DETECTING MULTIPLE BREAKS IN FINANCIAL MARKET VOLATILITY DYNAMICS

ELENA ANDREOU[a] AND ERIC GHYSELS[b]*

[a] *University of Cyprus*
[b] *University of North Carolina and CIRANO*

SUMMARY

The paper evaluates the performance of several recently proposed tests for structural breaks in the conditional variance dynamics of asset returns. The tests apply to the class of ARCH and SV type processes as well as data-driven volatility estimators using high-frequency data. In addition to testing for the presence of breaks, the statistics identify the number and location of multiple breaks. We study the size and power of the new tests for detecting breaks in the conditional variance under various realistic univariate heteroscedastic models, change-point hypotheses and sampling schemes. The paper concludes with an empirical analysis using data from the stock and FX markets for which we find multiple breaks associated with the Asian and Russian financial crises. These events resulted in changes in the dynamics of volatility of asset returns in the samples prior and post the breaks. Copyright © 2002 John Wiley & Sons, Ltd.

1. INTRODUCTION

On this twentieth anniversary of Rob Engle's seminal paper on ARCH it is worth reflecting on some of the outstanding questions in the literature. It has long been conjectured that stock market volatility exhibits occasional breaks. Diebold (1986), Hendry (1986) and Lamoureux and Lastrapes (1990) were among the first to suggest that persistence in volatility may be overstated with the presence of structural breaks. More recent and related evidence is provided by Diebold and Inoue (2001), Granger and Hyung (1999), Mikosch and Starica (1999), among others, which shows that the presence of breaks may also explain the findings of long memory, particularly in volatility.

There is a substantial literature on testing for the presence of breaks in linearly dependent stochastic processes (see, for instance, Bai, 1994, 1997; Bai and Perron, 1998, *inter alia*). There is a temptation to apply the tests for ARMA-type processes in the context of ARCH or Stochastic Volatility (SV) models. For instance, one could view squared returns as an ARMA process and proceed with the application of tests suggested for testing breaks in the mean. Unfortunately, things are not so simple. The resemblance between ARMA and GARCH or discrete time SV models is deceiving (see e.g. Francq and Zakoïan, 2000a,b). Carrasco and Chen (2001) present a comprehensive study which shows that most univariate GARCH processes are β-mixing. This result precludes the application of many aforementioned tests for structural breaks that require uniform mixing conditions.[1]

---

* Correspondence to: Eric Ghysels, Department of Economics, Gardner Hall, CB 3305, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599-3305, USA. E-mail: eghysels@email.unc.edu
[1] Most tests proposed for linear processes impose ϕ-mixing or uniform mixing conditions which are not satisfied by ARCH processes. For a general treatment of estimating the weak GARCH models, see Francq and Zakoïan (2000b).

The purpose of this paper is to explore recent advances in the theory of change-point estimation for strongly dependent processes including ARCH and SV models. Some early attempts to test for a break in a GARCH are found in Chu (1995) and Lundbergh and Terasvirta (1998). A number of recent papers have shown the consistency of CUSUM type change-point estimators for a single break and least squares tests for multiple breaks.[2] The tests are not model-specific and apply to a large class of (strongly) dependent ARCH and SV type specifications given appropriate stationarity conditions. The theoretical developments are described in a series of recent papers; see, in particular, Kokoszka and Leipus (1998, 1999, 2000) and Lavielle and Moulines (2000). So far only limited simulation and empirical evidence is reported about these tests. We enlarge the scope of applicability by suggesting several improvements that enhance the practical implementation of the proposed tests. This paper focuses on the Kokoszka and Leipus (2000) and Lavielle and Moulines (2000) tests and proposes three types of extensions. First, we find via simulations that the VARHAC estimator proposed by den Haan and Levin (1997) yields good properties for the CUSUM-type estimator of Kokoszka and Leipus (2000). Simulation evidence is also presented for the application of this test to the multiple breaks setting using a sequential sample segmentation approach similar to that of Inclán and Tiao (1994). Second, the series used in the tests so far are either squared or absolute returns. We suggest the application of these tests to more precise measures of volatility, including the high frequency data-driven processes studied by Andersen *et al.* (2001), Andreou and Ghysels (2002), Barndorff-Nielsen and Shephard (2002), among others. Third, the finite sample performance of these new tests is assessed via extensive Monte Carlo simulations for realistic univariate GARCH models, single and multiple breaks as well as different algorithms and information criteria for the multiple breaks case.

The empirical application examines various financial series, including equity index returns for several financial markets in the Hong Kong, Japan, the UK and US as well as FX market series. Our empirical analysis is particularly complementary to Granger and Hyung (1999) who use the tests proposed by Inclán and Tiao (1994) and Bai (1997) to examine breaks in the absolute returns. The advantage of the Kokoszka and Leipus as well as Lavielle and Moulines tests is their validity under a wide class of strongly dependent processes, including long memory, GARCH-type and non-linear models. The Inclán and Tiao test applies in principle to independent series and is designed to find a break in the (unconditional) variance with unknown location. We show via Monte Carlo that the Inclán and Tiao test has nevertheless power and only minor size distortions when applied to strongly dependent data, though it is not as powerful as the Kokoszka and Leipus and Lavielle and Moulines tests.

The paper is organized as follows. In Section 2 we describe the various tests. Section 3 presents the Monte Carlo design and results. Section 4 contains the empirical application and a final section concludes.

2. TEST STATISTICS FOR BREAKS IN VOLATILITY DYNAMICS

A classical statistical problem is to test the homogeneity of a process or the parameter constancy of models. There is a substantial literature on this question known as a change-point problem. The task is to test if a change or structural break has occurred somewhere in a sample and, if

---

[2] The tests proposed by Andrews (1993) require α-mixing and can be potentially applied to ARCH type models. They are, however, designed for a single break.

so, to estimate the time of its occurrence. The simplest form of departure from stationarity is a change in mean at some (unknown) point in the sample. This problem has received a great deal of attention; see, for instance, Csörgo and Horváth (1997) for a literature review. Financial returns series typically have constant mean, but exhibit noticeable and complex clustering patterns in volatility (see e.g. Bollerslev *et al.*, 1994, for a survey of stylized facts). Such processes pose some non-trivial challenges as detecting a change in variance in an ARCH model can be rather difficult.[3] This section provides a brief discussion of the Kokoszka and Leipus (2000) as well as the Lavielle and Moulines (2000) tests for single and multiple breaks as well as the volatility series to which the tests can be applied to in order to test for change points in the second-order dynamics of a process.

2.1. CUSUM Type Tests

Let the asset returns process, r_t, be a strongly dependent, e.g. β-mixing process with finite fourth moment. A large class of ARCH and SV models are β-mixing (see, for instance, Carrasco and Chen, 2001) that satisfy these assumptions.[4] Define the process of interest $X_t = |r_t|^\delta$ for $\delta = 1, 2$ which represents an observed measure of the variability of returns. Given that the measurable functions of mixing processes are mixing and of the same size (see White, 1984, Theorem 3.49) then $X_t = G(r_t, \ldots, r_{t-\tau})$, for finite τ, is also β-mixing. The choice of δ is of course important. For $\delta = 2$ we look at squared returns which is the parent process parametrically modelled in ARCH or SV-type models. Alternatively, when $\delta = 1$, we examine absolute returns, which is considered as another measure of risk; see, for instance, the Power-ARCH models (Ding *et al.*, 1993). It is worth mentioning that long memory features have also been established in the absolute returns (e.g. Ding *et al.*, 1993; Granger and Ding, 1996). Although the tests analysed here apply to some long memory volatility models they are beyond the scope of this paper.[5] Without an explicit functional form for the second conditional moment, the tests discussed in this section will examine whether there is evidence of structural breaks in the dynamics of stock returns volatility. If we find a break, one must conclude that when fitting ARCH or SV-type processes, there will be instability in their parametric structure. We can take this reasoning a step further and think of sampling returns intra-daily, denoted $r_{(i),t}$ for some intra-day frequency $i = 1, \ldots, m$, and form data-driven estimates of daily volatility by taking sums of squared intra-day returns. This is an example of $X_t = G(r^2_{(1),t}, \ldots, r^2_{(m),t})$. The high frequency process is β-mixing, and so is the daily sampled sum of intra-day squared returns, or various other empirical measures of quadratic variation. Using the notation of Andreou and Ghysels (2002) $X_t = (QVi)_t$ which are locally smoothed filters of the quadratic variation using i days of high-frequency data. The case of $QV1$ corresponds to the filters studied by Andersen *et al.* (2001) and Barndorff-Nielsen and Shephard (2002). The details of the various specifications for the X_t process will be discussed in the last subsection.

---

[3] One could, for instance, think of extreme cases, where there is no change in the unconditional moments but only a perturbation in the conditional variance dynamics.
[4] Examples that form exceptions in this class are Integrated GARCH (IGARCH) or Fractionally IGARCH (FIGARCH) models which are not covariance stationary.
[5] Results on change-point tests for volatility models with long memory can be found in Andreou (2002).

