varsigma}} \max_{1\leq k\leq T} |D_y(k)|, \tag{5}$$

where $\hat{\varsigma}^2$ is a consistent estimator for ς^2.

Obviously, the (assumptions concerning the) distributional properties of the time series y_t determine its long-run variance ς^2 and, furthermore, imply how it should be estimated. It is in this respect that the different CUSUM statistics that have been proposed differ. First, assuming that $\{y_t\}_{t=1}^T$ is a sequence of independent and identically distributed (iid) normal random variables, as in Inclán and Tiao (1994), the autocovariances of y_t^2 are all equal to zero, that is $\gamma_j = 0$, $\forall j \neq 0$, such that the long-run variance $\varsigma^2 = \gamma_0$. Due to the normality assumption ς in fact reduces to $\sigma^2\sqrt{2}$, where σ^2 is the variance of y_t,[1] which can be consistently estimated by $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^T y_t^2 = \frac{1}{T}C_y(T)$. It is then straightforward to show that the CUSUM statistic $U_y(k^*)$ as given in (5) is equivalent to

$$\sqrt{\frac{T}{2}} \max_{1\leq k\leq T} \left| \frac{C_y(k)}{C_y(T)} - \frac{k}{T} \right|,$$

[1] If y_t is iid normal and assuming that $\mathsf{E}[y_t] = 0$, the kurtosis of y_t equals $3 = \mathsf{E}[y_t^4]/((\mathsf{E}[y_t^2])^2)$. Consequently, $3\sigma^4 = E[y_t^4]$. Given that $\gamma_0 = \mathsf{E}[(y_t^2 - \mathsf{E}[y_t^2])^2] = \mathsf{E}[y_t^4] - (\mathsf{E}[y_t^2])^2$, it follows that $\gamma_0 = 2\sigma^4$.

which is the form used in Inclán and Tiao (1994).

Second, assuming that y_t is iid but not necessarily normally distributed, the long-run variance ς^2 is still equal to $= \gamma_0$, but now γ_0 should be estimated directly from the time series y_t as $\hat{\gamma}_0 = \frac{1}{T}\sum_{t=1}^{T}(y_t^2 - \sum_{t=1}^{T} y_t^2)^2 = \frac{1}{T}\sum_{t=1}^{T} y_t^4 - (\frac{1}{T}\sum_{t=1}^{T} y_t^2)^2$.

Third, one may relax the iid assumption and allow for various forms of dependence and heterogeneity in y_t. For example, Lee and Park (2001) allow for temporal dependence by assuming that y_t follows an MA(∞) process, that is $y_t = \sum_{j=1}^{\infty}\theta_j\varepsilon_{t-j} + \varepsilon_t$. Here we are mainly interested in cases where y_t displays conditional heteroskedasticity. In that respect, Kokoszka and Leipus (2000) assume that y_t follows an ARCH(∞) process,

$$
\begin{aligned}
y_t &= z_t\sqrt{h_t}, \\
h_t &= \omega + \sum_{j=1}^{\infty}\alpha_j y_{t-j}^2,
\end{aligned}
\tag{6}
$$

with α_j being non-negative constants and $z_t \sim$ iid $N(0,1)$. Alternatively, Kim *et al.* (2000) assume a GARCH(1,1) process for y_t,[2]

$$
\begin{aligned}
y_t &= z_t\sqrt{h_t}, \\
h_t &= \omega + \alpha y_{t-1}^2 + \beta h_{t-1},
\end{aligned}
\tag{7}
$$

with α, β positive constants such that $\alpha + \beta < 1$ and again $z_t \sim$ iid $N(0,1)$.

In all these cases, the squared series y_t^2 has non-zero autocorrelations γ_j, $j \neq 0$, at all lags and, consequently, $\hat{\gamma}_0$ does not provide a consistent estimate of the long run variance ς^2. One possible solution to this problem is to derive an explicit expression for γ_j, and thereby for ς^2, based on the specific parametric structure of the process that is assumed for y_t, as is done in Kim *et al.* (2000) for the GARCH(1,1) case. However, one can imagine that this procedure is rather sensitive to model misspecification.[3] An alternative and more robust approach is to use a nonparametric or data-based estimator of ς^2, as advocated in both Kokoszka and Leipus (2000) and Lee and Park (2001). There are several possibilities in this case. Andreou and Ghysels (2002), for example, use the autoregression heteroscedasticity and autocorrelation consistent (ARHAC) estimator of den Haan and Levin (1997). In our study, we use the popular Bartlett kernel estimator $\hat{\varsigma}^2 = \hat{\gamma}_0 + 2\sum_{j=1}^{l} w_{j,l}\hat{\gamma}_j$ where $w_{j,l} = j/(l+1)$,

[2] Alternative approaches to testing for parameter change in GARCH models have recently been developed by Chu (1995), Kokoszka and Teyssière (2002), and Lundbergh and Teräsvirta (2002), among others. A comparison of these tests with the CUSUM statistics considered here is of interest, but is left for future research.

[3] Lee *et al.* (2003) in fact observe that the parametric approach of Kim *et al.* (2000) does not work satisfactorily under all circumstances even when the DGP is a GARCH(1,1) process. In particular, the test suffers from size distortions and low power for certain parameterizations.

with automatic selection of the truncation lag or bandwidth $l > 0$ using an AR(1) model, as suggested in Andrews (1991).

In sum, each of the three types of assumptions discussed above lead to test statistics based on the same CUSUM process $D_y(k)$. Hence, they share the same limiting distribution under the null hypothesis and under correctness of the underlying assumptions, namely that of a (scaled) Brownian bridge. The only difference between the tests is the use of a different scaling factor or estimate of ς. Specifically, (i) under iid normality: $\hat{\varsigma}^2 = 2\hat{\sigma}^4$, (ii) under iid, but not necessarily normality: $\hat{\varsigma}^2 = \hat{\gamma}_0$, and (iii) under general dependence and heterogeneity: $\hat{\varsigma}^2 = \hat{\gamma}_0 + 2\sum_{j=1}^{l} w_{j,l}\hat{\gamma}_j$. In the Monte Carlo simulations reported below we consider all three statistics, which are denoted as $U_{y,\sigma}(k^*)$, $U_{y,\gamma_0}(k^*)$, and $U_{y,\varsigma}(k^*)$, respectively.

It is shown in Section 3.1 that all tests, including $U_{y,\varsigma}(k^*)$, suffer from severe size distortions in finite samples when y_t exhibits conditional heteroscedasticity, in particular when y_t follows a GARCH(1,1) process as given in (7). Hence, it seems advisable to filter the series first, in order to remove the conditional heteroskedasticity. Interestingly, nonparametric or "model-free" approaches of standardizing the series y_t either with volatility estimates based on high-frequency data (such as quadratic variation) or with Riskmetrics' volatility estimates (obtained as $\hat{h}_t = (1-\lambda)y_{t-1}^2 + \lambda\hat{h}_{t-1}$ with $\lambda = 0.94$) do not work well. In particular, this renders severely undersized test statistics; this corresponds with the findings of Andreou and Ghysels (2003) for CUSUM-type tests in changes in co-movement of conditionally heteroskedastic time series.[4] Here we explore the suggestion of Lee *et al.* (2003) to apply the statistic in (5) based on the CUSUM process $D_{\hat{z}}(k) = \frac{1}{\sqrt{T}}\sum_{t=1}^{k}\hat{z}_t^2 - \frac{k}{T\sqrt{T}}\sum_{t=1}^{T}\hat{z}_t^2$ of standardized residuals $\hat{z}_t \equiv y_t/\sqrt{\hat{h}_t}$, where $\hat{h}_t$ is the estimated conditional volatility of y_t obtained from a GARCH(1,1) model estimated with (quasi-)maximum likelihood ((Q)ML) assuming a normal distribution for z_t. The properties of (squared) standardized (G)ARCH residuals have been studied quite intensively in recent years, see Horváth *et al.* (2001), Berkes and Horváth (2003), and Berkes *et al.* (2003), among others. Lee *et al.* (2003) prove that, given the correct conditional volatility specification, the (scaled) CUSUM process $D_{\hat{z}}(k)$ converges to a Brownian bridge, such that the limiting distribution result as given in (4) continues to hold. Indeed, in the simulations reported below we find that the associated $U_{\hat{z},\cdot}(k^*)$ statistics have satisfactory size and power properties. One may doubt the practical usefulness of this parametric approach, as the properties of $U_{\hat{z},\cdot}(k^*)$ might be very sensitive to misspecification of the conditional volatility process. We explore this issue in depth in Section 3, and find that the CUSUM statistics based on standardized GARCH(1,1)-residuals are in

[4]Detailed simulation results demonstrating this result are available upon request.

fact remarkably robust to various forms of misspecification.

2.2 Testing for multiple structural changes

In the above we focused on testing for a single change in the unconditional variance of y_t. However, there is no reason why the volatility of a time series might not experience multiple changes. Testing for multiple changes in volatility has been addressed in a number of recent articles, including Chen and Gupta (1997) and Lavielle and Moulines (2000). Both studies develop an information criterion based penalized least-squares estimation approach to test for (and date) multiple breaks simultaneously. Similar to the testing framework developed by Bai (1997, 1999) and Bai and Perron (1998, 2003), CUSUM statistics can be applied in a sequential manner to test for and identify multiple volatility changes. The basic idea is that first the entire sample is tested for the presence of a single break in volatility using the CUSUM statistics discussed in Section 2.1. If a significant change is detected, the sample is split into two segments with the split point being equal to the identified change-point. Next, each subsample is examined separately for a volatility break, again using a CUSUM test. This procedure continues until no more changes are detected in any of the subsamples or until the number of identified changes reaches a pre-specified maximum. Sometimes a final step is added in which all identified breaks are re-evaluated and/or breakpoints re-estimated. In this context, Inclán and Tiao (1994) develop the Iterated Cumulative Sums of Squares (ICSS) algorithm which repeatedly applies their $U_{y,\sigma}(k^*)$ statistic.

We adopt a sequential approach here as well, based on the basic set-up discussed above. Our procedure works as follows. Suppose that at some point in the algorithm N volatility changes have been detected, for $N < M$ with M being the maximum allowed number of breaks. Consequently, the sample for y_t can be split into $N+1$ segments, according to the associated change-point estimates $1 = k_0^* < k_1^* < \ldots < k_N^* < k_{N+1}^* = T$. To test whether any of the segments contains an additional volatility change, we compute one of the CUSUM statistics $U_{\hat{z},\sigma}(k^*)$, $U_{\hat{z},\gamma_0}(k^*)$ or $U_{\hat{z},\varsigma}(k^*)$ for each subsample separately,[5] and select the segment for which the test statistic is largest. Suppose this occurs in the i-th segment for $1 \leq i \leq N+1$. If the value of the corresponding CUSUM statistic exceeds an appropriate critical value (see Section 2.3), we identify the $(N+1)$-th break in segment i. We repeat this procedure until either N equals M or the maximum of the test statistics across all segments is no longer significant. We control the overall significance level of the sequential procedure by using a significance level of $a/(N+1)$ when testing

[5] Note that this includes estimating separate GARCH(1,1) models for all segments.

for the $(N+1)$-th change in volatility. Finally, we re-estimate all change-points, where the location of the i-th volatility change is re-estimated based on the segment determined by the adjacent breakpoints k_{i-1}^* and k_{i+1}^*.[6] This corresponds with the "repartitioning" step in the Bai and Perron (1998) procedure.

Apart from a maximum allowed number of breaks, a second restriction that we impose in the algorithm is that adjacent change-points have to be at least δ observations apart. The latter restriction is to prevent breaks from being identified unrealistically close together. Although the precise value of δ clearly is a subjective decision, we feel that for daily data $\delta = 63$ or 126 business days (three and six months, respectively) seems appropriate. We impose the minimum distance restriction by calculating the maximum absolute value of the CUSUM test statistic in the i-th segment only using the permitted values of k, i.e. $k_{i-1}^* + \delta \leq k \leq k_i^* - \delta$, determined so far in the algorithm. Note that in the final step of the algorithm in which we re-estimate each change-point, we can actually not control the minimum distance between adjacent volatility changes. This would require treating the two adjacent change-points as fixed, whereas these can still be re-estimated at a different location. Hence, it may occur in practice that final breakpoint estimates are less than δ observations apart.

Our procedure as outlined above differs in a number of respects from the ICSS algorithm of Inclán and Tiao (1994). First, after detecting a first volatility change at $k = k_1^*$, the ICSS algorithm examines the first subsample $y_t, t = 1, \ldots, k_1^*$ exhaustively to identify the leftmost significant breakpoint, k_l^*, after which the same is done for the second subsample $y_t, t = k_1^* + 1, \ldots, T$ to identify the rightmost significant break point k_r^*. If the leftmost break point differs from the rightmost breakpoint, that is $k_l^* < k_r^*$, then the procedure is repeated for the subsamples $t = k_l^* + 1, \ldots, k_r^*$ until $k_l^* = k_r^*$. In our procedure, we consider all N segments when testing for a $(N+1)$-th break. Second, in the ICSS algorithm the same significance level is applied to each subsample, irrespective of how many breaks have already been found. Third, in the ICSS algorithm breaks can be arbitrarily close to each other, as no minimum distance restriction is imposed. Fourth, in the final step of the ICSS algorithm change-points are not only re-estimated, but the significance of all volatility breaks is determined again, using only the observations from the relevant segment. Earlier detected breaks are removed if they are no longer significant. Finally, and perhaps most important, the ICSS algorithm is based on the $U_{y,\sigma}(k^*)$ statistics, which does not account for possible non-normality and conditional heteroskedasticity.

[6] For the first volatility change we use the sample from the first observation k_0^* up to and including the second change-point k_2^*. For the last volatility change we use the sample from observation k_{N-1}^* until the last observation k_{N+1}^*.

2.3 Finite sample critical values

One issue we have not touched upon so far is the use of critical values. Especially when testing for multiple breaks, the length of the subsamples can quickly become quite small, which renders the use of asymptotic critical values questionable at the least. Therefore, we choose to use finite sample critical values. These are estimated through simulation using the response surface approach described in MacKinnon (2000).[7]

Suppose that we need the quantile of the distribution of the CUSUM test under the null hypothesis corresponding to a certain significance level a and for a specific finite sample length T, and denote this quantile by $q^a(T)$. This can be obtained by simulating a large number, R, of series of length T from the data-generating process under the null hypothesis and calculating the test statistic for each series. The simulated test statistics can be used to construct the appropriate finite sample distribution and the relevant quantile. Repeating this experiment a total of E times for this specific sample length results in E observations for $q^a(T)$. By repeating this process for different values of T we can then estimate the following type of response surface regressions

$$q_e^a(T) = \theta_\infty^a + \theta_1^a T^{-0.5} + \theta_2^a T^{-1} + \varepsilon_e, \tag{8}$$

where $q_e^a(T)$ denotes the quantile estimate obtained in the e-th experiment for sample size T. Subsequently, the estimated response surface regression can be used to determine the appropriate finite sample critical value (quantile) for *any* sample size T. Also note that $\hat{\theta}_\infty^a$ is an estimate of the asymptotic critical value $q^a(\infty)$. The parameter estimates θ_1^a and θ_2^a in our case typically are negative, such that finite sample quantiles are smaller than their asymptotic counterparts. Hence, if asymptotic critical values were used, the tests would appear to be undersized.

As discussed in the previous section we impose the restriction that two adjacent change-points should be at least δ observations apart, reducing the effective sample size. To account for this we modify the response surface specification by including powers of π, with π being the fraction of observations not considered at either side of the sample when calculating the test statistic, that is $\pi = \delta/T$. Specifically, we estimate response surface regressions of the form

$$q_e^a(T, \pi) = \theta_\infty^a + \theta_1^a T^{-0.5} + \theta_2^a T^{-1} + \phi_1^a \pi + \phi_2^a \pi^2 + \phi_3^a \pi^3 + \phi_4^a \pi^4 + \phi_5^a \pi^5 + \varepsilon_e. \tag{9}$$

To implement the response surface regression, we perform $E = 40$ experiments with $R = 50000$ iid $N(0, 1)$ replications each for sample sizes $T_i \in \{50, 60, \ldots, 100, 125, \ldots,$

[7] An alternative would be to consider bootstrap procedures for computing critical values or p-values, as in Kokoszka and Teyssière (2002). Given the extent of the Monte Carlo simulations conducted in the next section, however, the response surface technique is more convenient.

$250, 300, 350, \ldots, 500, 600, \ldots, 1000, 1500, \ldots, 3000, 4000, 5000\}$, trimming percentages $\pi \in \{0.025, 0.05, \ldots, 0.425, 0.45\}$, and for quantiles corresponding to significance levels $a = 10\%$, 5% and 1%. The fit of the response surface (9) generally is very good, with R^2 values never being below 97%. To illustrate Figure 1 shows the finite sample response surface for the $U_{\hat{z},\sigma}(k^*)$ test for a significance level of 5%. It is seen from the graph that finite sample critical values differ substantially from their asymptotic counterparts when either T is small or when π becomes close to 0.5.