In order to test for breaks in an ARCH(∞) Kokoszka and Leipus (1998, 2000) consider the following process:

$$U_T(k) = \left(1/\sqrt{T} \sum_{j=1}^{k} X_j - k/(T\sqrt{T}) \sum_{j=1}^{T} X_j \right) \qquad (1)$$

where $0 < k < T$, $X_j = r_t^2$. The returns process $\{r_t\}$ follows an ARCH(∞) process, $r_t = u_t\sqrt{h_t}$, $h_t = a + \sum_{j=1}^{\infty} b_j r_{t-j}^2$, $a \geq 0, b_j \geq 0$, with finite fourth moment and errors u_t that can be non-Gaussian. The CUSUM type estimator $\hat{k}$ of a change point k^* is defined as:

$$\hat{k} = \min\{k: |U_T(k)| = \max_{1 \leq j \leq T} |U_T(j)|\} \qquad (2)$$

The estimate $\hat{k}$ is the point at which there is maximal sample evidence for a break in the squared returns process. In the presence of a single break it is proved that $\hat{k}$ is a consistent estimator of the unknown change-point k^* with $P\{|k^* - \hat{k}| > \varepsilon\} \leq C/(\delta \varepsilon^2 \sqrt{n})$, where C is some positive constant and δ depends on the ARCH parameters and $|k^* - \hat{k}| = O_p(1/n)$ (Kokoszka and Leipus, 1998, 2000). Under the null hypothesis of no break:

$$U_T(k) \rightarrow_{D[0,1]} \sigma B(k) \qquad (3)$$

where $B(k)$ is a Brownian bridge and $\sigma^2 = \sum_{j=-\infty}^{\infty} \text{Cov}(X_j, X_0)$. Consequently, using an estimator $\hat{\sigma}$, one can establish that under the null:

$$\sup\{|U_T(k)|\}/\hat{\sigma} \rightarrow_{D[0,1]} \sup\{B(k): k\varepsilon[0, 1]\} \qquad (4)$$

which establishes a Kolmogorov–Smirnov type asymptotic distribution.[6]

The computation of the Kokoszka and Leipus (1998, 2000) test (henceforth K&L test) is relatively straightforward, with the exception of $\hat{\sigma}$ appearing in (4). The authors suggest to use a Heteroscedasticity and Autocorrelation Consistent (HAC) estimator applied to the X_j process. There are a number of such estimators, depending on the kernel function one uses (see, for instance, the chapter by den Haan and Levin, 1997 and references therein). We have experimented with a number of kernel estimators in addition to the procedure of den Haan and Levin (1997) who propose a HAC estimator without any kernel estimation, which is called the Vector Autoregression Heteroscedasticity and Autocorrelation Consistent (VARHAC) estimator. This estimator has an advantage over any estimator which involves kernel estimation in that the circular problem associated with estimating the optimal bandwidth parameter can be avoided. This estimator involves fitting a parametric autoregressive model and choosing the order of AR using, for instance, the AIC. The Monte Carlo evidence reported in den Haan and Levin (1997) indicates that the VARHAC estimator performs better than the non-prewhitened and prewhitened kernel estimators in many cases. Although we have not done a systematic study of various kernel HAC estimators versus the VARHAC estimator, we found via simulations that the latter is reliable. Hence all the results in the paper are based on the VARHAC estimator for $\hat{\sigma}$ appearing in (4).

The advantage of the K&L test is its validity under a wide class of processes, including long memory, GARCH-type and non-linear time series models. In a study closely related to ours Granger

---

[6] Critical values can be found in most textbooks on nonparametric methods. The 90%, 95% and 99% percentile (two-sided test) critical values are, respectively: 1.22, 1.36 and 1.63.

and Hyung (1999) use a different test, proposed by Inclán and Tiao (1994) for linear models with breaks such as those proposed by Chen and Tiao (1990) and Engle and Smith (1999). Aggarwal *et al.* (1999) also apply this test to GARCH models. The Inclán and Tiao test (henceforth I&T test) applies in principle to independent series and is designed to find a break in the (unconditional) variance with unknown location. The test statistic is defined as:

$$IT = \sqrt{T/2} \max_k |D_k| \tag{5}$$

where $D_k = \left[\left(\sum_{j=1}^k X_j / \sum_{j=1}^T X_j \right) - k/T \right]$. It is interesting to note that the asymptotic distribution of the statistic in (5) is the same as in (4), that is, the supremum of Brownian bridge and hence the same Kolmogorov-Smirnov type asymptotic distribution. In the Monte Carlo simulations we will examine how the Inclán and Tiao test performs in non-independent settings using transformations that yield independent processes and compare it to the Kokoszka and Leipus test.

The Kokoszka and Leipus test is also adapted for the multiple breaks hypothesis. The number of breaks is determined following a sequential sample segmentation approach similar to that of Inclán and Tiao (1994) and Bai (1997). The simulations present some encouraging results regarding the performance of the test for further theoretical investigation in the multiple breaks case. The test is applied to a few sample segments given the appropriate significance level adjustment.

2.2. Least Squares Type Tests

The change-point literature has recently dealt with the unknown multiple change points question in weakly dependent processes in a least-squares context. For instance, Bai (1994), Bai and Perron (1998) and Liu *et al.* (1997) use the Hájek–Rényi inequality to establish the asymptotic distribution of the test procedure. Recent work by Lavielle and Moulines (2000) has greatly increased the scope of testing for multiple breaks. They prove the Hájek–Rényi inequality results for weakly as well as strongly dependent processes, the latter being α-mixing which include long memory and ARCH-type processes.[7] The number of breaks is estimated via a penalized least-squares approach similar to Yao (1988). In particular, Lavielle and Moulines show that an appropriately modified version of the Schwarz criterion yields a consistent estimator of the number of change-points.

Consider the following generic model:

$$X_t = \mu_k^* + \varepsilon_t \quad t_{k-1}^* \le t \le t_k^* \quad 1 \le k \le r \tag{6}$$

where $t_0^* = 0$ and $t_{r+1}^* = T$, the sample size. The indices of the breakpoint and mean values $\mu_k^*, k = 1, \ldots, r$ are unknown. It is worth recalling that X_t is a generic process. In practical applications, equation (6) applies to squared returns, absolute returns, high-frequency data-driven volatility estimates, etc. The Lavielle and Moulines tests are based on the following least-squares computation:

$$Q_T(t) = \min_{\mu_k^*, k=1,\ldots,r} \sum_{k=1}^{r+1} \sum_{t=t_{k-1}+1}^{t_k} (X_k - \mu_k)^2 \tag{7}$$

Estimation of the number of break points involves the use of the Schwarz or Bayesian information criterion and hence a penalized criterion $Q_T(t) + \beta_T r$, where $\beta_T r$ is a penalty function to avoid

---

[7] Recall that the latter are β-mixing which imply α-mixing.

oversegmentation with r being the number of changes and $\{\beta_T\}$ a decreasing sequence of positive real numbers. We examine the properties of this test using both the Bayesian information criterion and that proposed in Liu *et al.* (1997). It is shown under mild conditions that the change-point estimator is strongly consistent with T rate of convergence. The Lavielle and Moulines simultaneously detects multiple breaks as opposed to the sequential adaptation of the Kokoszka and Leipus test. It is worth noting that the Kokoszka and Leipus statistic in (1) can be weighted by $(k(T-k)T^{-1})^\gamma$, $0 \le \gamma < 1$, as suggested in Kokoszka and Leipus (1998, equation 1.3, p. 386) which is then equivalent to the weighted least squares objective function when $\gamma = 0.5$. This brings out some similarities between the objective functions of the above two tests.

Given the asymptotic nature and sequential application of change-point tests in subsamples we may obviously end up with relatively small samples. It will therefore be important to appraise the power and size properties of these change-point tests in small samples via Monte Carlo simulations.

2.3. Empirical Processes

Following the above discussion the CUSUM and least squares tests are applied to the generic process X_t which represents the squared or absolute returns, both of which are regarded as alternative measures of risk. A departure from the limited number of applications found in the literature so far is to use estimates of conditional volatility which are based on high frequency data. The logic for considering such empirical processes is that squared returns can be viewed as noisy realizations of the underlying conditional volatility process (see Andersen and Bollerslev, 1998, for a discussion). Hence, instead of considering the daily return process and square it, we can take advantage of high frequency intra-daily data to obtain daily estimates of volatility.[8] Using the notation $r_{(m),t}$ to represent high frequency data on day t sampled with frequency m we can study sums of squared returns $r_{(m),t}^2$ for different values of m, to produce the daily volatility measure: (i) $\hat{\sigma}_t^{QV1} = \sum_{j=1}^m r_{(m),t+1-j/m}^2$, $t = 1, \ldots, T$, where for the 5-minute sampling frequency the lag length is $m = 288$ for financial markets open 24 hours per day (e.g. FX markets) as in Andersen *et al.* (2001), Andreou and Ghysels (2002) and Barndorff-Nielsen and Shephard (2002) or (ii) One-day Historical Quadratic Variation (introduced in Andreou and Ghysels, 2002) defined as the sum of m rolling QV estimates: $\hat{\sigma}_t^{HQV1} = 1/m \sum_{j=1}^m QV1_{(m),t+1-j/m}$, $t = 1, \ldots, T$. The intraday volatilities are denoted as QVi, $HQVi$ for window lengths $i = 1, 2, 3$. Clearly, the regularity conditions for squared daily returns can be transplanted to these more efficient filtering schemes like QVi (as discussed in Section 2.1).

3.   THE MONTE CARLO DESIGN AND RESULTS

The aim of this section is to evaluate the performance of the Kokoszka and Leipus (1998, 2000) in (4) as well as Inclán and Tiao (1994) tests in (5) (also referred to as K&L and I&T tests, respectively) in detecting breaks in the volatility dynamics of financial asset returns. The observed absolute or squared returns transformations are the series monitored for single and multiple breaks. The simulation design examines the size and power properties of the Kokoszka and Leipus and

---

[8] We refrain here from a discussion of the diffusion details of this class of estimators as well as definitions of quadratic variation. For details we refer the reader to Andersen *et al.* (2001), Andreou and Ghysels (2002) and Barndorff-Nielsen and Shephard (2000).

Lavielle and Moulines tests for GARCH type processes that can be considered as representative models of financial asset returns. Kokoszka and Leipus (2000) report simulation results that focus on the sampling distribution of the change-point estimator $\hat{k}$ for an ARCH(1) process. They find that its sampling distribution depends on the location of the change-point, the size of the variance change and its source. The extensive results presented in this section complement some early simulation evidence of these tests in establishing their power for univariate and bivariate GARCH processes and a number of alternative change-point hypotheses often encountered in asset returns. The robust character of the test is also examined in the presence of outliers given the stylized fact of jumps or extreme observations observed in volatility and absolute returns which may lead to spurious nonlinearities or IGARCH effects (e.g. Lamoureux and Lastrapes, 1990; van Dijk *et al.*, 1999).

The apparent similarity of the CUSUM-type statistics in K&L and I&T calls for an interesting comparison which brings about the connection between these two tests and their power in detecting change-points in GARCH processes as well as jumps in financial markets. For comparison purposes both tests in (4) and (5) are evaluated for absolute and squared returns whereas the I&T test is also applied to the residuals of a GARCH process given that this test is originally designed for independent processes. First we discuss the simulation design followed by an analysis of the results.

3.1. Simulation Design

The simulated returns processes are generated from a univariate Normal-GARCH process given by:

$$r_{i,t} = u_{i,t}\sqrt{h_{i,t}}$$
$$h_{i,t} = \omega_i + \alpha_i u_{i,t-1}^2 + \beta_i h_{i,t-1}, \quad t = 1, \dots, T \quad \text{and} \quad i = 0, 1 \qquad (8)$$

where $r_{i,t}$ is the returns process generated by the product of $u_{i,t}$ which is *i.i.d.*(0,1) and the volatility $h_{i,t}$ that has a GARCH(1,1) specification. This process without change points is denoted by $i = 0$ whereas a break in any of the parameters of the process is symbolized by $i = 1$. The models used in the simulation study are representative of financial markets data with the following set of parameters that capture a range of degrees of volatility persistence (measured by $\alpha_0 + \beta_0$). The vector parameters $(\omega_0, \alpha_0, \beta_0)$ in (8) describes the following Data Generating Processes: DGP1: (0.4,0.1,0.5) and DGP2: (0.1,0.1,0.8) which are characterized by low and high volatility persistence, respectively. The sample sizes of $T = 500, 1000, 3000$ are chosen so as to examine not only the asymptotic behaviour but also the small-sample properties of the tests. The small-sample features are particularly relevant for the sequential application of the tests in subsamples.

The model without breaks ($i = 0$) denotes the processes under the null hypothesis. Under the alternative hypothesis the returns process is assumed to exhibit breaks and four hypotheses are considered to evaluate the power of the tests. The simulation study first examines the single change-point followed by the multiple breaks hypothesis. In the context of (2.1) we study a single break in the conditional variance h_t which can also be thought as a permanent regime shift in volatility at change points πT ($\pi = .3, .5, .7$). Such breaks may have the following sources. H_1^A: a change in the volatility dynamics (or persistence), β_i. H_1^B: a change in the intercept, ω_i. H_1^C: a change in the tails of $u_{0,t}$ to $u_{1,t} \sim N(0, \sigma_u)$, ($\sigma_u = 1.1, 1.5$) at $t = \pi T + 1, \dots, T$. H_1^D: outliers in

the error, $u_{0,t}$ to $u_{1,t} \sim N(\mu_u, 1)$, with jump sizes $\mu_u = 4, 5$ and frequencies at given regular dates of a daily sample, $\Delta \cdot t_j$, (where $\Delta = 250, 500$ and $t_j = 1, 2, \ldots, \Delta/T$) and zero otherwise.[9] The sample sizes are $T = 1000$ and 3000 observations to match the empirical analysis as well as the large samples encountered in financial asset returns series of relatively high sampling frequency.

For the multiple breaks case we examine the alternative hypothesis that there are two breaks (or equivalently three segments) where the change points occur in model (8) as follows:

$$r_{1,t} = u_{1,t}\sqrt{h_{1,t}}, h_{1,t} = \omega_1 + \alpha_1 u_{1,t-1}^2 + \beta_1 h_{1,t-1} \quad \text{if } 1 < t \le [\pi T]$$

$$r_{2,t} = u_{2,t}\sqrt{h_{2,t}}, h_{2,t} = \omega_2 + \alpha_2 u_{2,t-1}^2 + \beta_2 h_{2,t-1} \quad \text{if } [\pi T] < t \le [2\pi T]$$

$$r_{3,t} = u_{3,t}\sqrt{h_{3,t}}, h_{3,t} = \omega_3 + \alpha_3 u_{3,t-1}^2 + \beta_3 h_{3,t-1} \quad \text{if } [2\pi] < t \le T \quad (9)$$

where $\pi = 0.33$ and $u_{j,t} \sim N(0, \sigma_{u_j})$, $j = 1, 2, 3$ and uncorrelated.[10] Under the null hypothesis the simulated process (8) holds whereas under the alternative hypothesis of multiple change points the process (9) is simulated with different sources of breaks as described under H_1^A, H_1^B, $H_1^C H_1^D$ above. In order to detect the two breaks at 0.33 (πT) and 0.66 of the sample ($2\pi T$) we first apply the Kokoszka and Leipus (2000) test to the total sample (T) of X_t and if a break is detected an algorithm similar to that of Inclán and Tiao (1994) is applied according to which the sample is segmented and the test is applied again to each subsample following sequential sample segmentation whenever a break is detected. The simulation design sequentially applies the test for upto 5 segments. Hence a 1% significance level is applied to each segment. The Lavielle and Moulines (2000) test is also applied to the multiple breaks process (as well as single break case mainly for comparison purposes with the K&L and I&T tests). This test simultaneously detects the number of change points and dates the breaks. The Lavielle and Moulines is a least squares type test and in order to detect multiple breaks one may apply a grid search approach or a more efficient algorithm based on dynamic programming suggested in Bellman and Roth (1969) and reintroduced in Bai and Perron (1998, 2002). Two information criteria are used for the penalty function: the BIC (Bayesian Information Criterion) and its modification by Liu *et al.* (1997) denoted LWZ. The simulation as well as empirical analysis is performed using the GAUSS programing language.

The simulation investigation is organized as follows: First, we consider the application of the I&T, K&L and L&M tests in model (8) to evaluate their size and power under the single break hypothesis. The alternative hypotheses aim to examine the power of the test in detecting breaks due to either changes in the parameters (defined by H_1^A, H_1^B) or the error of the GARCH process (defined by H_1^C, H_1^D). The former are interesting for studying the parameter constancy of the GARCH dynamics whereas the latter for examining the distributional homogeneity of the process. Both are the underlying assumptions in many asset pricing relationships and Value at Risk (VaR). Second, we evaluate the performance of the K&L and L&M tests for multiple breaks for the alternative hypotheses mentioned above.

---

[9] In our experiment the above simple jump process would facilitate the evaluation of the test's power in the presence of controlled outliers.
[10] The multiple change point model can be extended to more than two breaks and some preliminary simulation results with four breaks show that the tests share good properties, partly because of the large sample sizes encountered in financial asset returns. Therefore, for conciseness and comparison with other studies we report the two breaks case.

3.2. Simulation Results

The simulation results commence with the evaluation of the K&L test when the underlying process is a Normal-GARCH(1,1). Table I reports only minor size distortions for GARCH models with low persistence (e.g. DGP1 where $a_0 + \beta_0 = 0.6$). These minor distortions remain as the sample size increases from $T = 500$ to 3000. For the high-persistent process (DGP2) the test suffers more serious distortions up to 20%. The power of the K&L test is evaluated by a number of alternative hypotheses as defined in the previous section. The results in Table I suggest that the tests have good power in detecting breaks under the following alternative hypotheses: Break in the constant (H_1^B) or dynamics (H_1^A) of volatility. The power of the test is demonstrated even for small changes (e.g. a 0.1 increase) in β_0 for all DGPs. Similar results apply to the alternative of a small change in the error term (H_1^C). The power of the tests increases with T in all DGPs. Note that the high nominal power for the persistent GARCH process (DGP2) needs to be weighted by the size distortions for this process. In H_1^D the presence of outliers or short-lived jumps, which are evident in financial markets, do not seem to have an adverse effect on the test. The power of the K&L test is also evaluated for early change-points for H_1^A, H_1^B, H_1^C and the results show that the K&L test can detect breaks that occur as early as at $\pi = 0.3$ of the sample.

The size and power properties of the K&L test are compared with those of I&T. The latter is derived for independent series but has been applied to processes that exhibit dependence (Aggarwal *et al*., 1999; Granger and Hyung, 1999). Therefore we examine the properties of the test for $(r_t)^2$, $|r_t|$ as well as the errors of the GARCH process $(u_t)^2$ where $u_t = r_t/\sqrt{h_t}$ yields an independent series. Note that u_t refers to the simulated process and does not involve the estimation of the conditional variance. Table II presents the nominal size and power simulation results of the I&T test under the same null hypothesis of a Normal-GARCH(1,1) and all the alternative hypotheses discussed above. Let us first compare the performance of the I&T for $(r_t)^2$ and $|r_t|$. The I&T test for $(r_t)^2$ suffers from size distortions (above 10%) for all DGPs and sample sizes but appears to have good power in detecting even small changes in the GARCH coefficients or the error process (shown by the alternative hypotheses H_1^A, H_1^B, H_1^C) for large T. Nevertheless, its performance is adversely affected by outliers which appear to be consistently detected as change-points. If instead we adopt the $|r_t|$ transformation we note some interesting differences. The I&T test for $|r_t|$ appears seriously under-sized and with relatively less power, when compared with $(r_t)^2$ for any of the alternative hypotheses. However, it is interesting to note that for large T (e.g. 1000, 3000) and highly persistent GARCH processes (e.g. DGP2) the I&T test has good power properties and is not susceptible to outliers as opposed to $(r_t)^2$. Finally we examine the I&T test for $(u_t)^2$ which is by design an independent series. The size of the I&T statistic for $(u_t)^2$ is near the nominal 5% level. The I&T test for $(u_t)^2$ has power in detecting even small changes in the variance of error term (demonstrated by H_1^C) and is not seriously affected by outliers (H_1^D) for large samples, $T = 3000$. This evidence complements the results in Inclán and Tiao for *i.i.d.* series in that it shows that this test could be applied to the residuals of a GARCH for which it would have power to detect breaks only in the error term for very large samples. This statement is supported by the simulation results for the H_1^A and H_1^B for $(u_t)^2$ which show that it lacks power in detecting breaks in the conditional variance. The reason u_t lacks power is due to the standardization of the returns process $r_t/\sqrt{h_t}$ that offsets the corresponding changes in r_t and $\sqrt{h_t}$ and yields an *i.i.d.* error process, u_t.