3 Simulation design and results

In this section we report and discuss results from an extensive set of Monte Carlo simulations experiments, designed to examine the small-sample properties of the CUSUM tests and to assess their robustness to various types of misspecification. The (limited amount of) simulation results available in the literature typically only consider the properties of CUSUM tests for data-generating processes (DGPs) that match the assumptions under which a particular test was developed. For example, Inclán and Tiao (1994) evaluate the size and power properties of their test for iid normal series, while Kim *et al.* (2000) and Lee *et al.* (2003) perform simulations using a GARCH(1,1) process with normal shocks z_t as DGP. To some extent, Andreou and Ghysels (2002, 2004) form an exception, as they do consider alternative DGPs, namely GARCH(1,1) processes with possibly non-normal errors and with different degrees of volatility persistence, for the $U_{y,\sigma}(k^*)$ and $U_{y,\varsigma}(k^*)$ tests. They find that both tests suffer from positive size distortions and their ability to correctly identify breaks varies. Although this provides a reasonable indication of the small sample properties of these tests, we feel that there is scope for broadening these results. In particular, our simulation study has two purposes. First, we assess the small sample properties of the CUSUM tests when applied to standardized GARCH(1,1) residuals $\hat{z}_t$.[8] Second, we examine the robustness of the CUSUM tests to various forms of misspecification, including alternative error distributions as well as misspecification in the conditional variance dynamics.

3.1 Size properties

We start our analysis by gauging possible size distortions for the three CUSUM tests. For each DGP discussed below, we generate 10000 replications of length $T = 500$,

[8] Andreou and Ghysels (2002) do show results for the $U_{z,\varsigma}(k^*)$ test applied to the standardized series $z_t = y_t/\sqrt{h_t}$. However, when constructing the standardized series, the true simulated conditional variance series h_t is used. As a result, they do not take into account the effects of parameter estimation uncertainty and misspecification of the conditional volatility process.

1000, 2000 and 4000. We examine rejection frequencies at nominal significance levels $a = 10\%$, 5% and 1%, using finite sample critical values obtained from the response surface regression in (9), where we set $\pi = 0$ throughout. We consider a number of different DGPs where we focus mainly on those that relate to different types of potential misspecification.

First, we consider four DGPs under which the variance of y_t does not have conditional dependence. To be precise, we generated iid series from (i) a standard normal distribution, (ii) a Student-$t(\nu)$ distribution with the degrees of freedom parameter ν ranging from 4 to 8, (iii) a skewed-normal(λ) distribution, see Azzalini (1985), with the skewness parameter λ ranging from -5 (severe negative skewness) to -1 (moderate negative skewness) and finally (iv) normal-with-jumps. Under the latter DGP we add a jump component to the series in such a way that y_t jumps at random points in the sample, but with a fixed and predetermined jump size and jump intensity. Detailed results of these experiments are not shown here to save space but are available upon request. The results can be summarized as follows. When applied to the raw series as well as the standardized residuals from a GARCH(1,1) model, the $U_{\cdot,\sigma}(k^*)$ test is severely oversized for all distributions except the standard normal. This is not surprising given that this test critically depends on the normality assumption for y_t. The $U_{\cdot,\gamma_0}(k^*)$ and $U_{\cdot,\varsigma}(k^*)$ tests are almost always correctly sized, albeit rejection frequencies tend to be somewhat below the nominal significance levels for the Student-t distribution when the number of degrees of freedom is small ($\nu = 4, 5$) and the normal-with-jumps DGP when the jump size is substantial (jumps of 5 or 10 times the standard deviation of the regular component of y_t).

Second, we consider GARCH-type DGPs, such that the variance of y_t exhibits conditional dependence. We first employ a standard GARCH(1,1) model using various combinations of α and β and with different distributions for the errors z_t. We consider the same four distributions as above, albeit for the Student-t distribution we only use $\nu = 5$ and for the skewed normal only $\lambda = -5$. Table 1 shows the results of applying the tests to the raw series y_t and to standardized GARCH(1,1) residuals $\hat{z}_t$ for a GARCH(1,1) DGP with normal shocks z_t. The left panel shows that the $U_{y,\sigma}(k^*)$ test is severely oversized.[9] Again, this occurs because the iid normality assumption underlying the test is violated. The $U_{y,\gamma_0}(k^*)$ test is oversized as well, due to the fact that the nonzero (positive) autocorrelations of y_t^2 are not accounted for. What is surprising though is that the $U_{y,\varsigma}(k^*)$ test also suffers from substantial

[9] At first sight, it may seem odd that size distortions occur despite the use of finite sample critical values from the response surface (9). However, the response surface was created assuming a homoskedastic DGP, such that empirical rejection frequencies for heteroskedastic series can still differ from the nominal significance levels. Using asymptotic critical values renders even worse size distortions, given that finite sample critical values are smaller than asymptotic ones.

positive size distortions, which become larger when conditional volatility is more persistent (see also Table 1 of Andreou and Ghysels (2002)). Hence, although this CUSUM test theoretically is valid in the presence of heteroscedasticity, as shown in Kokoszka and Leipus (2000), it may require unrealistically large sample sizes for this asymptotic result to apply.[10]

Turning to the right panel, we observe that, when applied to GARCH(1,1) residuals, no substantial size distortions occur for all tests across all parameterizations. Given these results and to facilitate comparison with Andreou and Ghysels (2002), we only report results for the $U_{\hat{z},\varsigma}(k^*)$ test in the remainder of this section. Detailed results for other statistics are available upon request. Table 2 reports results for the other three distributions for z_t, showing that the $U_{\hat{z},\varsigma}(k^*)$ test statistic is properly sized for each of these. This might be expected of course, given that the normal QML estimator of the parameters in the GARCH(1,1) model is consistent. The unreported results for the other statistics show that the same holds for the $U_{\hat{z},\gamma_0}(k^*)$, while the $U_{\hat{z},\sigma}(k^*)$ is again oversized due to the non-normality of the shocks z_t (and hence also the standardized residuals $\hat{z}_t$).

We now turn to different types of misspecification of the conditional variance process. We consider first of all the asymmetric GARCH(1,1) process put forward by Glosten *et al.* (1993), where negative shocks have a different impact on conditional volatility than positive shocks of the same magnitude. The GJR-GARCH(1,1) model is specified as

$$y_t = z_t\sqrt{h_t}, \tag{10}$$

$$h_t = \omega + \alpha y_{t-1}^2 + \gamma y_{t-1}^2 \mathtt{I}[y_{t-1} < 0] + \beta h_{t-1}, \tag{11}$$

where we set $\alpha = 0$ and $\omega = 1 - \gamma/2 - \beta$, such that the unconditional variance of y_t equals 1.

Second, we examine a long-memory fractionally integrated GARCH process (FI-GARCH(1,1)), see Baillie *et al.* (1996),

$$y_t = z_t\sqrt{h_t}, \tag{12}$$

$$h_t = \omega + (1 - \beta - (1 - L)^d)y_{t-1}^2 + \beta h_{t-1}, \tag{13}$$

where d is the long memory parameter.

Finally, we consider a stochastic volatility (SV-AR(1)) DGP, see Taylor (1986),

$$y_t = z_t \exp(h_t/2) \tag{14}$$

$$h_t = \gamma_0 + \gamma_1 h_{t-1} + \eta_t \tag{15}$$

[10] The size distortion of the $U_{y,\varsigma}(k^*)$ does diminish as the sample size increases, such that the empirical rejection frequencies converge to the nominal significance levels, albeit very slowly. Note that for the $U_{y,\sigma}(k^*)$ and $U_{y,\gamma_0}(k^*)$, the empirical size actually becomes worse as T becomes larger.

where $\gamma_0 = -(1-\gamma_1)/2$, $\eta_t \sim \text{iid}\, N(0, \sigma_\eta^2)$ with $\sigma_\eta^2 = 1 - \gamma_1^2$, and z_t and η_t are independent. In all three models above, $z_t \sim \text{iid}\, N(0,1)$.

Rejection frequencies of the $U_{\hat{z},\varsigma}(k^*)$ statistic for the GJR-GARCH(1,1) DGP are quite close to the nominal significance levels used, especially for larger sample sizes $T \geq 2000$; see the upper panel of Table 3. By contrast, for the FI-GARCH(1,1) process the test suffers from, sometimes quite severe, positive size distortions, which worsen as the sample size T increases. Apparently the test gets confused when volatility undergoes longer lasting upswings and downswings which are mistakenly considered as structural breaks. Note that this is the reverse phenomenon of mistaking structural changes for long-memory, which has been discussed at considerably length in the literature, see Lamoureux and Lastrapes (1990), Liu (2000), Diebold and Inoue (2001), Franses *et al.* (2002) and Mikosch and Starica (2004), among others. Finally, the test is somewhat conservative for the SV-AR(1,1) DGP, in the sense that empirical rejection frequencies are a bit below the nominal significance levels. Nevertheless, overall the $U_{\hat{z},\varsigma}(k^*)$ test appears to be quite robust to various forms of misspecification.

3.2 Power properties

We now turn to the power properties of the $U_{\hat{z},\varsigma}(k^*)$ CUSUM test.[11] We first consider the case of a single break in volatility when the DGP is a GARCH(1,1) process with $z_t \sim N(0, \sigma_z^2)$. As the unconditional variance is given by $\sigma_y^2 = \omega\sigma_z^2/(1-\alpha-\beta)$ four potential causes for a variance change can be identified: a break in either ω, σ_z^2, α or β.[12] As the second cause is observationally equivalent to the first, we only consider breaks in the parameters ω, α and β. We allow for three different timings of the parameter change at τT for $\tau = 0.25$, 0.50 and 0.75, again using sample sizes $T = 500$, 1000, 2000 and 4000.

Table 4 shows rejection frequencies across 1000 replications from the GARCH(1,1) DGP with a break occurring in ω, where we only consider breaks that occur in the middle of the sample. A number of conclusions can be drawn from this table. These generally also hold true for subsequent tables so we discuss them in somewhat more detail here. First, power increases both with the magnitude of the change in ω (and thus in unconditional volatility) and with the sample size T, except when $\beta = 0.50$ and volatility after the change in ω is very small ($\sigma_a^2 = 0.50$). Second, there appears to exist asymmetry in the test's capability of detecting volatility changes, with volatility increases being picked up better than decreases or vice versa. The

[11] Again, detailed results for the other test statistics are available upon request.

[12] It can of course also happen that several parameters change at the same time. We do not consider that possibility here.

direction of the asymmetry depends on the volatility persistence as measured by β. Power is generally higher for volatility decreases for $\beta = 0.80$, whereas for $\beta = 0.50$ it is easier to detect volatility increases. Third, for the smaller sample sizes $T = 500$ and 1000 power is higher for low volatility persistence ($\beta = 0.50$). For the larger sample sizes $T = 2000$ and 4000 this continues to hold for small volatility changes, while large breaks in volatility are easier to detect under high volatility persistence.

Results for a single break in β are shown in Table 5. In addition to the increase in power with the magnitude of the change in unconditional volatility and with the sample size T, we observe that decreases in volatility now are easier detected under low volatility persistence ($\beta_b = 0.50$) as well. Furthermore, power is largest for breaks that occur in the middle of the series. For decreases in volatility, early changes ($\tau = 0.25$) are easier to detect than late ones ($\tau = 0.75$) while the reverse holds for volatility increases.[13]

Table 6 shows rejection frequencies when a break occurs in α. It is seen that a break in α is generally more difficult to detect than a break in either ω or β leading to the same change in unconditional volatility. Only for substantial changes and only when volatility is lower after the change does the power of the test seem reasonable. Again, when volatility increases, power is slightly better for breaks occurring late in the sample compared to early breaks while the reverse holds for decreases in volatility. Finally, the level of volatility persistence, reflected by β, is of influence with the rejection frequencies being higher for low persistence as compared to high persistence.

We also considered breaks in ω, α and β with different distributions for the shocks z_t, including Student-$t(\nu)$ and skewed-normal(λ). Typically, power goes down somewhat compared to DGPs with z_t being normally distributed. The same occurs when the DGP is a GJR-GARCH(1,1) process with a break in either ω or β. Detailed results for these experiments are available upon request.

Given that we wish to apply the CUSUM tests in the sequential procedure for multiple volatility changes, as described in Section 2.2, we consider their power in detecting multiple breaks in the GARCH parameters. To conserve space we only discuss a GARCH(1,1) DGP with two breaks in either ω or β. Results are reported in Tables 7 and 8, respectively. We first of all observe that power is reasonable when volatility first goes down and then jumps up again (or the reverse), but in such a way that it does not return to its initial level. For low volatility persistence the test detects the two breaks quite well. However, for high persistence, power is considerably lower. When volatility does return to its original level after the second

[13] The latter is noted and discussed in Inclán and Tiao (1994) by considering the expected value of $D_T(k)$.

change, the test typically identifies no breaks at all. In the latter case, the test only starts to pick up the changes in volatility when the series is fairly long ($T = 4000$). For a stepwise decrease in volatility from 1 to 0.7 to 0.4 we see that the test has difficulty in picking up the correct number of breaks, with two breaks being detected in only 20% of the cases for moderate samples sizes. The level of persistence again matters. Focusing for ease of discussion on $T = 2000$, we see that for $\beta = 0.50$ the test typically identifies no breaks at all, whereas for $\beta = 0.80$ a single change is identified in 70% of the replications. For a smaller step size (1 to 0.9 to 0.8) the pictures changes. Now for low volatility persistence either 0 or 1 break is identified equally frequent, whereas for high persistence zero breaks are identified most often (74%). The same result hold for small increases (1 to 1.1 to 1.2). For larger stepwise increases, the null hypothesis of no breaks is increasingly often rejected, but the ability to find two breaks is still quite low. Similar to the single break DGPs, power is higher for decreases in volatility compared to increases in volatility.

4 Volatility changes in emerging stock markets

Research in emerging market finance has been rapidly expanding over the past two decades, see Bekaert and Harvey (2002, 2003) for comprehensive surveys of the past, present and future of this area. Among others, this growing interest stems from the fact that emerging market assets have become increasingly important in international investment portfolios. This has been made possible thanks to the capital market liberalizations many of these countries experienced in the late 1980s and early 1990s. Typically, the liberalization measures that were implemented included substantial reduction or even complete removal of barriers-to-entry for foreign investors. The financial and macroeconomic effects of these liberalizations have been studied intensively, see Bekaert and Harvey (2003) for an overview. The consequences of liberalization and the resulting (or at least hoped-for) integration with developed markets for stock market volatility are not clear *a priori*. A common perception is that the growing influence of highly mobile foreign capital (so-called "hot money") might lead to higher volatility in liberalized markets. However, empirical studies have found little support for this hypothesis, but instead document either no significant changes or declines in stock market volatility following liberalizations, see Richards (1996), Bekaert and Harvey (1997, 2000a), De Santis and Imrohoroglu (1997), Aggarwal *et al.* (1999) and Kim and Singal (2000), and the references contained therein. The finding of lower post-liberalization volatility typically is attributed to increased market efficiency and diversification effects.

Analyzing the effect of liberalizations on stock market volatility is not without complications, however. First, in most countries liberalization has been a gradual process, with different measures taken at different points in time. Second, emerging markets volatility may change for a host of reasons other than financial liberalization, including (both local and global) social, political or economic events. Consequently, multiple sudden and substantial changes in volatility may be observed in these markets. In this section we aim to identify volatility changes in emerging stock market index returns by means of the CUSUM tests. Our empirical study resembles that of Aggarwal *et al.* (1999), although they analyze a considerably smaller set of countries over a shorter sample period and only use the original Inclán-Tiao CUSUM-statistic $U_{y,\sigma}(k^*)$. The latter difference may be most crucial as the $U_{y,\sigma}(k^*)$ statistic does not account for possible non-normality and conditional heteroskedasticity, which are relevant characteristics of emerging stock market returns, see Bekaert and Harvey (1997) and Bekaert *et al.* (1998), among others.

We examine daily returns on MSCI indexes for a total of 27 emerging stock markets. We select countries from each of the three emerging market clusters identified by MSCI: China, India, Indonesia, Korea, Malaysia, Pakistan, Philippines, Sri Lanka, Taiwan and Thailand (Asia), Argentina, Brazil, Chile, Colombia, Mexico, Peru and Venezuela (Latin America) and Czech Republic, Egypt, Hungary, Israel, Jordan, Morocco, Poland, Russia, South Africa and Turkey (Europe, Middle East and Africa).[14] The sample period runs from January 1, 1988 to December 31, 2003, resulting in a total of 4173 daily return observations, although not all series start on January 1, 1988. The second column of Table 9 shows the starting date of the returns series for each country. The countries with the shortest samples (Czech Republic, Egypt, Hungary, Morocco and Russia) still have over 2000 observations. Following Aggarwal *et al.* (1999), we consider returns measured in US dollars as well as in local currency. Unreported summary statistics confirm the importance of non-normality (in the form of significant skewness, excess kurtosis and infrequent large jumps, both positive and negative) and conditional heteroskedasticity for these stock return series.