In Table III we report the Lavielle and Moulines test simulation results for a single break. This table reports the frequency distribution of the number of change points—to make it comparable with the simulation results for multiple breaks that follow. The test appears to have good

Table I. Nominal size and power of the Kokoszka and Leipus (2000) test for a *single* change-point in the volatility based on a GARCH process

Samples, T Kokoszka & Leipus statistic Returns transformations	500 $U_{\max}/\widehat{\sigma}_{HAC}$		1000 $U_{\max}/\widehat{\sigma}_{HAC}$		3000 $U_{\max}/\widehat{\sigma}_{HAC}$	
	$(r_t)^2$	$|r_t|$	$(r_t)^2$	$|r_t|$	$(r_t)^2$	$|r_t|$
H_0: Univariate GARCH, $r_{0,t} = u_{0,t}\sqrt{h_{0,t}}$, $h_{0,t} = \omega_0 + \alpha_0 u_{0,t-1}^2 + \beta_0 h_{0,t-1}$, with $(\omega_0, \alpha_0, \beta_0)$:						
DGP1: (0.4, 0.1, 0.5)	0.059	0.072	0.061	0.078	0.061	0.067
DGP2: (0.1, 0.1, 0.8)	0.171	0.165	0.187	0.185	0.212	0.205
H_1^A: Break in the dynamics of volatility, β_0 (increase of 0.1) at $0.5T$.						
DGP1: $\beta_0 = 0.5$ to $\beta_1 = 0.6$	0.273	0.280	0.492	0.473	0.945	0.926
DGP2: $\beta_0 = 0.8$ to $\beta_1 = 0.9$	0.978	0.978	0.999	0.999	1.000	1.000
H_1^A: Break in the dynamics of volatility, β_0 (increase of 0.1) at $0.3T$.						
DGP1: $\beta_0 = 0.5$ to $\beta_1 = 0.6$	0.190	0.204	0.382	0.390	0.838	0.825
DGP2: $\beta_0 = 0.8$ to $\beta_1 = 0.9$	0.934	0.942	0.996	0.999	1.000	1.000
H_1^B: Break in the constant of volatility, ω_0 (increase of 0.1) at $0.5T$.						
DGP1: $\omega_0 = 0.4$ to $\omega_1 = 0.5$	0.210	0.204	0.353	0.353	0.809	0.787
DGP2: $\omega_0 = 0.1$ to $\omega_1 = 0.2$	0.718	0.702	0.913	0.915	1.000	1.000
H_1^B: Break in the constant of volatility, ω_0 (increase of 0.1) at $0.3T$.						
DGP1: $\omega_0 = 0.4$ to $\omega_1 = 0.5$	0.148	0.153	0.254	0.262	0.674	0.634
DGP2: $\omega_0 = 0.1$ to $\omega_1 = 0.2$	0.552	0.573	0.851	0.844	0.999	0.999
H_1^C: Break in the error, $u_0 \sim N(0, 1)$ (increase $\sigma_{u_1} = 1.1$) at $0.5T$						
DGP1: $u_1 \sim N(0, 1.1)$	0.287	0.277	0.548	0.520	0.921	0.917
DGP2: $u_1 \sim N(0, 1.1)$	0.449	0.437	0.710	0.700	0.982	0.975
H_1^C: Break in the error, $u_0 \sim N(0, 1)$ (increase $\sigma_{u_1} = 1.1$) at $0.3T$						
DGP1: $u_1 \sim N(0, 1.1)$	0.195	0.199	0.329	0.333	0.833	0.804
DGP2: $u_1 \sim N(0, 1.1)$	0.376	0.386	0.548	0.548	0.932	0.923
H_1^D: Outliers in the error, $u_0 \sim N(0, 1)$ ($\mu_{u_1} = 5$ every 250 observations).						
DGP1: $u_1 \sim N(5, 1)$	0.019	0.046	0.015	0.039	0.005	0.044
DGP2: $u_1 \sim N(5, 1)$	0.039	0.115	0.046	0.134	0.062	0.145

Notes: The Kokoszka and Leipus (2000) test statistic is defined as

$$U_k = \left(\frac{1}{\sqrt{T}} \sum_{j=1}^{k} r_j^2 - \frac{k}{T} \frac{1}{\sqrt{T}} \sum_{j=1}^{T} r_j^2 \right)$$

The max $U_T(k)$ is standardized by the VARHAC estimator, $\widehat{\sigma}_{HAC}$, which is applied to X_t that represents either squared or absolute returns of the GARCH model. The normalized statistic $U_{\max}/\widehat{\sigma}_{HAC}$ converges to the sup of a Brownian Bridge with asymptotic critical value 1.36 at the 5% significance level. The Normal GARCH (1,1) model is simulated (1000 replications) where the superscripts 1 and 0 in the variables and coefficients in the Table denote the cases with and without change-points, respectively.

size properties for either information criterion (BIC or LWZ) and for both DGPs and return transformations. For the persistent GARCH (DGP2) the LWZ criterion performs better than the BIC in terms of size. The test appears to have good power overall for a single break except when there is a small change in the coefficients of either the constant or dynamics of the GARCH model. This result is evident in the low persistent GARCH and in particular when the LWZ is used. However, the L&M test with the LWZ criterion seems to have relatively more power in detecting changes in the GARCH error than the BIC. Both criteria appear robust to the outliers. Comparing the L&M and K&L simulation results we observe that the latter performs relatively better when the size of change is small (e.g. 0.1 increase in the parameters of DGP1) but suffers from relatively higher size distortions in persistent GARCH processes.

Table II. Nominal Size and Power of the Inclán and Tiao (1994) test for a *single* change-point in the volatility based on a GARCH process

Samples, T Inclan & Tiao statistic Returns transformations	500 IT			1000 IT			3000 IT		
	$(r_t)^2$	$\|r_t\|$	$(u_t)^2$	$(r_t)^2$	$\|r_t\|$	$(u_t)^2$	$(r_t)^2$	$\|r_t\|$	$(u_t)^2$
H_0: Univariate GARCH, $r_{0,t} = u_{0,t}\sqrt{h_{0,t}}$, $h_{0,t} = \omega_0 + \alpha_0 u_{0,t-1}^2 + \beta_0 h_{0,t-1}$, with $(\omega_0, \alpha_0, \beta_0)$:									
DGP1: (0.4, 0.1, 0.5)	0.117	0.000	0.041	0.130	0.000	0.046	0.149	0.000	0.052
DGP2: (0.1, 0.1, 0.8)	0.280	0.001	0.044	0.301	0.005	0.046	0.373	0.003	0.052
H_1^A: Break in the dynamics of volatility, β_0 (increase of 0.1) at $0.5T$									
DGP1: $\beta_0 = 0.5$, $\beta_1 = 0.6$	0.415	0.009	0.040	0.622	0.027	0.045	0.961	0.296	0.052
DGP2: $\beta_0 = 0.8$, $\beta_1 = 0.9$	0.989	0.635	0.040	1.000	0.964	0.045	1.000	1.000	0.052
H_1^B: Break in the constant of volatility, ω_0 (increase of 0.1) at $0.5T$									
DGP1: $\omega_0 = 0.4$, $\omega_1 = 0.5$	0.474	0.008	0.040	0.493	0.008	0.045	0.875	0.130	0.052
DGP2: $\omega_0 = 0.1$, $\omega_1 = 0.2$	0.807	0.125	0.040	0.962	0.413	0.045	1.000	0.975	0.052
H_1^C: Break in the error, $u_0 \sim N(0, 1)$ (increase $\sigma_{u_1} = 1.1$) at $0.5T$									
DGP1: $u_1 \sim N(0, 1.1)$	0.442	0.009	0.242	0.663	0.029	0.451	0.969	0.331	0.914
DGP2: $u_1 \sim N(0, 1.1)$	0.675	0.054	0.242	0.817	0.158	0.451	0.989	0.675	0.914
H_1^D: Outliers in the error, $u_0 \sim N(0, 1)$ ($\mu_{u_1} = 5$ every 250 observations).									
DGP1: $u_1 \sim N(5, 1)$	0.365	0.000	0.232	0.275	0.000	0.131	0.227	0.000	0.083
DGP2: $u_1 \sim N(5, 1)$	0.526	0.006	0.232	0.481	0.005	0.131	0.450	0.002	0.083

Notes: The Inclán and Tiao (1994) statistic $IT = \sqrt{T/2} \max |D_k|$, where

$$D_k = \left[\left(\sum_{j=1}^{k} r_j^2 \Big/ \sum_{j=1}^{T} r_j^2 \right) - \frac{k}{T} \right]$$

is specified for independent processes. It also converges to the sup of a Brownian Bridge with asymptotic critical value 1.36 at the 5% significance level. The process u_t is the 'true' simulated process that does not involve volatility estimation. In the above table the seed of the random number generator is controlled for comparison purposes. We note that when the seed is not controlled there may be a variation in the simulation results that only affects the third decimal place of the reported results.