We start with the original Inclán-Tiao ICSS algorithm for detecting and dating multiple breaks in the unconditional volatility of demeaned returns.[15] Columns four and 11 of Table 9 show the number of breaks thus identified by the $U_{y,\sigma}(k^*)$ test, indicating an unrealistically large number of volatility changes. Furthermore,

[14]In the remainder of the analysis we do not need to differentiate between countries from different regions. Consequently, countries are given in alphabetical order in the tables and graphs below.

[15]We used the Schwarz Information Criterion (SIC) to determine the optimal lag order p in an autoregressive (AR(p)) model for each return series. It turned out that in general a lag order of $p = 0$ was selected and consequently we did not use any autoregressive lags when demeaning the return series.

sometimes the identified change-points are only a few weeks or even days apart. It is therefore hardly justifiable to classify these as genuine shifts in the level of volatility.

We proceed with our sequential testing algorithm as described in Section 2.2, allowing for a maximum number of 10 breaks, which each have to be at least 126 (trading) days apart[16]. Appropriate finite sample critical values are obtained from (9), using an initial nominal significance level $a = 0.05$. We implement the algorithm with each of the three CUSUM statistics $U_{y,\sigma}(k^*)$, $U_{y,\gamma_0}(k^*)$ and $U_{y,\varsigma}(k^*)$ to the demeaned returns. The number of detected volatility changes is drastically reduced, as shown by the results in columns 5-7 for US dollar returns and columns 12-14 for local currency returns. The $U_{y,\sigma}(k^*)$ test still often identifies the maximum number of 10 breaks, which is due to the fact that it cannot account for the non-normality of the return series. The $U_{y,\gamma_0}(k^*)$ test, which scales down the centered sum of squares with the variance of the squared returns series, only hits the upper bound for four countries (Argentina, Chile, Thailand and Turkey) for the US dollar returns and for only two counties (Argentina and Thailand) for the local currency returns. The $U_{y,\varsigma}(k^*)$ statistic on the other hand is never constrained by the imposed maximum number of breaks. The maximum number of identified breaks across all returns series is 8, for Turkey.

Although the number of variance changes based on the $U_{y,\varsigma}(k^*)$ test appear to be quite reasonable, it seems only natural to apply the CUSUM tests to standardized returns in light of the size distortions documented in Section 3.1. Doing so using a GARCH(1,1) model yields the results shown in columns 8-10 and the final three columns of Table 9. Compared to the results for demeaned returns, the number of breaks further declines and to such an extent that for some countries no volatility changes are identified at all when using either the $U_{\hat{z},\gamma_0}(k^*)$ or $U_{\hat{z},\varsigma}(k^*)$ statistics. Furthermore, these two tests now yield the exact same number of breaks for all countries. The $U_{\hat{z},\sigma}(k^*)$ statistic on the other hand still always identifies a positive number of breaks (except for US dollar returns in Brazil and local currency returns in Pakistan), although considerably less than before and also less than the number of changes identified in Aggarwal *et al.* (1999).

The magnitude and timing of the identified volatility changes is examined graphically in Figure 5, which presents plots of the daily returns in local currency. The horizontal lines in these graphs indicate ± 3 times the unconditional standard deviation between consecutive change-points, as identified by the $U_{\hat{z},\sigma}(k^*)$ and $U_{\hat{z},\varsigma}(k^*)$ statistics, in the upper and lower panels, respectively. Thick vertical lines correspond with the "official liberalization dates" as determined by Bekaert and Harvey

[16]We also considered a minimum distance of 63 business days which resulted in the same number of change points for all but a few countries.

(2000b) and Bekaert *et al.* (2003). It is seen that most identified breaks indeed correspond with marked shifts in volatility, many of which can be related to economic and political events such as, for example, the Asian and Russian financial crises in 1997 and 1998, respectively. Changes close to the official liberalization dates are found only for Chile and Indonesia, where in both cases volatility declined. What also becomes clear from Figure 5 is that whereas the $U_{\hat{z},\sigma}(k^*)$ test is inclined to find more, if not too many breaks, the $U_{\hat{z},\gamma_0}(k^*)$ and $U_{\hat{z},\varsigma}(k^*)$ tests have a tendency to be conservative. For example, it is surprising to see that no volatility changes are identified for Indonesia, Korea and Malaysia around the middle of 1997 when the Asian crisis occurred. The time series plots suggest that the stock markets in these countries experienced a substantial and prolonged volatility increase around that time. This more or less confirms the reduced power for the $U_{\hat{z},\varsigma}(k^*)$ statistic when testing for multiple breaks, as reported in Section 3.2.

5 Concluding remarks

In this paper we have examined CUSUM-based tests for changes in the unconditional volatility of conditionally heteroskedastic time series. A prominent conclusion from our analysis is that application of these tests to the raw time series observations leads to severe size distortions, rendering the tests highly unreliable. Remarkably, this was also found to be the case for the CUSUM test of Kokoszka and Leipus (2000), which at least theoretically allows for the presence of heteroskedasticity. Our simulation results show that it may require unrealistically large sample sizes for this asymptotic result to apply. Consequently, it appears necessary to filter the series in order to remove the heteroskedasticity prior to applying the CUSUM test. As a practical way to accomplish this, we adopt the suggestion of Lee *et al.* (2003) to use a GARCH(1,1)-volatility filter. Put differently, we recommend to apply the CUSUM test to standardized residuals from an estimated GARCH(1,1) model. Extensive Monte Carlo simulations showed that this results in correctly sized tests with good power properties when testing for a single break. Furthermore, the tests were found to be reasonably robust against various forms of model misspecification. The CUSUM tests appear to have difficulty to detect multiple changes in volatility and, hence, developing a more powerful procedure for testing for multiple breaks is an interesting topic for future research. The general properties of the CUSUM tests were confirmed in an application to emerging stock market returns, where the GARCH-filtered tests led to a considerably smaller, and much more realistic number of volatility changes than the original CUSUM statistics.

References

Aggarwal, R., C. Inclan, and R. Leal (1999), Volatility In Emerging Stock Markets, *Journal of Financial and Quantitative Analysis*, **34**, 33–55.

Andreou, E. and E. Ghysels (2002), Detecting Multiple Breaks in Financial Market Volatility Dynamics, *Journal of Applied Econometrics*, **17**, 579–600.

Andreou, E. and E. Ghysels (2003), Tests For Breaks in the Conditional Co-Movements of Asset Returns, discussion Paper 2003-03, University of Cyprus.

Andreou, E. and E. Ghysels (2004), The Impact of Sampling Frequency and Volatility Estimators on Change-Point Tests, *Journal of Financial Econometrics*, **2**, 290–318.

Andrews, D. (1991), Heteroscedasticity and Autocorrelation Consistent Covariance Matrix Estimation, *Econometrica*, **59**, 817–858.

Azzalini, A. (1985), A Class of Distributions Which Includes the Normal Ones, *Scandinavian Journal of Statistics*, **12**, 171–178.

Bai, J. (1997), Estimating Multiple Breaks One at a Time, *Econometric Theory*, **13**, 315–352.

Bai, J. (1999), Likelihood Ratio Tests for Multiple Structural Changes, *Journal of Econometrics*, **91**, 299–323.

Bai, J. and P. Perron (1998), Estimating and Testing Linear Models with Multiple Structural Changes, *Econometrica*, **66**, 47–78.

Bai, J. and P. Perron (2003), Computation and Analysis of Multiple Structural Change Models, *Journal of Applied Econometrics*, **18**, 1–22.

Baillie, R. T., T. Bollerslev, and H. Mikkelsen (1996), Fractionally Integrated Generalized Autoregressive Conditional Heteroscedasticity, *Journal of Econometrics*, **14**, 3–30.

Bekaert, G., C. B. Erb, C. R. Harvey, and T. E. Viskanta (1998), Distributional Characteristics of Emerging Market Returns and Asset Allocation, *Journal of Portfolio Management*, **24**, 102–116.

Bekaert, G. and C. R. Harvey (1997), Emerging Equity Market Volatility, *Journal of Financial Economics*, **43**, 29–77.

Bekaert, G. and C. R. Harvey (2000a), Capital Flows and the Behavior of Emerging Equity Market Returns, in S. Edwards (ed.), *Capital Flows and the Emerging Economies: Theory, Evidence and Controversies*, Chicago: University of Chicago Press, pp. 159–194.

Bekaert, G. and C. R. Harvey (2000b), Foreign Speculators and Emerging Equity Markets, *Journal of Finance*, **55**, 565–614.

Bekaert, G. and C. R. Harvey (2002), Research in Emerging Markets Finance: Looking to the Future, *Emerging Markets Review*, **3**, 429–448.

Bekaert, G. and C. R. Harvey (2003), Emerging Markets Finance, *Journal of Empirical Finance*, **10**, 3–55.

Bekaert, G., C. R. Harvey, and C. T. Lundblad (2003), Equity Market Liberalization in Emerging Markets, *Journal of Financial Research*, **26**, 275–299.

Berkes, I. and L. Horváth (2003), Limit Results for the Empirical Process of Squared Residuals in GARCH(p,q) Models, *Stochastic Processes and their Applications*, **105**, 271–298.

Berkes, I., L. Horváth, and P. Kokoszka (2003), GARCH(p,q) Processes: Structure and Estimation, *Bernoulli*, **9**, 201–227.

Boswijk, H. P. (2004), Asymptotic Theory for Integrated Processes, manuscript in preparation, Oxford University Press, parts available at `http://www.fee.uva.nl/ke/boswijk/book.pdf`.

Chen, J. and A. K. Gupta (1997), Testing and Locating Variance Change-Points with Application to Stock Prices, *Journal of the American Statistical Association*, **92**, 739–747.

Chu, C. S. J. (1995), Detecting Parameter Shifts in GARCH Models, *Econometric Theory*, **14**, 241–266.

De Santis, G. and S. Imrohoroglu (1997), Stock Returns and Volatility in Emerging Financial Markets, *Journal of International Money and Finance*, **16**, 561–579.

den Haan, W. J. and A. Levin (1997), A Practioner's Guide to Robust Covariance Matrix Estimation, in C. R. Rao and G. S. Maddala (eds.), *Handbook of Statistics - Vol. 15*, Elsevier Science, Amsterdam, pp. 291–341.

Diebold, F. X. (1986), Modeling the Persistence of Conditional Variances: A Comment, *Econometric Reviews*, **5**, 51–56.

Diebold, F. X. and A. Inoue (2001), Long Memory and Structural Change, *Journal of Econometrics*, **105**, 131–159.

Franses, P. H., M. van der Leij, and R. Paap (2002), Modelling and Forecasting Level Shifts in Absolute Returns, *Journal of Applied Econometrics*, **17**, 601–616.

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

Horváth, L., P. Kokoszka, and G. Teyssière (2001), Empirical Process of the Squared Residuals of an ARCH Sequence, *Annals of Statistics*, **29**, 445–469.

Hsu, D. A. (1977), Tests for Variance Shift at an Unknown Time Point, *Applied Statistics*, **26**, 279–284.

Inclán, C. and G. Tiao (1994), Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance, *Journal of the American Statistical Association*, **89**, 913–923.

Kim, E. H. and V. Singal (2000), Opening Up of Stock Markets: Lessons from Emerging Economies, *Journal of Business*, **73**, 25–66.

Kim, S., S. Cho, and S. Lee (2000), On the CUSUM Test for Parameter Changes in GARCH(1,1) Models, *Communications in Statistics. Theory and Methods*, **29**, 445–462.

Kokoszka, P. and R. Leipus (2000), Change-point Estimation in ARCH Models, *Bernoulli*, **6**, 513–539.

Kokoszka, P. and G. Teyssière (2002), Change-Point Detection in GARCH Models: Asymptotic and Bootstrap Tests, working paper, Utah State University.

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

Lavielle, M. and E. Moulines (2000), Least-Squares Estimation of an Unknown Nummber of Shifts in a Time Series, *Journal of Time Series Analysis*, **21**, 33–59.

Lee, S. and S. Park (2001), The CUSUM of Squares Test for Scale Changes in Infinite Order Moving Order Average Processes, *Scandinavian Journal of Statistics*, **28**, 625–644.

Lee, S., Y. Tokutsu, and K. Maekawa (2003), The Residual CUSUM Test for the Constancy of Parameters in GARCH(1,1) Models, working paper, Seoul National University.

Liu, M. (2000), Modeling Long Memory in Stock Market Volatility, *Journal of Econometrics*, **99**, 139–171.

Lundbergh, S. and T. Teräsvirta (2002), Evaluating GARCH Models, *Journal of Econometrics*, **110**, 417–435.

MacKinnon, J. G. (2000), Computing Numerical Distribution Functions in Econometrics, in A. Pollard, D. Mewhort, and D. Weaver (eds.), *High Performance Computing Systems and Applications*, Kluwer, Amsterdam, pp. 455–470.

Mikosch, T. and C. Starica (2004), Nonstationarities in Financial Time Series, the Long-Range Dependence, and the IGARCH Effects, *Review of Economics and Statistics*, **86**, 378–390.

Richards, A. J. (1996), Volatility and Predictability in National Stock Markets: How Do Emerging and Mature Markets Differ?, *International Monetary Fund Staff Papers*, **43**, 461–501.

Sakata, T. (1988), Detecting a Change in Variances, *Communications in Statistics. Theory and Methods*, **17**, 641–655.

Talwar, P. P. and J. E. Gentle (1981), Detecting a Scale Shift in a Random Sequence at an Unknown Time Point, *Applied Statistics*, **30**, 301–304.

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

Wichern, D. W., R. B. Miller, and D. A. Hsu (1976), Changes of Variance in First-Order Autoregressive Time Series Models, *Applied Statistics*, **25**, 248–256.

Table 1: Empirical rejection frequencies of CUSUM tests for a single change in volatility when DGP is GARCH(1,1)-N

			Tests applied to "raw" series y_t												Tests applied to standardized GARCH(1,1) residuals											
			T=500			T=1000			T=2000			T=4000			T=500			T=1000			T=2000			T=4000		
α	β	a	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010
			$U_{y,\sigma}(k^*)$												$U_{\hat{z},\sigma}(k^*)$											
0.10	0.50		0.265	0.173	0.065	0.278	0.184	0.067	0.290	0.184	0.069	0.296	0.192	0.070	0.070	0.030	0.004	0.083	0.035	0.005	0.089	0.044	0.009	0.094	0.047	0.009
0.10	0.60		0.324	0.219	0.092	0.339	0.236	0.098	0.363	0.246	0.101	0.368	0.254	0.105	0.072	0.029	0.004	0.082	0.036	0.005	0.090	0.043	0.009	0.095	0.046	0.008
0.10	0.70		0.429	0.313	0.158	0.463	0.343	0.174	0.500	0.371	0.182	0.510	0.385	0.192	0.076	0.030	0.003	0.084	0.038	0.006	0.091	0.043	0.009	0.094	0.047	0.008
0.10	0.80		0.655	0.541	0.347	0.724	0.614	0.415	0.781	0.676	0.471	0.816	0.712	0.504	0.081	0.030	0.003	0.085	0.035	0.005	0.090	0.042	0.008	0.094	0.045	0.008
0.20	0.50		0.566	0.445	0.255	0.613	0.493	0.291	0.658	0.538	0.327	0.683	0.557	0.341	0.082	0.034	0.005	0.090	0.040	0.005	0.093	0.045	0.009	0.096	0.047	0.009
0.20	0.60		0.703	0.598	0.399	0.769	0.664	0.467	0.823	0.728	0.531	0.856	0.764	0.566	0.084	0.034	0.003	0.088	0.042	0.006	0.092	0.044	0.009	0.097	0.048	0.008
0.20	0.70		0.877	0.817	0.672	0.934	0.888	0.770	0.971	0.940	0.854	0.985	0.969	0.902	0.085	0.033	0.003	0.088	0.040	0.006	0.091	0.044	0.009	0.095	0.046	0.008
			$U_{y,\gamma_0}(k^*)$												$U_{\hat{z},\gamma_0}(k^*)$											
0.10	0.50		0.245	0.155	0.055	0.256	0.164	0.055	0.266	0.167	0.057	0.270	0.168	0.060	0.074	0.031	0.004	0.084	0.037	0.006	0.092	0.045	0.009	0.095	0.046	0.009
0.10	0.60		0.297	0.194	0.076	0.310	0.210	0.080	0.328	0.217	0.082	0.334	0.224	0.090	0.073	0.029	0.004	0.085	0.038	0.005	0.090	0.044	0.009	0.095	0.046	0.008
0.10	0.70		0.395	0.279	0.127	0.423	0.301	0.145	0.451	0.322	0.150	0.461	0.334	0.156	0.079	0.029	0.003	0.086	0.039	0.005	0.092	0.043	0.008	0.094	0.046	0.007
0.10	0.80		0.611	0.490	0.292	0.668	0.548	0.339	0.722	0.604	0.388	0.750	0.635	0.413	0.084	0.031	0.003	0.086	0.035	0.004	0.091	0.044	0.007	0.096	0.045	0.007
0.20	0.50		0.468	0.344	0.167	0.499	0.374	0.192	0.534	0.404	0.209	0.547	0.416	0.218	0.085	0.034	0.005	0.091	0.042	0.007	0.093	0.046	0.010	0.097	0.047	0.009
0.20	0.60		0.602	0.475	0.276	0.649	0.519	0.315	0.695	0.574	0.351	0.716	0.592	0.370	0.087	0.036	0.004	0.090	0.041	0.007	0.092	0.045	0.009	0.096	0.047	0.007
0.20	0.70		0.800	0.705	0.512	0.857	0.776	0.584	0.902	0.833	0.662	0.932	0.870	0.702	0.088	0.034	0.003	0.090	0.040	0.006	0.093	0.046	0.009	0.096	0.045	0.008
			$U_{y,\varsigma}(k^*)$												$U_{\hat{z},\varsigma}(k^*)$											
0.10	0.50		0.157	0.088	0.024	0.154	0.084	0.020	0.143	0.078	0.019	0.134	0.073	0.016	0.075	0.032	0.005	0.084	0.037	0.006	0.092	0.046	0.009	0.096	0.047	0.008
0.10	0.60		0.193	0.111	0.035	0.186	0.107	0.029	0.175	0.099	0.026	0.158	0.091	0.023	0.076	0.030	0.005	0.086	0.039	0.006	0.090	0.044	0.009	0.095	0.047	0.008
0.10	0.70		0.264	0.164	0.060	0.251	0.161	0.051	0.240	0.146	0.047	0.219	0.132	0.040	0.081	0.031	0.003	0.087	0.039	0.006	0.093	0.043	0.009	0.094	0.046	0.007
0.10	0.80		0.449	0.327	0.153	0.447	0.324	0.152	0.439	0.310	0.142	0.405	0.284	0.123	0.087	0.032	0.003	0.088	0.035	0.005	0.091	0.044	0.008	0.096	0.045	0.007
0.20	0.50		0.186	0.107	0.030	0.172	0.096	0.024	0.158	0.085	0.021	0.144	0.079	0.018	0.086	0.036	0.006	0.092	0.042	0.007	0.094	0.046	0.010	0.097	0.047	0.008
0.20	0.60		0.260	0.161	0.054	0.231	0.141	0.043	0.211	0.124	0.036	0.184	0.105	0.030	0.089	0.037	0.004	0.090	0.041	0.007	0.091	0.045	0.009	0.096	0.047	0.007
0.20	0.70		0.428	0.306	0.135	0.397	0.270	0.115	0.362	0.239	0.091	0.310	0.200	0.071	0.090	0.035	0.003	0.091	0.040	0.006	0.093	0.046	0.009	0.096	0.045	0.007

Note: Table entries indicate fractions of rejection of the null hypothesis of constant volatility against a single structural change across 10000 replications at nominal significance level a, using finite sample critical values obtained from the response surface (9) with $\pi = 0$. Series of length T are generated from a GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1},$$

$z_t \sim \mathsf{iid}\, N(0,1)$, and $\omega = 1 - \alpha - \beta$. Test statistics are applied to the "raw" series y_t (left panel) and to standardized QML residuals $\hat{z}_t$ from a GARCH(1,1)-N model (right panel).

Table 2: Empirical rejection frequencies of the $U_{\hat{z},\varsigma}(k^*)$ test for a single change in volatility when DGP is GARCH(1,1) with alternative error distributions

			T=500			T=1000			T=2000			T=4000		
α	β	a	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010
								GARCH(1,1)-t(5)						
0.10	0.50		0.064	0.027	0.003	0.071	0.028	0.005	0.079	0.035	0.006	0.082	0.042	0.006
0.10	0.60		0.066	0.026	0.003	0.071	0.028	0.005	0.080	0.036	0.005	0.081	0.041	0.006
0.10	0.70		0.066	0.028	0.003	0.074	0.028	0.004	0.079	0.036	0.006	0.080	0.041	0.006
0.10	0.80		0.070	0.028	0.003	0.075	0.030	0.003	0.077	0.035	0.004	0.081	0.040	0.006
0.20	0.50		0.074	0.032	0.004	0.076	0.033	0.005	0.083	0.038	0.007	0.083	0.042	0.006
0.20	0.60		0.075	0.032	0.003	0.078	0.033	0.005	0.083	0.038	0.006	0.083	0.042	0.007
0.20	0.70		0.073	0.030	0.004	0.079	0.032	0.004	0.082	0.038	0.006	0.083	0.042	0.007
								GARCH(1,1)-SN(-5)						
0.10	0.50		0.069	0.027	0.004	0.082	0.038	0.005	0.089	0.043	0.006	0.096	0.046	0.010
0.10	0.60		0.073	0.027	0.004	0.080	0.037	0.005	0.089	0.042	0.006	0.095	0.045	0.010
0.10	0.70		0.075	0.026	0.003	0.083	0.037	0.005	0.090	0.042	0.005	0.093	0.047	0.009
0.10	0.80		0.080	0.029	0.002	0.085	0.034	0.004	0.092	0.040	0.004	0.094	0.046	0.009
0.20	0.50		0.082	0.034	0.005	0.088	0.042	0.006	0.093	0.044	0.006	0.097	0.046	0.010
0.20	0.60		0.082	0.033	0.004	0.088	0.041	0.005	0.093	0.044	0.006	0.095	0.047	0.010
0.20	0.70		0.083	0.030	0.002	0.088	0.038	0.005	0.093	0.043	0.004	0.095	0.046	0.009
								GARCH(1,1)-N with jumps						
0.10	0.50		0.077	0.034	0.005	0.085	0.040	0.007	0.093	0.044	0.007	0.100	0.047	0.008
0.10	0.60		0.080	0.032	0.005	0.088	0.041	0.007	0.093	0.043	0.007	0.100	0.047	0.008
0.10	0.70		0.083	0.034	0.004	0.091	0.043	0.006	0.095	0.044	0.006	0.099	0.048	0.008
0.10	0.80		0.087	0.033	0.004	0.092	0.042	0.005	0.094	0.043	0.005	0.097	0.047	0.007
0.20	0.50		0.090	0.041	0.006	0.094	0.045	0.009	0.096	0.045	0.008	0.101	0.049	0.009
0.20	0.60		0.089	0.041	0.005	0.095	0.045	0.008	0.096	0.046	0.007	0.100	0.049	0.009
0.20	0.70		0.090	0.038	0.003	0.093	0.044	0.006	0.096	0.044	0.006	0.098	0.048	0.008

Note: Table entries indicate fractions of rejection of the null hypothesis of constant volatility against a single structural change across 10000 replications at nominal significance level a, using finite sample critical values obtained from the response surface (9) with $\pi = 0$. Series of length T are generated from (i) a GARCH(1,1)-$t(5)$ process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1},$$

$z_t \sim \mathsf{iid}\, t(5)$ (top panel), (ii) a GARCH(1,1)-$SN(-5)$ process where $z_t \sim \mathsf{iid}\, SN(-5)$ (middle panel) and (iii) a GARCH(1,1)-N-with-jumps process where $y_t = z_t\sqrt{h_t} + \delta D_t$ with $z_t \sim \mathsf{iid}\, N(0,1)$ and D_t is a dummy variable taking the values 1 or -1 (with equal probability) at random time points $t_1, t_2, \ldots, t_{\pi T}$, and 0 otherwise, where $\pi = 0.005$ and $\delta = 5$ (top panel). For all models $\omega = 1 - \alpha - \beta$. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized QML residuals $\hat{z}$ from a GARCH(1,1)-N model.

Table 3: Empirical rejection frequencies of the $U_{\hat{z},\varsigma}(k^*)$ test for a single change in volatility when conditional volatility model is misspecified

			T=500			T=1000			T=2000			T=4000		
γ / d	β	a	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010
							GJR-GARCH(1,1)-N							
0.10	0.50		0.066	0.024	0.005	0.072	0.032	0.006	0.089	0.040	0.008	0.091	0.045	0.008
0.10	0.60		0.066	0.025	0.005	0.072	0.032	0.005	0.089	0.041	0.007	0.092	0.043	0.008
0.10	0.70		0.070	0.026	0.004	0.075	0.035	0.006	0.089	0.042	0.008	0.092	0.044	0.007
0.10	0.80		0.078	0.031	0.005	0.083	0.036	0.006	0.092	0.041	0.007	0.089	0.043	0.007
0.20	0.50		0.073	0.029	0.004	0.082	0.039	0.006	0.097	0.045	0.010	0.094	0.048	0.009
0.20	0.60		0.076	0.029	0.004	0.086	0.039	0.007	0.098	0.046	0.010	0.094	0.045	0.008
0.20	0.70		0.082	0.031	0.004	0.089	0.042	0.007	0.099	0.046	0.010	0.094	0.044	0.008
							FI-GARCH(1,1)-N							
0.40	0.10		0.320	0.186	0.040	0.480	0.351	0.136	0.583	0.461	0.243	0.682	0.560	0.342
0.40	0.30		0.203	0.079	0.009	0.328	0.184	0.029	0.417	0.281	0.095	0.493	0.359	0.160
0.60	0.30		0.251	0.123	0.016	0.393	0.256	0.066	0.484	0.355	0.154	0.574	0.446	0.235
0.60	0.50		0.176	0.057	0.002	0.266	0.128	0.014	0.352	0.220	0.057	0.428	0.303	0.128
0.80	0.50		0.169	0.074	0.009	0.258	0.146	0.032	0.314	0.209	0.075	0.377	0.264	0.122
0.80	0.70		0.138	0.043	0.005	0.181	0.077	0.016	0.259	0.157	0.069	0.381	0.292	0.190
							SV-AR(1)							
0.75			0.064	0.027	0.003	0.070	0.030	0.004	0.073	0.030	0.004	0.072	0.033	0.006
0.80			0.062	0.025	0.003	0.068	0.029	0.004	0.067	0.028	0.003	0.070	0.032	0.005
0.85			0.059	0.022	0.002	0.066	0.026	0.003	0.064	0.026	0.004	0.067	0.029	0.005
0.90			0.055	0.020	0.002	0.058	0.023	0.003	0.057	0.023	0.003	0.060	0.025	0.005
0.95			0.047	0.015	0.001	0.048	0.017	0.001	0.046	0.017	0.001	0.048	0.019	0.002
0.975			0.056	0.015	0.002	0.051	0.019	0.004	0.056	0.026	0.009	0.061	0.034	0.010

Note: Table entries indicate fractions of rejection of the null hypothesis of constant volatility against a single structural change across 10000 replications at nominal significance level a, using finite sample critical values obtained from the response surface (9) with $\pi = 0$. Series of length T are generated from
(i) a GJR-GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha y_{t-1}^2 + \gamma y_{t-1}^2 \mathtt{I}[y_{t-1} < 0] + \beta h_{t-1},$$

$\alpha = 0$, and $\omega = 1 - \gamma/2 - \beta$ (top panel),
(ii) a FI-GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + (1 - \beta - (1 - L)^d) y_{t-1}^2 + \beta h_{t-1},$$

$\omega = 0.10$ (middle panel) and
(iii) a SV-AR(1) process, $y_t = z_t \exp(h_t/2)$, where

$$h_t = \gamma_0 + \gamma_1 h_{t-1} + \eta_t,$$

$\gamma_0 = -(1 - \gamma_1)/2$, $\eta_t \sim \mathsf{iid}\, N(0, \sigma_\eta^2)$ with $\sigma_\eta^2 = 1 - \gamma_1^2$, and z_t and η_t are independent (bottom panel). For all models $z_t \sim \mathsf{iid}\, N(0, 1)$. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized residuals $\hat{z}$ from a GARCH(1,1)-N model.

Table 4: Empirical rejection frequencies of the $U_{\hat{z},\varsigma}(k^*)$ test for a single change in volatility when DGP is GARCH(1,1)-N with break in ω

			T=500			T=1000			T=2000			T=4000		
σ_a^2	β	a	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010
0.50	0.50		0.659	0.473	0.272	0.654	0.514	0.397	0.630	0.481	0.380	0.542	0.417	0.333
0.60	0.50		0.683	0.497	0.225	0.837	0.741	0.595	0.852	0.793	0.729	0.836	0.793	0.747
0.70	0.50		0.521	0.345	0.118	0.840	0.725	0.450	0.962	0.940	0.883	0.988	0.985	0.980
0.80	0.50		0.265	0.158	0.038	0.519	0.389	0.161	0.813	0.722	0.458	0.976	0.968	0.874
0.90	0.50		0.125	0.050	0.011	0.191	0.109	0.019	0.318	0.231	0.079	0.525	0.392	0.201
1.10	0.50		0.110	0.064	0.007	0.143	0.077	0.021	0.284	0.190	0.070	0.471	0.334	0.156
1.20	0.50		0.219	0.117	0.028	0.390	0.246	0.071	0.651	0.528	0.296	0.921	0.868	0.680
1.30	0.50		0.351	0.224	0.055	0.622	0.471	0.226	0.897	0.832	0.616	0.999	0.995	0.964
1.40	0.50		0.504	0.320	0.107	0.783	0.667	0.398	0.950	0.921	0.849	0.993	0.991	0.984
1.50	0.50		0.598	0.425	0.169	0.861	0.778	0.528	0.931	0.897	0.853	0.966	0.949	0.935
0.50	0.80		0.505	0.252	0.053	0.874	0.716	0.293	0.996	0.987	0.916	1.000	1.000	1.000
0.60	0.80		0.426	0.212	0.041	0.776	0.608	0.196	0.981	0.956	0.794	1.000	1.000	0.998
0.70	0.80		0.271	0.117	0.013	0.540	0.351	0.086	0.835	0.724	0.443	0.988	0.977	0.891
0.80	0.80		0.166	0.064	0.010	0.283	0.164	0.022	0.495	0.352	0.150	0.779	0.673	0.407
0.90	0.80		0.098	0.034	0.003	0.129	0.053	0.004	0.207	0.122	0.033	0.296	0.202	0.061
1.10	0.80		0.095	0.044	0.007	0.094	0.049	0.006	0.187	0.111	0.024	0.258	0.168	0.052
1.20	0.80		0.140	0.055	0.010	0.206	0.090	0.015	0.382	0.248	0.085	0.627	0.511	0.249
1.30	0.80		0.208	0.099	0.016	0.345	0.202	0.031	0.601	0.467	0.213	0.891	0.810	0.594
1.40	0.80		0.263	0.129	0.017	0.484	0.314	0.063	0.796	0.680	0.380	0.976	0.960	0.859
1.50	0.80		0.342	0.161	0.026	0.617	0.432	0.114	0.908	0.844	0.550	0.999	0.996	0.970

Note: Table entries indicate fractions of rejection of the null hypothesis of constant volatility against a single structural change across 1000 replications at nominal significance level a, using finite sample critical values obtained from the response surface (9) with $\pi = 0$. Series of length T are generated from a GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega_t + \alpha y_{t-1}^2 + \beta h_{t-1},$$

$z_t \sim \mathsf{iid}\, N(0,1)$, $\alpha = 0.10$, $\omega_t = \omega_b = 1-\alpha-\beta$ if $t \leq \tau T$ and $\omega_t = \sigma_a^2\omega_b$ if $t > \tau T$, with $\tau = 0.50$. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized residuals $\hat{z}$ from a GARCH(1,1)-N model.