The remaining simulation analysis addresses the multiple breaks hypotheses given in model (9). Table IV reports the K&L results using a sequential sample segmentation approach. The frequency distribution of the number of breaks under the alternative hypotheses is reported. The results show that the K&L has good power only for large and non-monotone (rather than small and gradual) changes in the GARCH parameters for any of the DGPs but for the absolute rather than the squared returns transformations. Similarly it shares good power for detecting changes in the variance of the error term in the GARCH process. As the sample size (T) increases the performance of the test improves even for small change points. The Lavielle and Moulines multiple breaks results are reported in Table V. The frequency distribution of the number of change points is reported for the alternative hypotheses, the two information criteria and two breaks. The results show that the L&M has good power in detecting the two change points especially when the BIC is applied to r_t^2 for small and gradual changes and when applied to either return transformation for large changes. In contrast, the LWZ underestimates the number of breaks in the GARCH parameters. Generally the L&M test has less power in detecting small and/or monotone (gradual) changes in the GARCH parameters as opposed to large and non-monotone changes. This is a common feature shared by the L&M and K&L tests. It is interesting to note that in detecting changes in the distribution of the GARCH process the absolute rather than squared returns transformation yields

Table III. Size, power and frequency distribution of the number of change-points obtained with the Lavielle and Moulines (2000) test when there is a *single* break at $0.5T$ of the sample in the GARCH process

Samples, T Lavielle & Moulines Segments, $t_k = 5$ Returns Number of breaks	1000						1000					
	BIC			LWZ			BIC			LWZ		
	$(r_t)^2$						$\|r_t\|$					
	0	1	≥ 2	0	1	≥ 2	0	1	≥ 2	0	1	≥ 2
H_0: Univariate GARCH, $r_{0,t} = u_{0,t}\sqrt{h_{0,t}}$, $h_{0,t} = \omega_0 + \alpha_0 u_{0,t-1}^2 + \beta_0 h_{0,t-1}$, with $(\omega_0, \alpha_0, \beta_0)$:												
DGP1: (0.4, 0.1, 0.5)	**0.96**	0.03	0.01	**1.00**	0.00	0.00	0.98	0.02	0.00	1.00	0.00	0.00
DGP2: (0.1, 0.1, 0.8)	**0.88**	0.07	0.05	**1.00**	0.00	0.00	0.93	0.07	0.00	1.00	0.00	0.00
H_1^A: Break in the dynamics of volatility with parameters (β_0, β_1)												
DGP1: (0.5,0.6)	0.72	**0.24**	0.04	1.00	**0.00**	0.00	0.79	**0.20**	0.01	1.00	**0.00**	0.00
DGP1: (0.5,0.8)	0.00	**0.95**	0.05	0.00	**1.00**	0.00	0.00	**0.93**	0.07	0.00	**1.00**	0.00
DGP2: (0.8,0.7)	0.21	**0.75**	0.03	0.85	**0.15**	0.00	0.20	**0.75**	0.05	0.84	**0.16**	0.00
DGP2: (0.8,0.4)	0.00	**0.72**	0.28	0.00	**1.00**	0.00	0.00	**0.86**	0.14	0.00	**1.00**	0.00
H_1^B: Break in the constant of volatility with parameters (ω_0, ω_1)												
DGP1: (0.4, 0.5)	0.85	**0.14**	0.01	1.00	**0.00**	0.00	0.82	**0.18**	0.00	1.00	**0.00**	0.00
DGP1: (0.4, 0.8)	0.00	**0.94**	0.06	0.38	**0.62**	0.00	0.00	**1.00**	0.00	0.36	**0.64**	0.00
DGP2: (0.1, 0.3)	0.00	**0.94**	0.06	0.18	**0.82**	0.00	0.00	**0.99**	0.01	0.13	**0.87**	0.00
DGP2: (0.1, 0.5)	0.00	**0.86**	0.14	0.00	**1.00**	0.00	0.00	**0.95**	0.05	0.00	**1.00**	0.00
H_1^C: Break in the variance of the error with parameters $(\sigma_{u_0}, \sigma_{u_1})$												
DGP1: (0, 1.1)	0.01	**0.49**	0.50	0.01	**0.94**	0.05	0.01	**0.57**	0.42	0.01	**0.95**	0.04
DGP1: (0, 1.5)	0.00	**0.63**	0.37	0.00	**0.97**	0.03	0.00	**0.58**	0.42	0.00	**0.97**	0.03
DGP1: (0, 3)	0.00	**0.60**	0.40	0.00	**0.98**	0.02	0.00	**0.53**	0.47	0.00	**0.93**	0.07
H_1^D: Outliers in the error, $u_0 \sim N(0, 1)$ ($\mu_{u_1} = 5$ every 250 observations).												
DGP1: $u_1 \sim N(5, 1)$	**0.99**	0.01	0.00	**1.00**	0.00	0.00	**0.99**	0.01	0.00	**1.00**	0.00	0.00
DGP2: $u_1 \sim N(5, 1)$	**0.98**	0.02	0.00	**1.00**	0.00	0.00	**0.92**	0.06	0.02	**1.00**	0.00	0.00

Notes: The Lavielle and Moulines (2000) test is described in Section 2.2. The Bayesian Information Criterion (BIC) and its modification by Liu *et al.* (1997) denoted as LWZ are used. The simulations focus on DGP1, DGP2, $T = 1000$ for 500 trials. For comparison purposes the alternative hypotheses of change points are similar to the K&L simulations (Table I) and extended to larger breaks. Reported is the frequency distribution of the breaks detected. The bold numbers refer to the true number of change-points in the simulated process.

relatively higher power with the BIC. Overall increasing the sample size, T, and the number of segments, t_k, improves the power of the test especially when the size of change is small. For large breaks the L&M and K&L share similar power. However for small or monotone changes the K&L has relatively more power and is computationally less demanding than the L&M test. The latter however has better size properties and does not overpredict the number of breaks.

4. EMPIRICAL RESULTS

There is a plethora of empirical evidence that squared asset returns exhibit dynamic heteroscedasticity (e.g. Bollerslev *et al.*, 1994) and absolute returns feature long-range dependence (e.g. Granger and Ding, 1996). Empirical studies recognize that the existence of breaks or regime changes in financial markets affects volatility and long-range dependence in stock returns (e.g. Lamoureux and Lastrapes, 1990; Mikosch and Starica, 1999; Granger and Hyung, 1999; Diebold and Inoue, 2001).

Table IV. Size, power and frequency distribution of the number of change-points obtained with the Kokoszka and Leipus (2000) test when there are *two* breaks at $0.33T$ and $0.67T$ of the sample in the GARCH process

Samples, T Returns Number of breaks	1000								3000							
	$(r_t)^2$				$|r_t|$				$(r_t)^2$				$|r_t|$			
	0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3	0	1	2	≥ 3
H_1^A: Break in the dynamics of volatility with parameters $(\beta_0, \beta_1, \beta_2)$																
DGP1:																
(0.5,0.6,0.8)	0.00	0.88	**0.01**	0.11	0.00	0.00	**0.87**	0.13	0.00	0.87	**0.13**	0.00	0.00	0.00	**1.00**	0.00
(0.5,0.6,0.3)	0.00	0.60	**0.39**	0.01	0.00	0.00	**0.99**	0.01	0.00	0.00	**0.99**	0.01	0.00	0.00	**1.00**	0.00
DGP2:																
(0.8,0.5,0.8)	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00
H_1^B: Break in the constant of volatility with parameters $(\omega_0, \omega_1, \omega_2)$																
DGP1:																
(0.4,0.5,0.8)	0.00	0.39	**0.61**	0.00	0.00	0.00	0.71	0.29	0.00	0.00	1.00	0.00	0.00	0.00	1.00	0.00
(0.4,0.8,0.4)	0.00	0.53	**0.39**	0.08	0.00	0.00	0.77	0.23	0.00	0.30	0.66	0.04	0.00	0.00	0.99	0.01
DGP2:																
(0.1,0.2,0.5)	0.00	0.01	**0.99**	0.00	0.00	0.00	**0.99**	0.01	0.00	0.00	**0.99**	0.01	0.00	0.00	**1.00**	0.00
(0.1,0.5,0.8)	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00
(0.1,0.5,0.1)	0.00	0.06	**0.94**	0.00	0.00	0.00	**0.98**	0.02	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00
(0.1,0.3,0.1)	0.00	0.20	**0.78**	0.02	0.00	0.00	**0.97**	0.03	0.00	0.00	**0.99**	0.01	0.00	0.00	**0.99**	0.00
H_1^C: Break in the variance of the error with parameters $(\sigma_{u_0}, \sigma_{u_1}, \sigma_{u_2})$																
DGP1:																
(0,1.5,3)	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00
(0,3,5)	0.00	1.00	**0.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00

Notes: The K&L test (see notes in Table I) is applied following a sequential sample segmentation approach and the frequency distribution of the change-points is reported. The bold numbers refer to the true number of breaks in the simulated process.

The empirical analysis aims to complement the simulation evidence in the following directions. We examine the change-point hypothesis in volatility dynamics of international stock market indices and FX returns. The empirical performance of the tests, discussed in the previous sections, is evaluated by examining the relation of the change-points to economic events detected not only in the squared and absolute returns but also to a family of data-driven volatility filters. Moreover, we estimate the volatility in subsamples prior and post breaks in an attempt to verify changes in the dependence of the series. The empirical analysis also complements the simulation results to tests for multiple breaks using the Lavielle and Moulines least-squares test as well as the Kokoszka and Leipus sequential sample segmentation test (henceforth K&L and L&M, respectively).