Table 5: Empirical rejection frequencies of the $U_{\hat{z},\varsigma}(k^*)$ test for a single change in volatility when DGP is GARCH(1,1)-N with break in β

				T=500			T=1000			T=2000			T=4000		
σ_a^2	β_b	τ	a	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010
0.60	0.50	0.25		0.535	0.358	0.135	0.799	0.713	0.498	0.876	0.814	0.746	0.897	0.827	0.784
0.60	0.50	0.50		0.698	0.528	0.262	0.833	0.738	0.620	0.832	0.764	0.698	0.800	0.733	0.693
0.60	0.50	0.75		0.487	0.298	0.064	0.826	0.697	0.389	0.912	0.881	0.822	0.935	0.922	0.915
0.80	0.50	0.25		0.185	0.077	0.014	0.351	0.248	0.065	0.637	0.513	0.255	0.909	0.845	0.649
0.80	0.50	0.50		0.263	0.155	0.036	0.522	0.396	0.161	0.811	0.725	0.467	0.978	0.970	0.878
0.80	0.50	0.75		0.156	0.071	0.012	0.317	0.186	0.043	0.560	0.438	0.181	0.885	0.788	0.530
1.20	0.50	0.25		0.133	0.067	0.006	0.201	0.104	0.021	0.414	0.287	0.101	0.692	0.554	0.298
1.20	0.50	0.50		0.213	0.113	0.028	0.377	0.233	0.063	0.629	0.500	0.269	0.908	0.838	0.646
1.20	0.50	0.75		0.146	0.084	0.016	0.207	0.129	0.035	0.453	0.323	0.129	0.740	0.644	0.358
1.40	0.50	0.25		0.259	0.139	0.030	0.502	0.333	0.093	0.853	0.744	0.428	0.995	0.988	0.928
1.40	0.50	0.50		0.451	0.292	0.081	0.756	0.634	0.344	0.947	0.915	0.813	0.997	0.994	0.991
1.40	0.50	0.75		0.290	0.163	0.038	0.561	0.400	0.144	0.877	0.783	0.542	0.996	0.993	0.962
0.60	0.80	0.25		0.299	0.135	0.023	0.640	0.440	0.098	0.942	0.867	0.583	1.000	0.998	0.978
0.60	0.80	0.50		0.460	0.257	0.041	0.832	0.661	0.278	0.989	0.975	0.872	1.000	1.000	1.000
0.60	0.80	0.75		0.235	0.093	0.006	0.533	0.315	0.070	0.902	0.790	0.423	0.999	0.998	0.960
0.80	0.80	0.25		0.116	0.045	0.006	0.201	0.093	0.005	0.341	0.224	0.071	0.578	0.441	0.213
0.80	0.80	0.50		0.162	0.058	0.011	0.281	0.159	0.018	0.486	0.349	0.148	0.783	0.674	0.407
0.80	0.80	0.75		0.100	0.031	0.003	0.165	0.082	0.007	0.310	0.193	0.043	0.529	0.377	0.136
1.20	0.80	0.25		0.105	0.043	0.003	0.097	0.048	0.004	0.214	0.119	0.029	0.344	0.215	0.060
1.20	0.80	0.50		0.126	0.052	0.006	0.175	0.080	0.012	0.332	0.208	0.060	0.569	0.429	0.183
1.20	0.80	0.75		0.100	0.043	0.005	0.110	0.056	0.006	0.241	0.131	0.033	0.372	0.252	0.085
1.40	0.80	0.25		0.140	0.049	0.003	0.220	0.085	0.011	0.446	0.278	0.075	0.768	0.617	0.292
1.40	0.80	0.50		0.229	0.104	0.010	0.401	0.226	0.035	0.678	0.535	0.248	0.946	0.890	0.697
1.40	0.80	0.75		0.136	0.057	0.007	0.232	0.112	0.019	0.466	0.329	0.097	0.787	0.692	0.393

Note: Table entries indicate fractions of rejection of the null hypothesis of constant volatility against a single structural change across 1000 replications at nominal significance level a, using finite sample critical values obtained from the response surface (9) with $\pi = 0$. Series of length T are generated from a GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha y_{t-1}^2 + \beta_t h_{t-1},$$

$z_t \sim$ iid $N(0,1)$, $\beta_t = \beta_b$ if $t \leq \tau T$ and $\beta_t = \beta_a$ if $t > \tau T$, where β_a is such that the unconditional volatility after the break is equal to $(\omega/(1-\alpha-\beta_a) =)\ \sigma_a^2$, $\alpha = 0.10$, and $\omega = 1-\alpha-\beta_b$ such that the unconditional volatility before the break is equal to 1. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized residuals $\hat{z}$ from a GARCH(1,1)-N model.

Table 6: Empirical rejection frequencies of the $U_{\hat{z},\varsigma}(k^*)$ test for a single change in volatility when DGP is GARCH(1,1)-N with break in α

				T=500			T=1000			T=2000			T=4000		
σ_a^2	β	τ	a	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010	0.100	0.050	0.010
0.60	0.70	0.25		0.326	0.164	0.026	0.683	0.506	0.167	0.935	0.870	0.650	0.992	0.981	0.954
0.60	0.70	0.50		0.502	0.299	0.053	0.845	0.717	0.382	0.990	0.977	0.891	0.999	0.998	0.994
0.60	0.70	0.75		0.248	0.105	0.011	0.541	0.353	0.078	0.893	0.780	0.448	0.998	0.995	0.961
0.80	0.50	0.25		0.146	0.060	0.011	0.276	0.192	0.041	0.529	0.384	0.176	0.812	0.717	0.481
0.80	0.50	0.50		0.194	0.113	0.020	0.406	0.282	0.093	0.690	0.546	0.310	0.935	0.871	0.716
0.80	0.50	0.75		0.130	0.054	0.006	0.228	0.123	0.026	0.418	0.292	0.097	0.711	0.579	0.283
0.80	0.70	0.25		0.119	0.042	0.005	0.200	0.090	0.009	0.322	0.206	0.078	0.543	0.415	0.185
0.80	0.70	0.50		0.143	0.061	0.007	0.260	0.149	0.025	0.435	0.302	0.130	0.715	0.597	0.331
0.80	0.70	0.75		0.104	0.029	0.001	0.162	0.073	0.008	0.268	0.159	0.040	0.447	0.308	0.101
1.20	0.50	0.25		0.105	0.066	0.005	0.128	0.071	0.014	0.251	0.142	0.044	0.386	0.254	0.096
1.20	0.50	0.50		0.150	0.083	0.014	0.207	0.114	0.037	0.390	0.278	0.108	0.629	0.520	0.263
1.20	0.50	0.75		0.117	0.065	0.011	0.152	0.085	0.024	0.289	0.188	0.061	0.453	0.333	0.141
1.20	0.70	0.25		0.088	0.043	0.006	0.100	0.052	0.009	0.170	0.108	0.028	0.247	0.152	0.040
1.20	0.70	0.50		0.112	0.058	0.007	0.140	0.078	0.015	0.250	0.168	0.058	0.423	0.294	0.122
1.20	0.70	0.75		0.098	0.050	0.010	0.106	0.060	0.010	0.209	0.116	0.031	0.306	0.201	0.059
1.40	0.50	0.25		0.163	0.089	0.014	0.239	0.129	0.029	0.474	0.345	0.124	0.785	0.650	0.371
1.40	0.50	0.50		0.275	0.158	0.042	0.455	0.321	0.108	0.726	0.618	0.370	0.962	0.923	0.780
1.40	0.50	0.75		0.194	0.114	0.023	0.305	0.208	0.060	0.567	0.449	0.219	0.852	0.768	0.572
1.40	0.70	0.25		0.119	0.066	0.007	0.157	0.082	0.014	0.315	0.191	0.059	0.529	0.395	0.151
1.40	0.70	0.50		0.178	0.086	0.014	0.282	0.165	0.042	0.508	0.387	0.168	0.785	0.693	0.456
1.40	0.70	0.75		0.128	0.069	0.012	0.191	0.110	0.019	0.366	0.251	0.081	0.638	0.492	0.250

Note: Table entries indicate fractions of rejection of the null hypothesis of constant volatility against a single structural change across 1000 replications at nominal significance level a, using finite sample critical values obtained from the response surface (9) with $\pi = 0$. Series of length T are generated from a GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha_t y_{t-1}^2 + \beta h_{t-1},$$

$z_t \sim \text{iid}\, N(0,1)$, $\alpha_t = \alpha_b = 0.15$ if $t \leq \tau T$ and $\alpha_t = \alpha_a$ if $t > \tau T$, where α_a is such that the unconditional volatility after the break is equal to $(\omega/(1-\alpha_a-\beta) =)\ \sigma_a^2$, and $\omega = 1-\alpha_b-\beta$ such that the unconditional volatility before the break is equal to 1. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized residuals $\hat{z}$ from a GARCH(1,1)-N model.

Table 7: Number of identified change points for $U_{\hat{z},\varsigma}(k^*)$ test for changes in volatility when the DGP is GARCH(1,1)-N with two breaks in ω

				T=1000				T=2000				T=4000			
σ^2_{a1}	σ^2_{a2}	β	l	0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3
0.70	0.40	0.50		0.627	0.282	0.088	0.003	0.701	0.086	0.200	0.013	0.746	0.004	0.180	0.070
0.80	0.60	0.50		0.245	0.725	0.030	0.000	0.101	0.727	0.163	0.009	0.063	0.327	0.551	0.059
0.90	0.80	0.50		0.724	0.260	0.016	0.000	0.482	0.493	0.025	0.000	0.160	0.790	0.049	0.001
0.80	1.00	0.50		0.915	0.070	0.015	0.000	0.782	0.099	0.115	0.004	0.528	0.064	0.385	0.023
0.80	1.20	0.50		0.639	0.266	0.095	0.000	0.256	0.330	0.403	0.011	0.022	0.111	0.824	0.043
1.20	0.80	0.50		0.570	0.318	0.112	0.000	0.167	0.458	0.365	0.010	0.010	0.290	0.670	0.030
1.20	1.00	0.50		0.938	0.048	0.014	0.000	0.850	0.081	0.068	0.001	0.722	0.084	0.182	0.012
1.10	1.20	0.50		0.835	0.155	0.010	0.000	0.616	0.371	0.013	0.000	0.320	0.650	0.028	0.002
1.20	1.40	0.50		0.494	0.489	0.017	0.000	0.150	0.820	0.029	0.001	0.009	0.836	0.144	0.011
1.30	1.60	0.50		0.277	0.700	0.023	0.000	0.079	0.806	0.111	0.004	0.042	0.474	0.439	0.045
0.70	0.40	0.80		0.377	0.567	0.055	0.001	0.036	0.703	0.252	0.009	0.000	0.218	0.715	0.067
0.80	0.60	0.80		0.556	0.424	0.020	0.000	0.166	0.783	0.049	0.002	0.006	0.839	0.138	0.017
0.90	0.80	0.80		0.881	0.106	0.012	0.001	0.742	0.243	0.014	0.001	0.502	0.470	0.028	0.000
0.80	1.00	0.80		0.959	0.034	0.007	0.000	0.910	0.060	0.029	0.001	0.819	0.083	0.093	0.005
0.80	1.20	0.80		0.886	0.096	0.018	0.000	0.639	0.244	0.113	0.004	0.264	0.293	0.415	0.028
1.20	0.80	0.80		0.834	0.146	0.020	0.000	0.581	0.315	0.098	0.006	0.199	0.495	0.302	0.004
1.20	1.00	0.80		0.964	0.032	0.004	0.000	0.919	0.068	0.013	0.000	0.886	0.073	0.040	0.001
1.10	1.20	0.80		0.933	0.065	0.002	0.000	0.812	0.181	0.007	0.000	0.638	0.334	0.025	0.003
1.20	1.40	0.80		0.782	0.211	0.007	0.000	0.509	0.481	0.010	0.000	0.170	0.780	0.046	0.004
1.30	1.60	0.80		0.605	0.378	0.017	0.000	0.220	0.756	0.021	0.003	0.021	0.862	0.110	0.007

Note: Table entries indicate fractions of replications for which l structural changes in volatility were found across 1000 replications using the sequential procedure described in Section 2.2. Series are generated from a GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha y^2_{t-1} + \beta_t h_{t-1},$$

$z_t \sim$ iid $N(0,1)$, $\alpha = 0.1$, $\omega_t = \omega_b = 1 - \alpha - \beta$ if $t \leq \tau_1 T$, $\omega_t = \sigma^2_{a1}\omega_b$ if $\tau_1 T \leq t \leq \tau_2 T$ and $\omega_t = \sigma^2_{a2}\omega_b$ if $t \geq \tau_2 T$ with $\tau_1 = 0.33$ and $\tau_1 = 0.67$. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized residuals $\hat{z}$ from a GARCH(1,1)-N model. Finite sample critical values are obtained from the response surface (9) with $\pi = 0.15$, for initial nominal significance level $a = 0.05$.

Table 8: Number of identified change points for the $U_{\hat{z},\varsigma}(k^*)$ test for changes in volatility when the DGP is GARCH(1,1)-N with two breaks in β

				T=1000				T=2000				T=4000			
σ^2_{a1}	σ^2_{a2}	β	l	0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3
0.70	0.40	0.50		0.666	0.241	0.093	0.000	0.789	0.059	0.141	0.011	0.775	0.007	0.149	0.069
0.80	0.60	0.50		0.227	0.735	0.038	0.000	0.123	0.689	0.174	0.014	0.092	0.285	0.560	0.063
0.90	0.80	0.50		0.727	0.256	0.017	0.000	0.473	0.500	0.027	0.000	0.156	0.789	0.053	0.002
0.80	1.00	0.50		0.911	0.073	0.016	0.000	0.779	0.101	0.116	0.004	0.521	0.063	0.394	0.022
0.80	1.20	0.50		0.642	0.266	0.092	0.000	0.272	0.322	0.398	0.008	0.020	0.110	0.826	0.044
1.20	0.80	0.50		0.575	0.316	0.109	0.000	0.172	0.470	0.349	0.009	0.009	0.311	0.652	0.028
1.20	1.00	0.50		0.941	0.046	0.013	0.000	0.863	0.075	0.061	0.001	0.735	0.078	0.177	0.010
1.10	1.20	0.50		0.849	0.139	0.012	0.000	0.633	0.354	0.013	0.000	0.341	0.632	0.025	0.002
1.20	1.40	0.50		0.536	0.449	0.015	0.000	0.186	0.782	0.031	0.001	0.014	0.864	0.113	0.009
1.30	1.60	0.50		0.319	0.652	0.029	0.000	0.063	0.851	0.082	0.004	0.007	0.579	0.381	0.033
0.70	0.40	0.80		0.295	0.614	0.087	0.004	0.075	0.582	0.333	0.010	0.011	0.150	0.725	0.114
0.80	0.60	0.80		0.504	0.476	0.020	0.000	0.123	0.822	0.051	0.004	0.002	0.792	0.188	0.018
0.90	0.80	0.80		0.877	0.112	0.011	0.000	0.753	0.234	0.012	0.001	0.511	0.464	0.025	0.000
0.80	1.00	0.80		0.959	0.033	0.008	0.000	0.902	0.071	0.026	0.001	0.821	0.087	0.087	0.005
0.80	1.20	0.80		0.889	0.095	0.016	0.000	0.670	0.235	0.091	0.004	0.314	0.289	0.378	0.019
1.20	0.80	0.80		0.850	0.129	0.021	0.000	0.629	0.292	0.077	0.002	0.240	0.518	0.239	0.003
1.20	1.00	0.80		0.965	0.031	0.004	0.000	0.928	0.062	0.010	0.000	0.906	0.062	0.031	0.001
1.10	1.20	0.80		0.938	0.058	0.004	0.000	0.838	0.156	0.006	0.000	0.710	0.267	0.021	0.002
1.20	1.40	0.80		0.866	0.129	0.005	0.000	0.630	0.359	0.011	0.000	0.297	0.659	0.042	0.002
1.30	1.60	0.80		0.733	0.255	0.012	0.000	0.390	0.595	0.014	0.001	0.062	0.870	0.063	0.005

Note: Table entries indicate fractions of replications for which l structural changes in volatility were found across 1000 replications using the sequential procedure described in Section 2.2. Series are generated from a GARCH(1,1)-N process, $y_t = z_t\sqrt{h_t}$, where

$$h_t = \omega + \alpha y_{t-1}^2 + \beta_t h_{t-1},$$

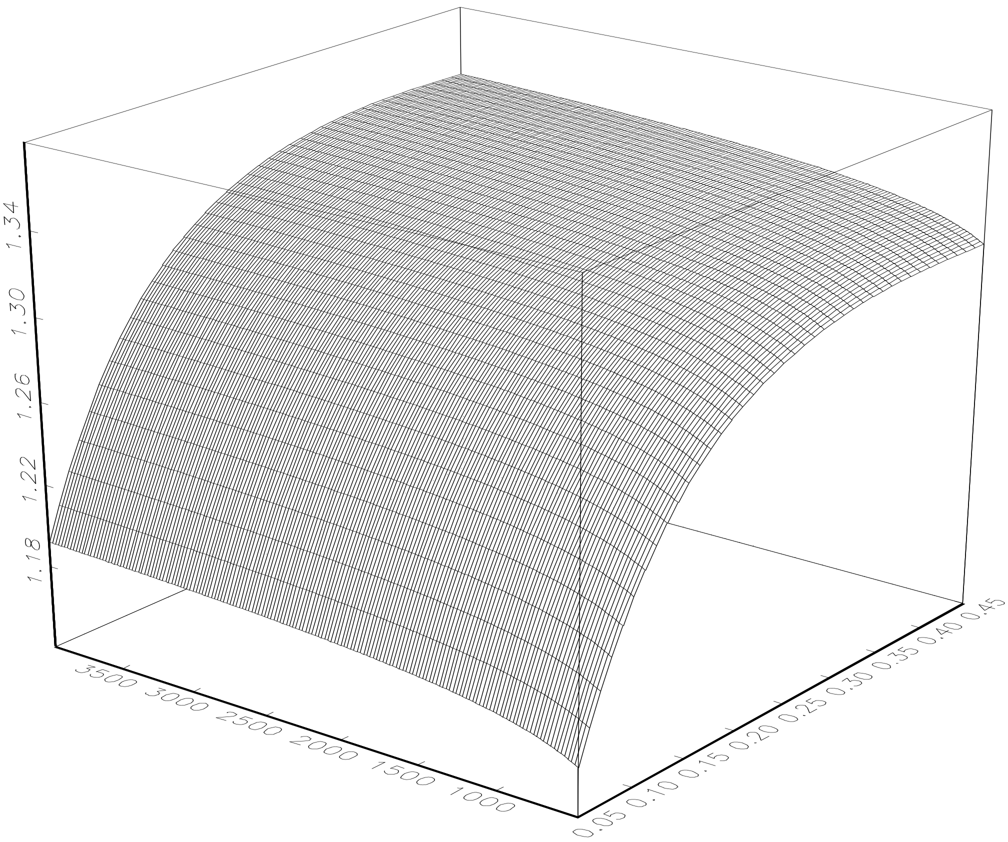
$z_t \sim \text{iid } N(0,1)$, $\beta_t = \beta_b$ if $t \leq \tau_1 T$, $\beta_t = \beta_{a_1}$ if $\tau_1 T < t \leq \tau_2 T$ and $\beta_t = \beta_{a_2}$ if $\tau_2 T < t$, $\alpha = 0.10$, and $\omega = 1-\alpha-\beta_b$ with $\tau_1 = 0.33$ and $\tau_1 = 0.67$. β_{a_1} and β_{a_2} are such that the unconditional volatility $\omega/(1-\alpha-\beta_{a_i}) = \sigma^2_{a_i}, i = 1, 2$. The $U_{\hat{z},\varsigma}(k^*)$ statistic is applied to standardized residuals $\hat{z}$ from a GARCH(1,1)-N model. Finite sample critical values are obtained from the response surface (9) with $\pi = 0.15$, for initial nominal significance level $a = 0.05$.