The empirical analysis is performed using data from the stock and FX markets. The four international stock market returns indices, the Financial Times Stock Exchange 100 index (FTSE), the Hang-Seng Index (HSI), the Nikkei 500 index (NIKKEI) and the Standard and Poors 500 index (S&P500) are studied over the period 4/1/1989–19/10/2001 at daily frequency (sample size, $T = 3338$). The data source is Datastream. The choice of the sample is based on the recent experience of the Asian and Russian financial crises. We also study the Yen *vis-à-vis* the US dollar returns over the period 1/12/1986–30/11/1996 at 5-minute sampling frequency. The data source is Olsen and Associates. The original sample is 1,052,064 five-minute return observations (2653 days·288 five-minute intervals per day). The returns for some days were removed from the sample to avoid having regular and predictable market closures which affect the characterization

Table V. Size, power and frequency distribution of the number of change-points obtained with the Lavielle and Moulines (2000) test when there are *two* breaks at $0.33T$ and $0.67T$ of the sample in the GARCH process

Samples, T Lavielle & Moulines Returns Number of breaks Segments		1000							1000							
		BIC				LWZ				BIC				LWZ		
		$(r_t)^2$								$|r_t|$						
		0	1	2	≥3	0	1	2	≥3	0	1	2	≥3	0	1	2	≥3

H_1^A: Break in the dynamics of volatility with parameters $(\beta_0, \beta_1, \beta_2)$

		0	1	2	≥3	0	1	2	≥3	0	1	2	≥3	0	1	2	≥3
DGP1:																	
(0.5,0.6,0.8)	5	0.00	0.95	**0.05**	0.00	0.01	0.99	**0.00**	0.00	0.00	0.98	**0.02**	0.00	0.06	0.94	**0.00**	0.00
	3	0.00	1.00	**0.00**	0.00	0.01	0.99	**0.00**	0.00	0.00	0.97	**0.03**	0.00	0.02	0.98	**0.00**	0.00
(0.5,0.6,0.3)	5	0.14	0.47	**0.39**	0.00	0.94	0.04	**0.02**	0.00	0.16	0.56	**0.28**	0.00	0.93	0.07	**0.00**	0.00
	3	0.20	0.50	**0.28**	0.00	0.97	0.03	**0.00**	0.00	0.19	0.62	**0.19**	0.00	0.96	0.04	**0.00**	0.00
DGP2:																	
(0.8,0.5,0.8)	5	0.00	0.03	**0.90**	0.06	0.70	0.19	**0.11**	0.00	0.01	0.00	**0.99**	0.36	0.06	0.58	**0.00**	0.00
	3	0.03	0.97	**0.00**	0.00	0.68	0.32	**0.00**	0.00	0.00	0.01	**0.99**	0.00	0.51	0.08	**0.41**	0.00

H_1^B: Break in the constant of volatility with parameters $(\omega_0, \omega_1, \omega_2)$

		0	1	2	≥3	0	1	2	≥3	0	1	2	≥3	0	1	2	≥3
DGP1:																	
(0.4,0.5,0.8)	5	0.05	0.91	**0.04**	0.00	0.66	0.34	**0.00**	0.00	0.04	0.94	**0.02**	0.00	0.63	0.37	**0.00**	0.00
	3	0.02	0.97	**0.01**	0.00	0.52	0.48	**0.00**	0.00	0.05	0.94	**0.01**	0.00	0.69	0.31	**0.00**	0.00
(0.4,0.8,0.4)	5	0.09	0.00	**0.90**	0.01	0.92	0.00	**0.08**	0.00	0.09	0.02	**0.89**	0.00	0.92	0.00	**0.08**	0.00
	3	0.02	0.00	**0.98**	0.00	0.90	0.00	**0.10**	0.00	0.09	0.01	**0.90**	0.00	0.97	0.00	**0.03**	0.00
DGP2:																	
(0.1,0.2,0.5)	5	0.00	0.82	**0.18**	0.00	0.01	0.99	**0.00**	0.00	0.00	0.67	**0.30**	0.03	0.00	1.00	**0.00**	0.00
	3	0.00	0.91	**0.09**	0.00	0.00	1.00	**0.00**	0.00	0.00	0.72	**0.28**	0.00	0.00	1.00	**0.00**	0.00
(0.1,0.5,0.8)	5	0.00	0.18	**0.79**	0.03	0.00	0.99	**0.01**	0.00	0.00	0.36	**0.60**	0.04	0.00	0.96	**0.04**	0.00
	3	0.00	0.20	**0.80**	0.00	0.00	1.00	**0.00**	0.00	0.00	0.34	**0.66**	0.00	0.00	0.91	**0.09**	0.00
(0.1,0.5,0.1)	5	0.00	0.00	**0.95**	0.05	0.01	0.00	**0.99**	0.00	0.00	0.00	**0.94**	0.06	0.00	0.00	**1.00**	0.00
	3	0.00	0.00	**1.00**	0.00	0.01	0.00	**0.99**	0.00	0.00	0.00	**1.00**	0.00	0.02	0.00	**0.98**	0.00
(0.1,0.3,0.1)	5	0.00	0.00	**0.99**	0.01	0.77	0.00	**0.23**	0.00	0.02	0.02	**0.93**	0.03	0.75	0.02	**0.23**	0.00
	3	0.01	0.02	**0.97**	0.00	0.68	0.00	**0.32**	0.00	0.04	0.00	**0.96**	0.00	0.71	0.02	**0.27**	0.00

H_1^C: Break in the variance of the error with parameters $(\sigma_{u_0}, \sigma_{u_1}, \sigma_{u2})$

		0	1	2	≥3	0	1	2	≥3	0	1	2	≥3	0	1	2	≥3
DGP1:																	
(0,1.5,3)	5	0.00	0.78	**0.22**	0.00	0.00	0.98	**0.02**	0.00	0.00	0.37	**0.56**	0.07	0.00	0.96	**0.04**	0.00
	3	0.00	0.97	**0.03**	0.00	0.00	1.00	**0.00**	0.00	0.00	0.53	**0.47**	0.00	0.00	1.00	**0.00**	0.00
(0,3,5)	5	0.00	0.76	**0.24**	0.00	0.00	0.96	**0.04**	0.00	0.00	0.00	**0.81**	0.19	0.00	0.00	**0.98**	0.02
	3	0.00	0.96	**0.04**	0.00	0.00	1.00	**0.00**	0.00	0.00	0.00	**1.00**	0.00	0.00	0.00	**1.00**	0.00

of the volatility dynamics. For the description of the data removed refer to Andersen *et al.* (2001). The final sample includes 705,024 five-minute returns reflecting $T = 2448$ trading days.

The empirical analysis commences with investigating the hypothesis of a single break in the four international stock market indices. The results in Table VI provide evidence that neither the K&L nor the I&T tests support the null hypothesis of homogeneity in the absolute or squared returns of the stock market indices over the sample 1989–2001. These results hold for two alternative non-parametric estimators of $(r_t)^2$ and $|r_t|$ used for standardizing the max $U_T(k)$ statistic defined in Section (1.1): the VARHAC estimator and the Nonlinear Least Squares (NLS) variance estimator of the ARMA(1,1) for squared and absolute returns (Francq and Zakoian, 2000b). The overall picture of the four stock market returns indices dates the change point in 1997 and in particular in the summer months of 1997 for the FTSE, HKI and NIKKEI. The same change-point dates are

Table V. (*Continued*)

Samples, T Lavielle & Moulines Returns Number of breaks Segments	t_k	3000								3000							
		BIC				LWZ				BIC				LWZ			
		$(r_t)^2$								$\lvert r_t \rvert$							
		0	1	2	≥3	0	1	2	≥3	0	1	2	≥3	0	1	2	≥3
H_1^A: Break in the dynamics of volatility with parameters $(\beta_0, \beta_1, \beta_2)$																	
DGP1:																	
(0.5,0.6,0.8)	5	0.00	0.92	**0.08**	0.00	0.00	1.00	**0.00**	0.00	0.00	0.79	**0.21**	0.00	0.00	1.00	**0.00**	0.00
(0.5,0.6,0.3)	5	0.00	0.91	**0.09**	0.00	0.00	1.00	**0.00**	0.00	0.00	0.77	**0.22**	0.01	0.00	1.00	**0.00**	0.00
DGP2:																	
(0.8,0.5,0.8)	5	0.00	0.00	**0.96**	0.04	0.00	0.00	**1.00**	0.00	0.00	0.00	**0.94**	0.06	0.00	0.00	**1.00**	0.00
H_1^B: Break in the constant of volatility with parameters $(\omega_0, \omega_1, \omega_2)$																	
DGP1:																	
(0.4,0.5,0.8)	5	0.00	0.90	**0.10**	0.00	0.01	0.99	**0.00**	0.00	0.00	0.92	**0.08**	0.00	0.03	0.97	**0.00**	0.00
(0.4,0.8,0.4)	5	0.00	0.00	**0.99**	0.01	0.09	0.00	**0.91**	0.00	0.00	0.09	**0.91**	0.00	0.00	0.98	**0.02**	0.00
DGP2:																	
(0.1,0.2,0.5)	5	0.00	0.45	**0.54**	0.01	0.00	1.00	**0.00**	0.00	0.00	0.16	**0.82**	0.02	0.00	0.98	**0.02**	0.00
(0.1,0.5,0.8)	5	0.00	0.01	**0.99**	0.00	0.00	0.35	**0.65**	0.00	0.00	0.01	**0.94**	0.05	0.00	0.39	**0.61**	0.00
(0.1,0.5,0.1)	5	0.00	0.00	**0.99**	0.01	0.01	0.00	**0.99**	0.00	0.00	0.00	**0.96**	0.04	0.00	0.00	**1.00**	0.00
(0.1,0.3,0.1)	5	0.00	0.00	**0.99**	0.01	0.01	0.00	**0.99**	0.00	0.00	0.00	**0.97**	0.03	0.01	0.00	**0.99**	0.00
H_1^C: Break in the variance of the error with parameters $(\sigma_{u_0}, \sigma_{u_1}, \sigma_{u2})$																	
DGP1:																	
(0,1.5,3)	5	0.00	0.74	**0.26**	0.00	0.00	0.98	**0.02**	0.00	0.00	0.00	**0.75**	0.25	0.00	0.87	**0.13**	0.00
(0,3,5)	5	0.00	0.64	**0.34**	0.02	0.00	0.99	**0.01**	0.00	0.00	0.00	**0.83**	0.17	0.00	0.00	**0.99**	0.00

Notes: Refer to the notes in Table III.

also supported by the I&T test. Using the simulation evidence in Table II we note that for large sample sizes T the I&T test for $\lvert r_t \rvert$ is well-behaved in terms of size and power and is not distorted by outliers. It is interesting to note that the extension of the I&T statistic by Kim *et al.* (2000) (also reported in Table VI as $B_T(C)$) does not detect any change-points. One possible explanation can be the poor power performance of the test in the presence of highly persistent GARCH processes as documented in Kim *et al.* and as is supported by the estimation of GARCH models for the four stock market indices (presented in the last table).

The change-points detected in the three international stock market indices in Table VI refer to the Asian crisis period. However, the single change point hypothesis can mask the existence of multiple breaks which implies that in dating change-points it is advisable to follow a multiple breaks procedure. The results of two tests are summarized in Table VII. In the Lavielle and Moulines test we adopt two penalty function criteria, the first is the Bayesian Information Set (BIC) and the second is a modified BIC as proposed in Liu *et al.* (1997) (denoted by LWZ in Table VII) and we set the number of segments t_k equal to 3 and 5. The empirical findings show that irrespective of the choice of t_k the L&M test consistently detects the same number of breaks. Specifically the combination of the BIC and $\lvert r_t \rvert$ tends to predict the largest number of breaks whereas the pair of LWZ and $(r_t)^2$ the smallest number of change-points. The latter result is consistent with the conservatism of the LWZ found in the simulation analysis. The Asian crisis period appears to be a common break in the above combinations (of processes and information criteria) and in all stock market indices that is revealed in different months of 1997. In July and August 1997 we detect the first change-points associated with the Asian crisis in the FTSE, HSI

Table VI. Testing for a *single* change-point in the volatility of daily stock market Indices (SMI) over the period 1989-2001

SMI	Returns	Change-point $k*$	Change-point Statistics				
			Kokoszka & Leipus test			Inclan & Tiao tests	
			$\max U_{k*}$	$\dfrac{\max U_{k*}}{\widehat{\sigma}_{HAC}}$	$\dfrac{\max U_{k*}}{\widehat{\sigma}_{ARMA}}$	IT	$B_T(C)$
FTSE	$\lvert r_t \rvert$	05/06/97	1.917	5.862*	6.665*	4.414*	0.599
	$(r_t)^2$	04/08/97	2.238	5.266*	3.511*	9.195*	1.249
HSI	$\lvert r_t \rvert$	14/08/97	3.460	4.619*	5.828*	4.954*	0.321
	$(r_t)^2$	18/08/97	7.104	2.181*	1.291	8.583*	0.556
NIKKEI	$\lvert r_t \rvert$	31/07/97	1.521	3.091*	3.806*	2.905*	0.449
	$(r_t)^2$	21/10/97	1.836	1.972*	1.305	4.427*	0.684
S&P500	$\lvert r_t \rvert$	04/02/97	2.395	6.882*	7.181*	5.837*	0.356
	$(r_t)^2$	26/03/97	2.718	4.888*	1.665*	11.103*	0.678

Notes:(1) The Stock Market Index (SMI) series refer to the Financial Times Stock Exchange index 100 (FTSE100), the Hang-Seng Index (HSI), the Nikkei 500 (NIKKEI), the Standards and Poors 500 index (S&P500). The daily sample over the period 4/1/1989 to 19/10/2001 yields $T = 3338$ observations. The series $r_t := \log p_t - \log p_{t-1}$ represents the returns on each index. The change-point tests are applied to the $(r_t)^2$ and $\lvert r_t \rvert$ transformations as well as $(u_t)^2$ where u_t is the residual from the GARCH. (2) The Kokoszka and Leipus (1998, 2000) reported $\max U_{k*}$ is the maximum of the statistic,

$$U_k = \left(\frac{1}{\sqrt{T}} \sum_{j=1}^{k} r_j^2 - \frac{k}{T} \frac{1}{\sqrt{T}} \sum_{j=1}^{T} r_j^2 \right)$$

which is standardized by non-parametric estimators, Heteroscedastic Consistent ($\widehat{\sigma}_{HAC}$) and ARMA estimators ($\widehat{\sigma}_{ARMA}$) of squared and absolute returns. The normalized statistic $\max U_{k*} / \widehat{\sigma}$ converges to the sup of a Brownian Bridge. (3) The Inclán and Tiao (1994) statistic

$$D_k = \left(\sum_{j=1}^{k} r_j^2 / \sum_{j=1}^{T} r_j^2 \right) - \frac{k}{T}$$

specified for *i.i.d.* processes normalized as $IT = \sqrt{T/2} \max \lvert D_k \rvert$ also converges to the sup of a Brownian Bridge and is extended in Kim *et al.* (2000) for GARCH processes to be $B_T(C) = C\sqrt{T} \max \lvert D_k \rvert$ where C^2 and κ are constants that are estimated by substituting the quasi-MLEs of the GARCH(1,1) $\widehat{\omega}$, $\widehat{\alpha}$, $\widehat{\beta}$ and $T^{-1} \sum_{j=1}^{T} r_j^4$ to ω, α, β and $E(r_j^4)$. (4) $k*$ refers to the location of the break and the * symbol attached to statistics denotes that the null hypothesis of no structural change is rejected using the asymptotic critical value of 1.36.

and NIKKEI followed by the October 1997 change-point in the S&P500 as well as the NIKKEI.[11] A second common break in the stock indices that is revealed in the L&M procedure is associated with the Russian crisis. In July 1998 we detect change-points in the FTSE and the S&P500 followed by the August 1998 break in the NIKKEI. Table VIII reports the results of multiple breaks from the Kokoszka and Leipus test sequential application. Comparing the results from the two tests we observe that the latter test detects a larger number of breaks especially when applied to the $\lvert r_t \rvert$ process even at the 1% significance level. The two multiple change-point tests detect some common breaks in the same year mainly that of 1997.

As a final empirical application we test for change-points in the FX market applying the K&L test to the family of high-frequency volatility filters that estimate the Quadratic Variation (QV)

---

[11] A detailed chronology of the Asian financial crisis events in 1997 and 1998 produced by N. Roubini can be found at http://www.stern.nyu.edu/nroubini/asia/AsiaChronology1.hml

Table VII. Testing for *multiple* change-points in the volatility of daily stock market Indices (SMI) over the period 1989-2001

SMI	Process	Selection criterion		Number & location of breaks	
			Lavielle and Moulines test		
FTSE	$|r_t|$	BIC	$-2.616(2), -2.610(1)$	2	3/11/92, 1/8/97
		LWZ	$-2.599(1), -2.549(0)$	1	1/8/97
	$(r_t)^2$	BIC	$-2.123(1), -2.070(0)$	1	10/7/98
		LWZ	$-2.112(1), -2.069(0)$	1	10/7/98
HSI	$|r_t|$	BIC	$-1.121(3), -1.117(2)$	3	3/7/92, 24/1/95, 15/8/97
		LWZ	$-1.108(1), -1.074(0)$	1	15/8/97
	$(r_t)^2$	BIC	$2.005(1), 2.009(0)$	1	15/8/97
		LWZ	$2.010(0)$	0	
NIKKEI	$|r_t|$	BIC	$-1.874(2), -1.867(1)$	2	15/9/92, 30/7/97
		LWZ	$-1.857(1), -1.851(0)$	1	20/8/98
	$(r_t)^2$	BIC	$-0.457(2), -0.452(1)$	2	15/9/92, 14/10/97
		LWZ	$-0.448(0)$	0	
S&P500	$|r_t|$	BIC	$-2.525(3), -2.513(2)$	3	27/12/91, 5/1/96, 28/7/98
		LWZ	$-2.492(2), -2.491(1)$	2	20/8/91, 3/2/97
	$(r_t)^2$	BIC	$-1.602(1), -1.559(0)$	1	14/10/97
		LWZ	$-1.591(1), -1.559(0)$	1	14/10/97

Notes: For brief data description refer to note 1, Table VI. The Lavielle and Moulines test is described in Section 2.2. The number of segments for multiple breaks denoted by m is set equal to 3. The selection criteria BIC and LWZ refer to the Bayesian or Schwarz Information Criterion and modified BIC proposed in Liu *et al.* (1997).

of diffusion processes with stochastic volatility (briefly discussed in Section 2.3). These high-frequency volatility estimates have been introduced by Merton (1980) and applied in Poterba and Summers (1986), French *et al.* (1987) and Hsieh (1991) *inter alia*. More recently Andersen and Bollerslev (1998) reintroduced these filters using intraday data, similar to Hsieh (1991). Based on a continuous time diffusion process Andersen and Bollerslev (1998) estimate the one-day Quadratic Variation ($QV1$) which is also called integrated volatility and defined as the sum of the squared returns, $r^2_{(m),t}$, for the intraday frequency m, to produce the daily volatility measure: $QV1$, discussed in Section 2.3, using 5-minute sampling frequency the lag length is $m = 288$ for financial markets open 24 hours per day (e.g. FX markets). $QV1$ can be considered as an efficient estimate of the quadratic variation of a stock returns process. One reason for their efficiency being that they utilize the high-frequency intraday data information. The $QV1$ filter is generalized in Andreou and Ghysels (2002) using the results in Foster and Nelson (1996) to increase the window length $k = 2, 3$ days in QVk and to suggest rolling instead of block sampling schemes. The rolling estimation method yields the one-day Historical Quadratic Variation ($HQV1$) defined as the sum of m rolling QV estimates, as discussed in Section 2.3, which is also extended to a k window length, $HQVk$. The rolling estimation method yields smooth volatility filters which answers one of the criticisms of the $QV1$ filter (see, for instance, Andreou and Ghysels, 2002, Barndorff-Nielsen and Shephard, 2002). The K&L and I&T tests are applied to these estimates of the quadratic variation and compared with the results for $(r_t)^2$. The results in Table IX reveal the existence of a single change-point that is detected in all the QV type filters by the $U_{\max}/\hat{\sigma}_{VARHAC}$ and IT even at the 1% significance level as opposed to the mixed evidence of a change-point in $(r_t)^2$ and $|r_t|$. This change-point in the quadratic variation of the YN/US$ series is consistently estimated by the