Table 9: Number of identified break points for emerging stock market returns

	Start date	T	US Dollar returns							local currency returns						
				y_t			$\hat{z}_t$				y_t			$\hat{z}_t$		
			ICSS	σ	γ_0	ς	σ	γ_0	ς	ICSS	σ	γ_0	ς	σ	γ_0	ς
Argentina	1/1/1988	4007	36	10	10	4	2	0	0	36	10	10	1	2	0	0
Brazil	1/1/1988	4006	16	10	6	4	0	0	0	19	9	7	5	1	0	0
Chile	1/1/1988	4035	25	10	10	1	4	1	1	27	10	3	1	3	1	1
China	1/1/1993	2835	19	9	4	3	3	0	0	19	9	4	3	3	0	0
Colombia	1/1/1993	2739	24	8	2	2	3	0	0	20	7	3	2	3	0	0
Czech Republic	1/2/1995	2325	9	7	7	5	1	0	0	16	7	7	5	1	0	0
Egypt	1/2/1995	1945	18	8	7	2	2	0	0	23	8	7	3	1	0	0
Hungary	1/2/1995	2326	14	6	3	3	4	0	0	23	8	6	4	5	0	0
India	1/1/1993	2790	13	9	6	4	2	0	0	14	10	6	4	2	0	0
Indonesia	1/1/1988	4017	38	10	4	3	8	3	3	48	10	0	0	6	4	4
Israel	1/1/1993	2851	19	10	9	4	4	0	0	21	10	7	6	4	0	0
Jordan	1/1/1988	2958	36	8	0	0	3	0	0	29	9	1	0	4	0	0
Korea	1/1/1988	4012	12	9	7	7	3	2	2	15	8	6	6	3	1	1
Malaysia	1/1/1988	4069	40	10	5	4	5	0	0	30	10	4	4	7	0	0
Mexico	1/1/1988	4078	24	10	6	3	7	0	0	22	10	6	0	4	0	0
Morocco	1/2/1995	2334	20	4	2	2	3	1	1	21	9	5	1	2	1	1
Pakistan	1/1/1993	2506	25	9	4	4	3	0	0	25	10	4	3	0	0	0
Peru	1/1/1993	2785	26	10	6	1	4	1	1	32	10	2	1	2	1	1
Philippines	1/1/1988	4043	32	10	7	4	6	0	0	32	10	8	6	6	0	0
Poland	1/1/1993	2822	9	9	8	5	3	0	0	10	8	8	6	3	0	0
Russia	1/2/1995	2288	21	10	7	4	2	2	2	21	10	7	4	2	2	2
South Africa	1/1/1993	2839	19	10	7	3	4	0	0	19	7	3	3	2	2	2
Sri Lanka	1/1/1993	2679	36	9	6	4	7	0	0	33	10	6	1	8	0	0
Taiwan	1/1/1988	4006	25	10	6	4	2	0	0	18	10	6	4	2	0	0
Thailand	1/1/1988	4068	36	10	10	3	4	3	3	40	10	10	3	4	3	3
Turkey	1/1/1988	4057	30	10	10	8	2	1	1	30	10	8	3	3	1	1
Venezuela	1/1/1993	2731	31	9	0	0	7	0	0	25	10	0	0	6	0	0

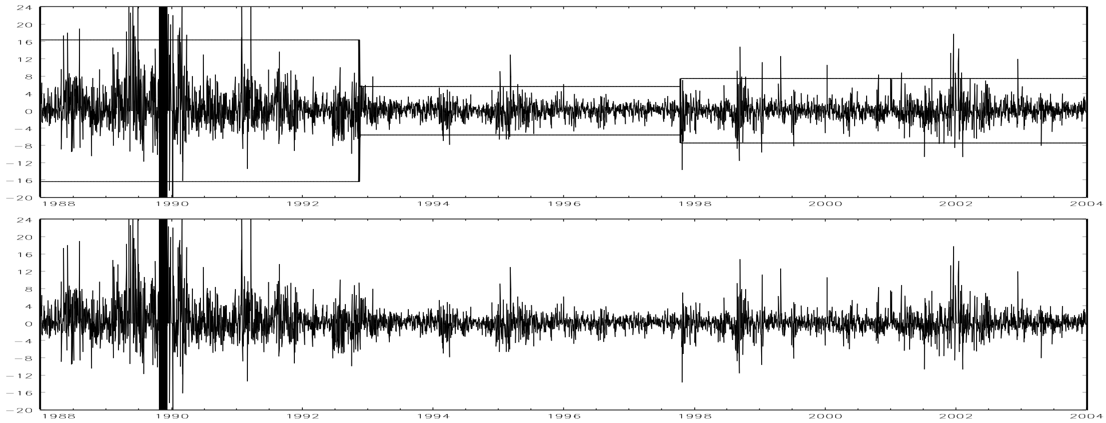
Note: Table entries represent the number of identified break in volatility of daily emerging stock market returns, both in US dollars and local currency. Results are reported for the original ICSS algorithm of Inclán and Tiao (1994), and for the sequential procedure described in Section 2.2 using the $U_{\cdot,\sigma}(k^*)$, $U_{\cdot,\gamma_0}(k^*)$, and $U_{\cdot,\varsigma}(k^*)$ statistics (in columns headed σ, γ_0 and ς, respectively), applied to demeaned returns y_t as well as GARCH(1,1)-standardized returns $\hat{z}_t$. The maximum allowed number of breaks is set equal to 10, while consecutive breaks have to be at least 126 (trading) days apart. Finite sample critical values obtained from (9), with an initial nominal significance level $a = 0.05$.

Figure 1: Response Surface

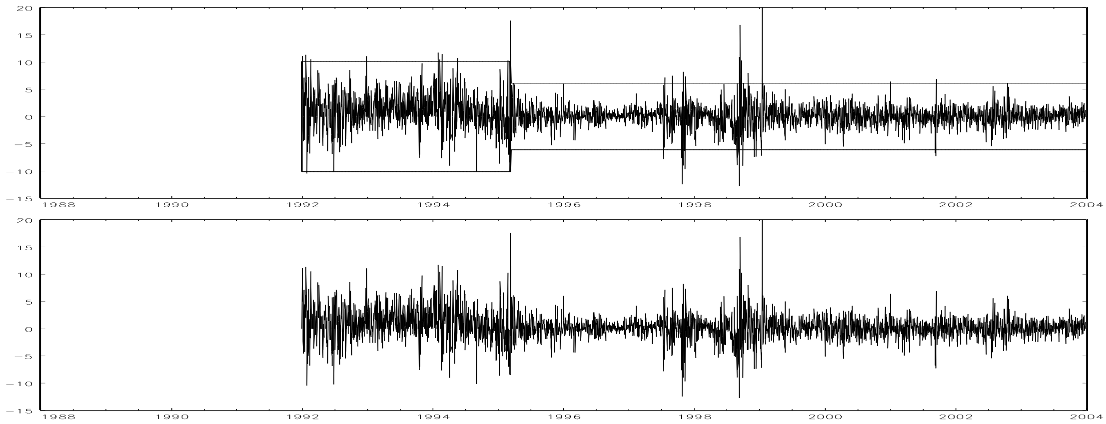


Note: Response surface for the .95 quantile $q_e^a(T, \pi)$ of the distribution of $U_{y,\varsigma}(k^*)$ test when applied to a sample of length T, discarding a fraction of π observations at both ends of the sample when computing the test statistic.

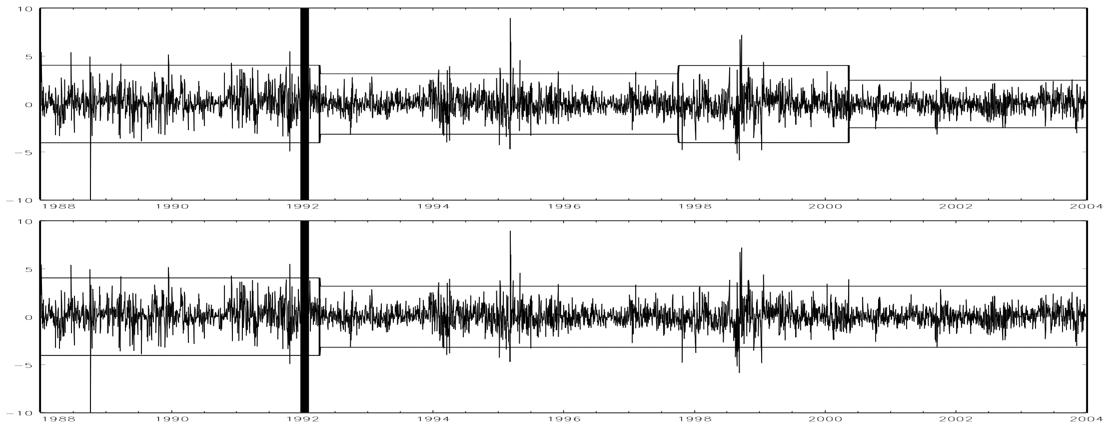
Figure 2: Break Point Locations



(i) Argentina



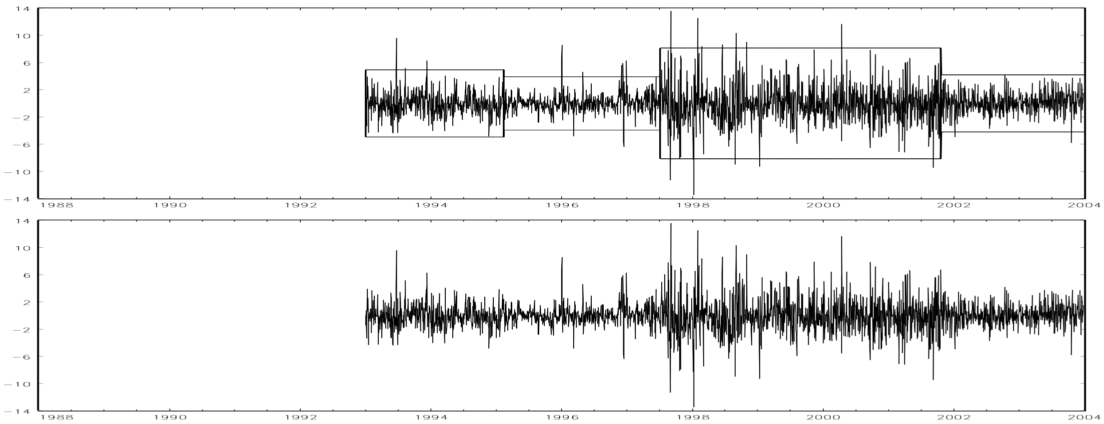
(ii) Brazil



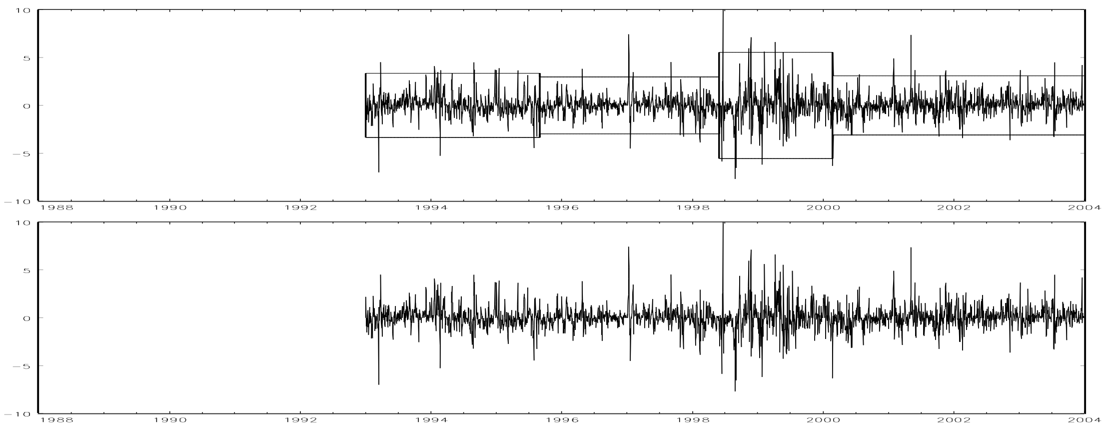
(iii) Chile

Note: The figure shows the daily emerging stock market returns in local currency. The full sample period is January 4, 1988 - December 31, 2003. The horizontal lines in these graphs indicate ± 3 times the unconditional standard deviation between consecutive change-points, as identified by the $U_{\hat{z},\sigma}(k^*)$ and $U_{\hat{z},\varsigma}(k^*)$ statistics, in the upper and lower panels, respectively. Thick vertical lines correspond with the "official liberalization dates" as determined by Bekaert and Harvey (2000b) and Bekaert *et al.* (2003).

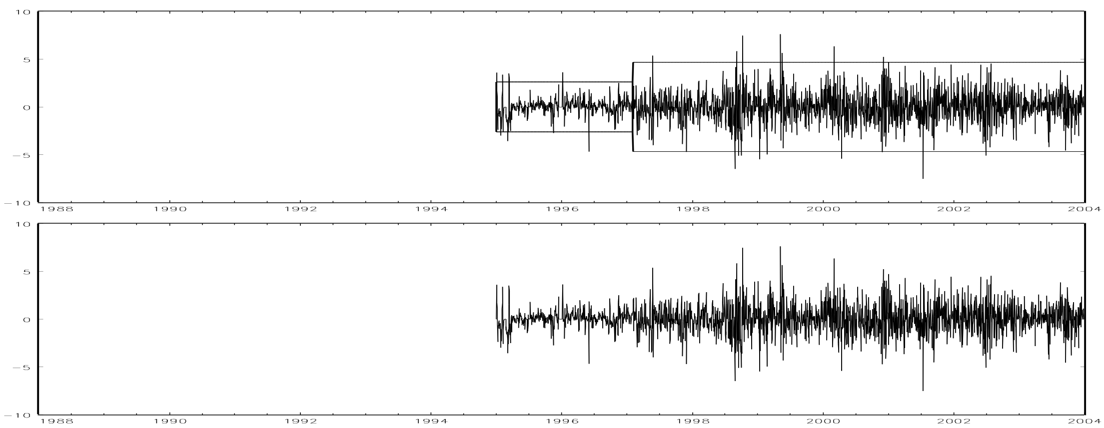
Figure 2 (continued)



(iv) China

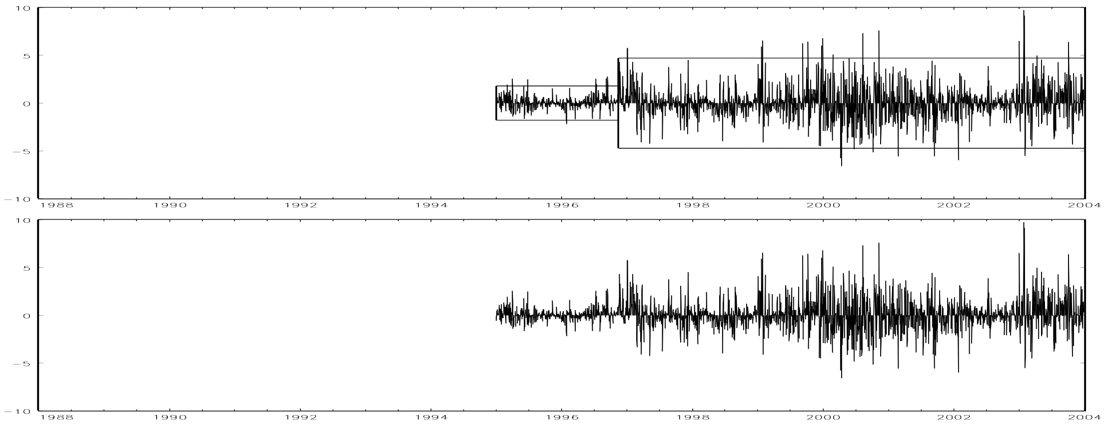


(v) Colombia

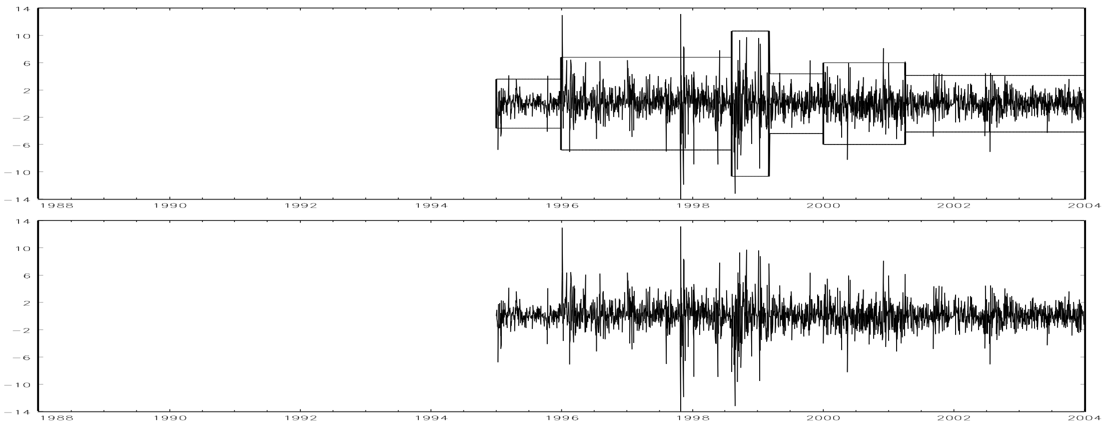


(vi) Czech Republic

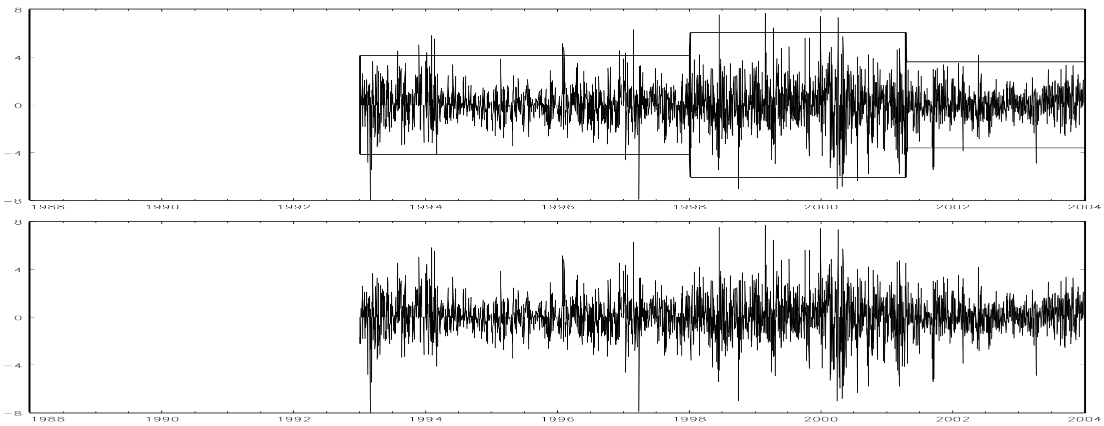
Figure 2 (continued)



(vii) Egypt

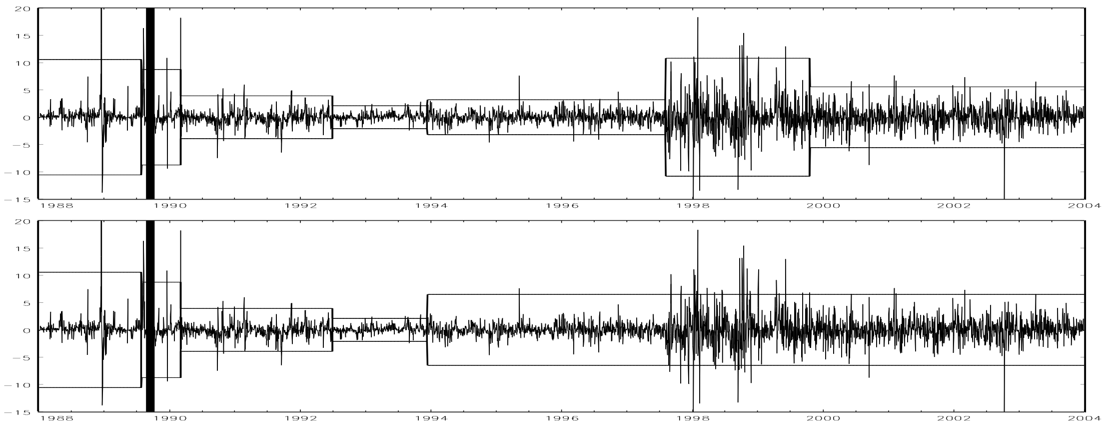


(viii) Hungary

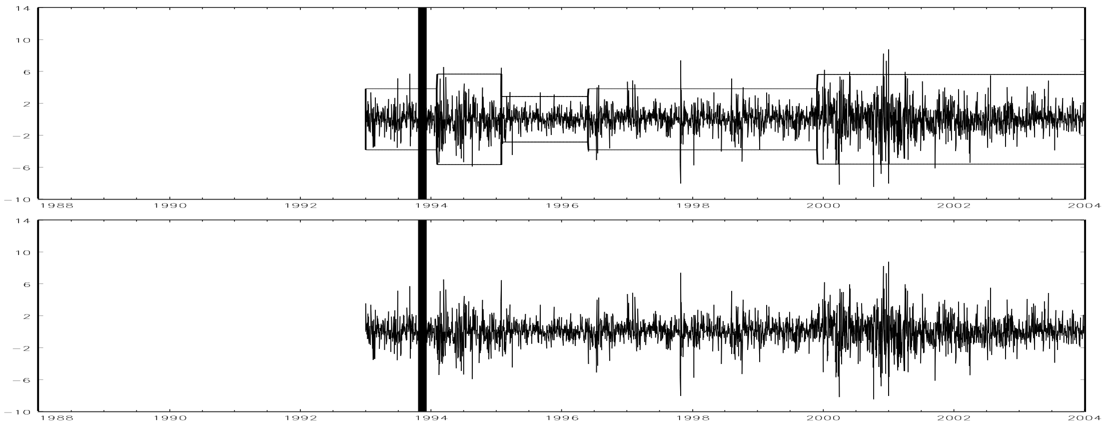


(ix) India

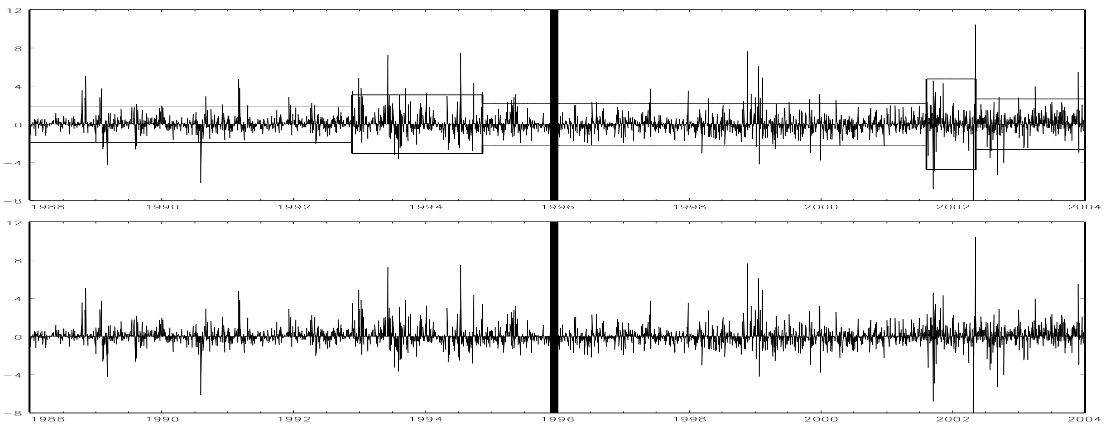
Figure 2 (continued)



(x) Indonesia

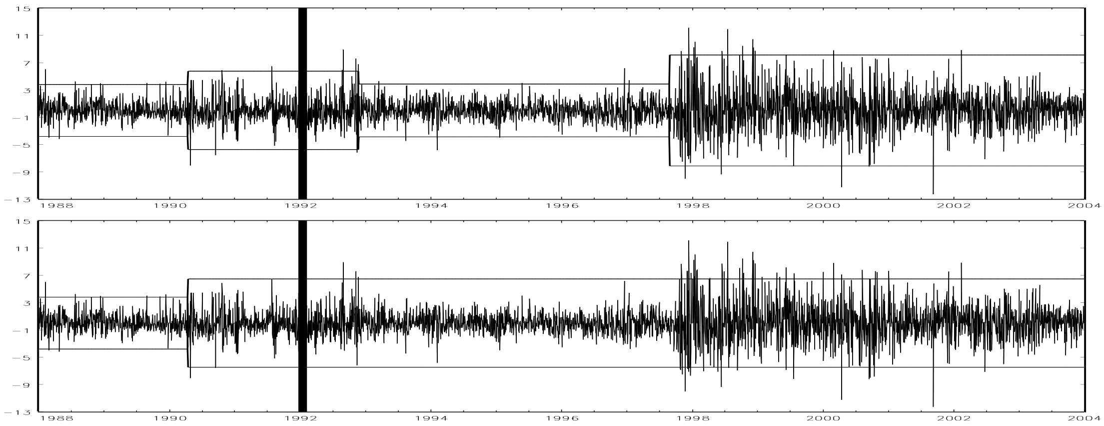


(xi) Israel

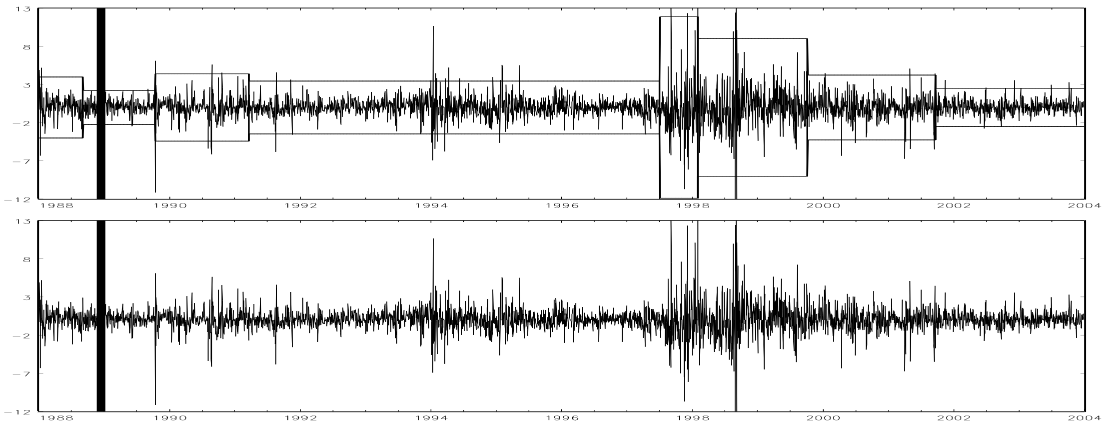


(xii) Jordan

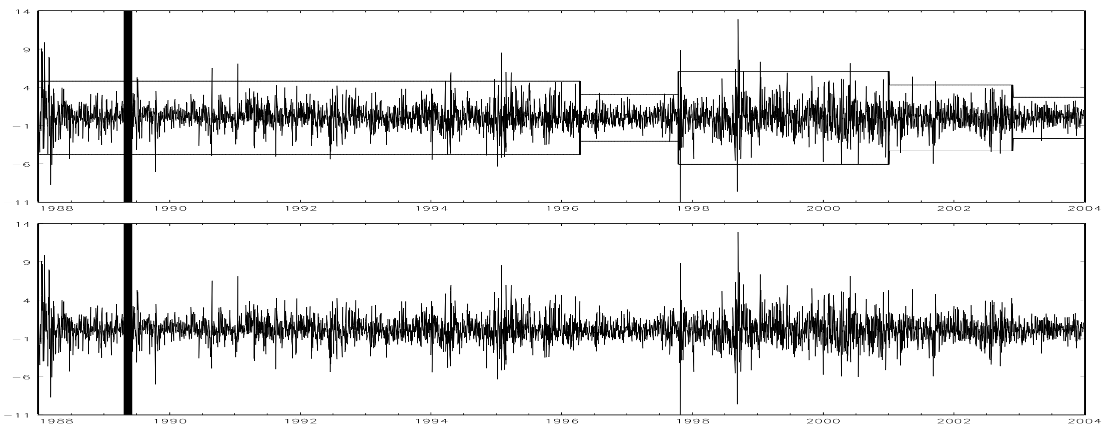
Figure 2 (continued)



(xiii) Korea

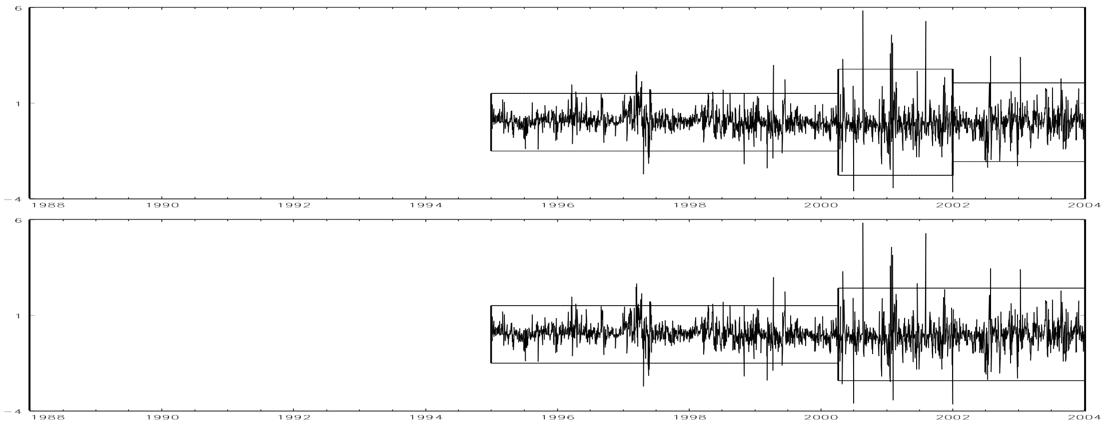


(xiv) Malaysia

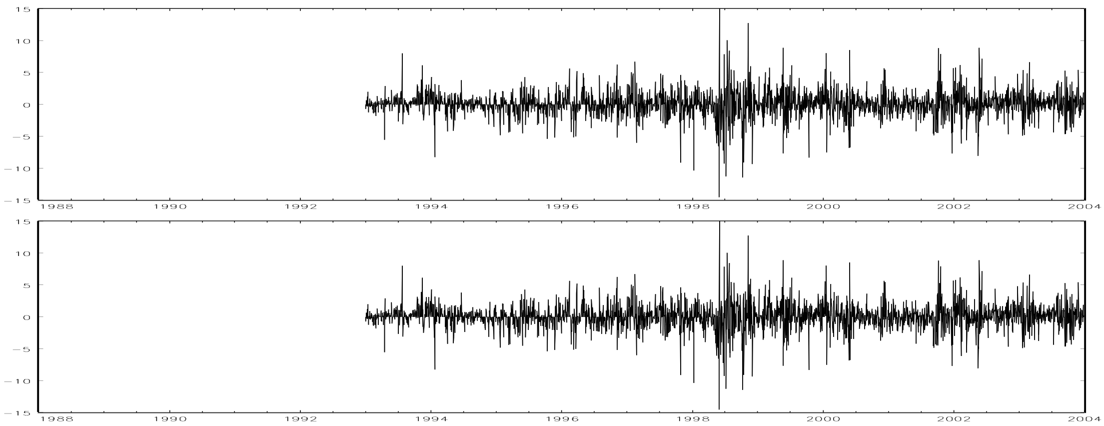


(xv) Mexico

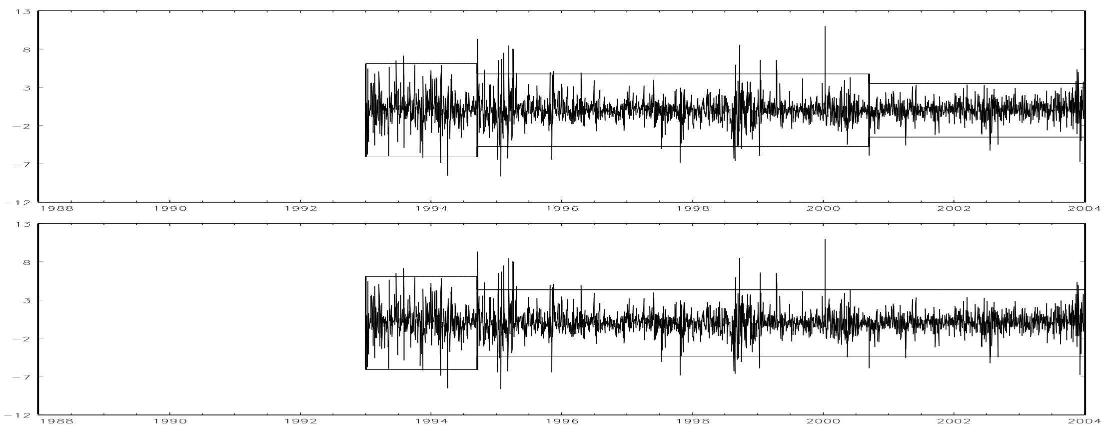
Figure 2 (continued)



(xvi) Morocco

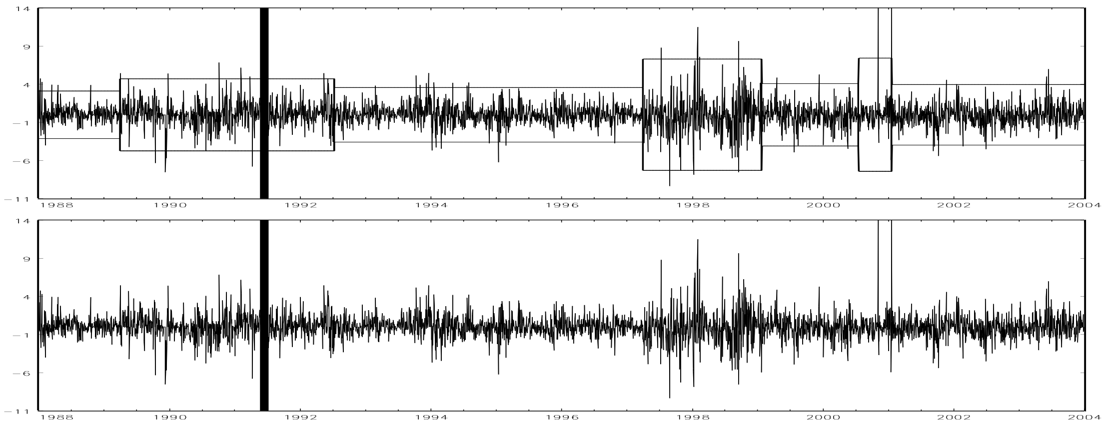


(xvii) Pakistan

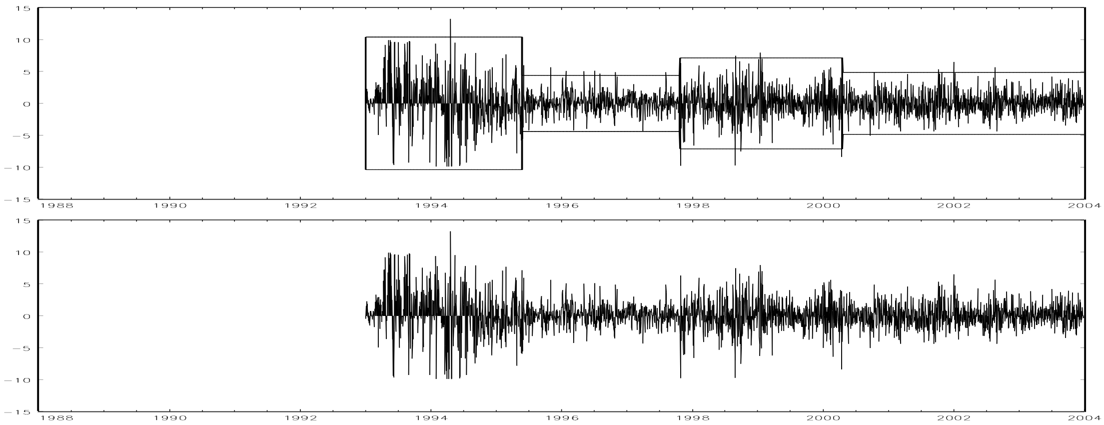


(xviii) Peru

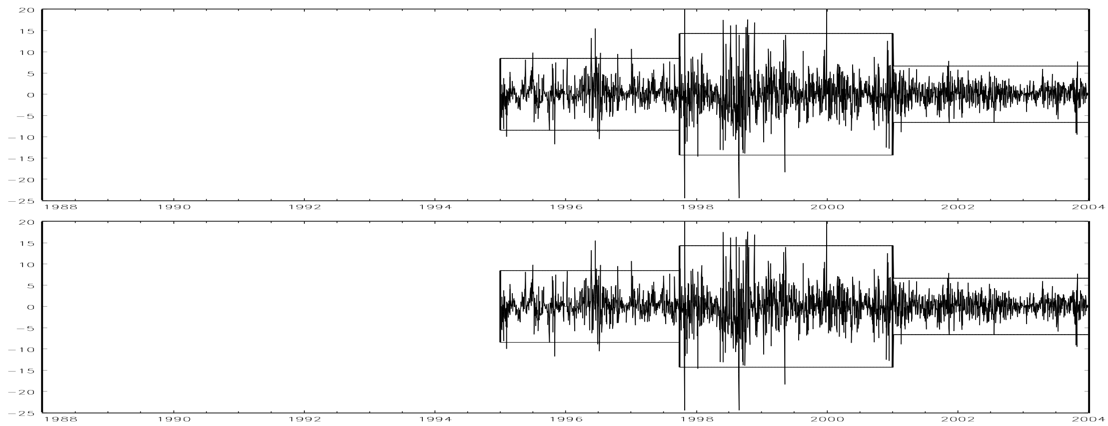
Figure 2 (continued)



(xix) Philippines

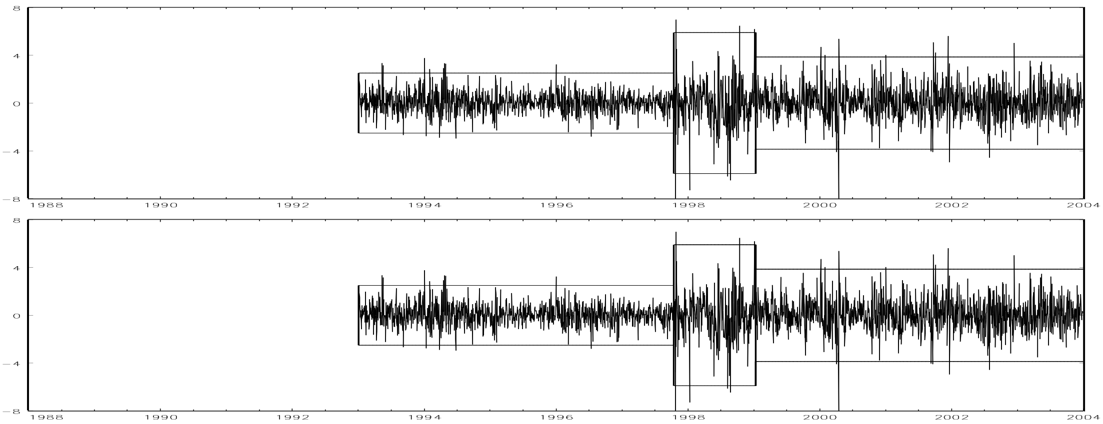


(xx) Poland

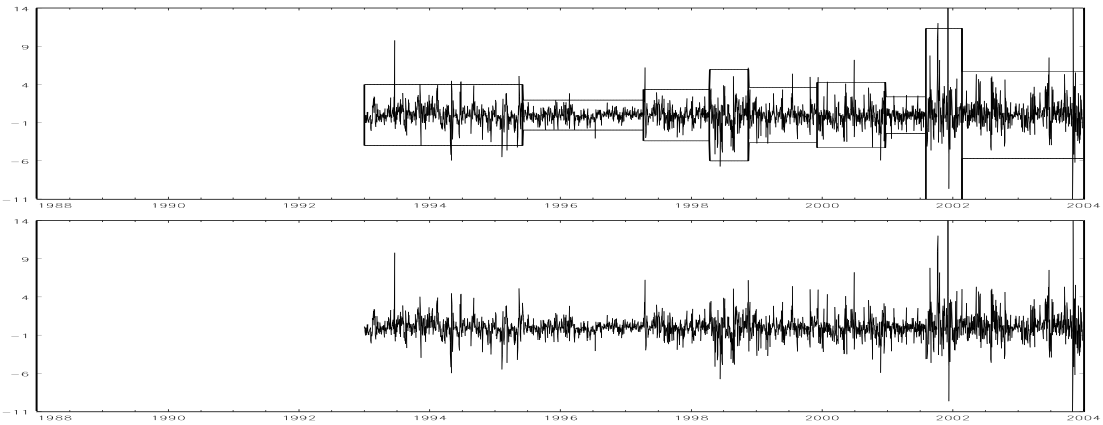


(xxi) Russia

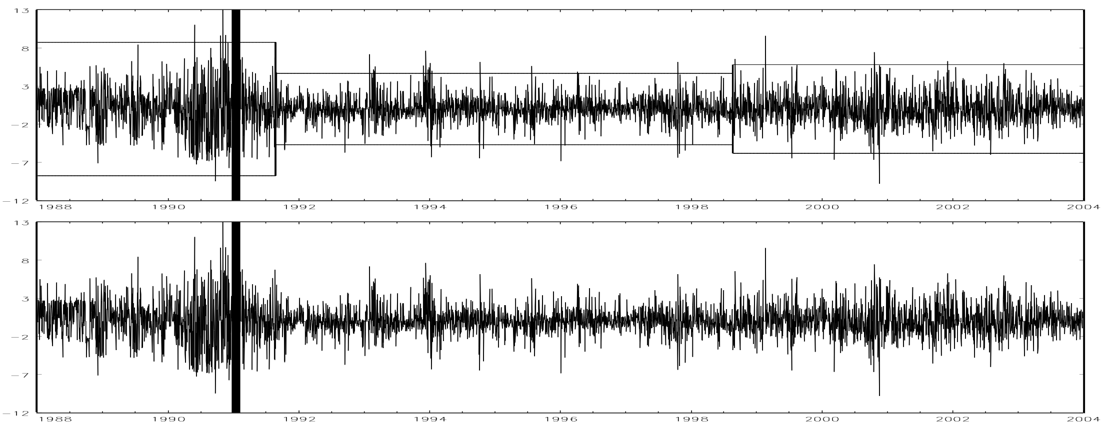
Figure 2 (continued)



(xxii) South Africa

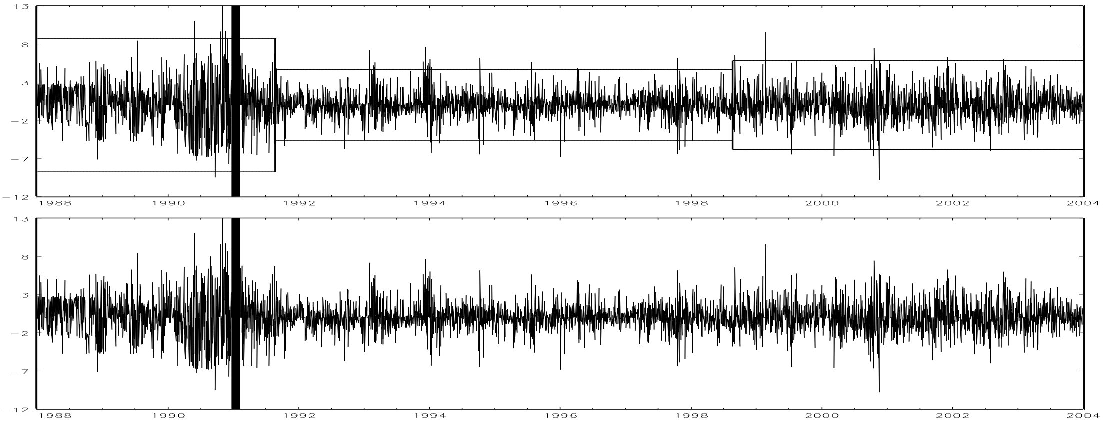


(xxiii) Sri Lanka

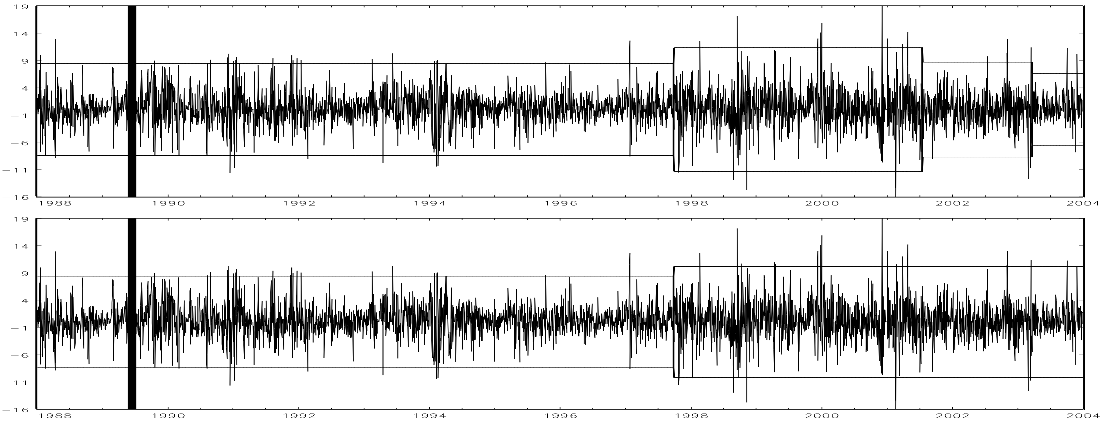


(xxiv) Taiwan

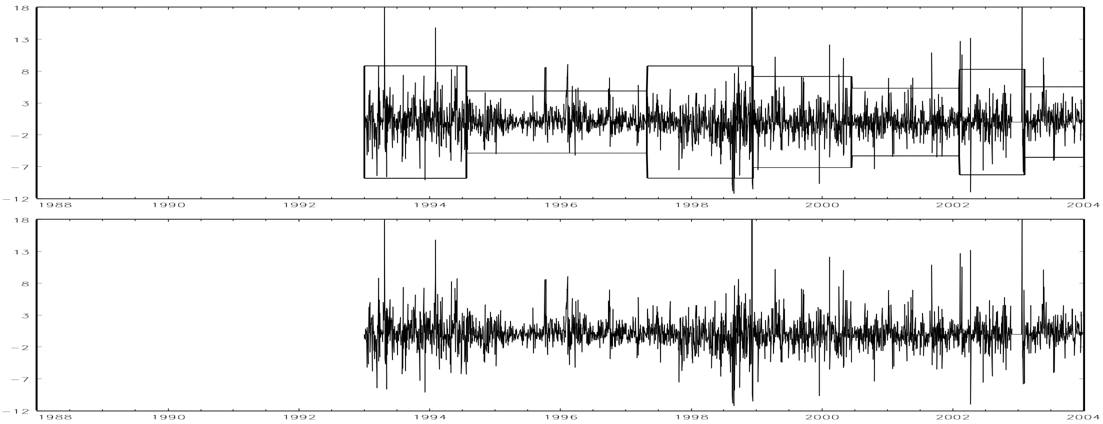
Figure 2 (continued)



(xxv) Thailand



(xxvi) Turkey



(xxvii) Venezuela