Table VIII. Sequential sample segmentation approach of the Kokoszka and Leipus (2000) test for detecting multiple breaks in stock market returns

SMI	Process	Subsamples (observat.)	Kokoszka and Leipus test			
			k		$\dfrac{\max U_T(k^*)}{\hat{\sigma}_{HAC}}$	Number of breaks
			Observ.	Date		
FTSE	$\lvert r_t \rvert$	1-3338	2197	5/6/97	5.861*	4
		1-2197	1021	2/12/92	3.508*	
		2198-3338	2497	30/7/98	1.366*	
		1021-2197	1631	5/4/95	2.041*	
	$(r_t)^2$	1-3338	2239	4/8/97	5.265*	4
		1-2239	991	21/10/92	2.744*	
		992-2239	1556	21/12/94	1.734*	
		992-1556	1279	29/11/93	2.507*	
HSI	$\lvert r_t \rvert$	1-3338	2247	18/8/97	4.619*	6
		1-2247	1630	4/4/95	1.815*	
		2248-3338	2554	19/10/98	2.818*	
		1-1630	1239	4/10/93	2.416*	
		1240-1630	1404	23/5/94	2.540*	
		1405-1630	1534	21/11/94	1.953*	
	$(r_t)^2$	1-3338	2249	18/8/97	2.181*	2
		2250-3338	2554	22/10/98	2.072*	
NIKKEI	$\lvert r_t \rvert$	1-3338	2237	31/7/97	3.091*	5
		1-2237	966	16/9/92	2.476*	
		2238-3338	2792	16/9/99	2.161*	
		1-966	293	16/2/90	3.119*	
		2238-2792	2365	27/1/98	1.716*	
	$(r_t)^2$	1-3338	2295	21/10/97	1.972*	4
		1-2295	966	16/9/92	2.470*	
		2296-3338	2861	22/12/99	1.757*	
		1-966	295	20/2/90	2.162*	
S&P500	$\lvert r_t \rvert$	1-3338	2110	4/2/97	6.841*	4
		1-2110	776	30/12/91	4.143*	
		2111-3338	2495	28/7/98	1.863*	
		780-2110	1776	25/10/95	2.130*	
	$(r_t)^2$	1-3338	2146	26/3/97	4.888*	3
		1-2146	779	31/12/91	3.159*	
		780-2146	1814	18/12/95	2.343*	

Notes: The test is based on a sequential sample segmentation approach with a 1% significance level for each segment.

high-frequency volatility filters to be located on the 8/2/1993 and 9/2/1993 and is associated with the highest increase of the YN *vis-à-vis* the US dollar since the 1970s and the possibility of Central Bank interventions (as published in the *Asian Wall Street Journal* dated 23 February, 1993).

The empirical analysis so far applied single and multiple breaks test procedures and identified the common dates estimated by the above tests as change-points. In an approach to verify that there was indeed a structural change in the asset returns processes we examine the volatility characteristics of the series in alternative subsamples—prior and after the breaks. The results in Table X report the estimated MLE parameters from a Normal GARCH(1,1). The varying estimated coefficients of volatility persistence and unconditional variance over the subsamples can be considered as further supportive empirical evidence that complements the change-point tests.

Table IX. Testing for a *single* change-point in high-frequency volatility filters in the YN/US$ during the period 1986–1996

Volatility filters	$k*$	Change-point statistics				
		Kokoszka & Leipus			Inclan & Tiao & extention	
		$\max U_{k*}$	$\dfrac{\max U_{k*}}{\widehat{\sigma}_{HAC}}$	$\dfrac{\max U_{k*}}{\widehat{\sigma}_{ARMA}}$	IT	$B_T(C)$
$|r_t|$	26/4/91	0.3538	1.493*	1.589*	1.996*	0.451
$(r_t)^2$	—	0.2676	1.120	1.273	1.151	0.260
$QV1$	9/2/93	0.3445	1.925*	3.845*	2.302*	—
$QV2$	9/2/93	0.3443	1.262	7.685*	2.301*	—
$QV3$	9/2/93	0.3442	1.021	11.212*	2.300*	—
$HQV1$	8/2/93	0.3428	1.804*	4.222*	2.291*	—
$HQV2$	8/2/93	0.3429	1.207	8.467*	2.292*	—
$HQV3$	9/2/93	0.3432	0.948	12.435*	2.294*	—

Notes: (1) The Yen *vis-a-vis* the US dollar returns over the period 1/12/1986-30/11/1996 at the 5-minute sampling frequency is analysed for structural changes. The data source is Olsen and Associates. The original sample is 1,052,064 five-minute return observations ($2,653 \times 288$ five-minute intervals per day). The returns for some days were removed from the sample to avoid having regular and predictable market closures which affect the characterization of volatility dynamics. The final sample includes 705,024 five-minute returns reflecting $N = 2448$ days. (2) The one-day Quadratic Variation ($QV1$) is the sum of squared returns $r_{(m),t}$ for the intraday frequency m, to produce the daily volatility measure: $QV1 = \sum_{j=1}^{m} r^2_{(m),t+1-j/m}$, $t = 1 \cdots, T_{days}$, where for the 5-minute sampling frequency the lag length is $= 288$ observations for financial markets open 24 hours per day. In $QV2$ and $QV3$ the window length is $k = 2, 3$ days, respectively. The rolling estimation method yields the one-day Historical Quadratic Variation ($HQV1$) defined as the sum of m rolling QV estimates: $HQV1 = 1/m \sum_{j=1}^{m} QV1_{(m),t+1-j/m}$, $t = 1 \cdots, T_{days}$, which is also extended to a k window length, $HQVk$. (3) The tests are described in the notes of Table VI.

5. CONCLUSIONS

There is a substantial literature on testing for the presence of breaks in linearly dependent stochastic processes. The purpose of this paper is to explore recent advances in the theory of change-point estimation, using various new CUSUM type change-point estimators and tests for multiple breaks in the context of volatility models. The tests are not model-specific and apply to a large class of strongly dependent processes such as ARCH and SV type processes and were developed in a series of recent papers in particular by Kokoszka and Leipus (1998, 1999, 2000) and Lavielle and Moulines (2000). We focus on the Kokoszka and Leipus (2000) and Lavielle and Moulines (2000) tests which monitor non-linear transformations of returns processes (in square and absolute returns) without the need to specify any particular, restrictive functional form of the process. Moreover, the CUSUM type test of Kokoszka and Leipus and the RSS minimization type test of Lavielle and Moulines are characterized by relative computational simplicity which is an additional advantage for the complex nonlinear structure of financial time series. So far only limited simulation and empirical evidence is reported about these tests. We enlarge the scope of applicability by suggesting several improvements that enhance the practical implementation of the proposed tests. The extensive simulation investigation regarding the performance of the Kokoszka and Leipus test provides evidence that the test has good power properties in detecting even small changes in all the GARCH parameters and the error and appears robust to outliers, but suffers

Table X. Estimating volatility dynamics in subsamples prior and post the change-points of stock market returns indices

SMI	Process	Subsamples (observations)	k^* Date	Normal GARCH(1,1) Estimates		
				ω	α	β
FTSE	$(r_t)^2$	1-3338		0.003[4.409]	0.067[10.46]	0.919[111.8]
		1-2484	13/7/98	**0.003**[3.669]	**0.058**[6.744]	**0.915**[64.18]
		2485-3338		**0.017**[2.839]	**0.100**[4.410]	**0.845**[22.40]
HSI	$(r_t)^2$	1-3338		0.016[15.80]	0.124[19.29]	0.855[128.6]
		1-2248	15/8/97	0.023[18.71]	**0.134**[17.04]	0.810[97.42]
		2249-3338		0.029[4.387]	**0.092**[7.462]	0.878[63.48]
NIKKEI	$(r_t)^2$	1-3338		0.005[6.979]	0.134[16.33]	0.859[105.4]
		1-966	16/9/92	**0.002**[2.206]	**0.229**[11.11]	0.801[58.14]
		967-2290	14/10/97	**0.007**[4.959]	**0.080**[7.805]	0.877[52.79]
		2291-3338		**0.011**[3.125]	**0.100**[5.752]	0.875[41.36]
S&P500	$(r_t)^2$	1-3338		0.0007[4.712]	0.039[12.979]	0.958[294.4]
		1-2291	15/10/97	**0.0003**[3.256]	**0.018**[7.796]	**0.979**[409.1]
		2292-3338		**0.013**[3.440]	**0.075**[5.801]	**0.883**[42.41]

Note: The Moulines and Lavielle (2000) multiple breaks results in Table VII for the absolute and squared returns processes are used to create various subsamples of each stock market return index. The estimated Normal GARCH(1,1) coefficients as well as the Power-ARCH coefficients are reported for the total sample ($T = 1 - 3338$) as the various subsamples determined by the estimated break points. Although not all subsamples have equal size some are approximately equal which allow for a better comparison of the estimated parameters. The bold parameters emphasize the change in the size of the volatility estimates in most subsamples (especially the parameters referring to the constant and ARCH effects of dynamic volatility).

some size distortions in the persistent GARCH case.[12] For the multiple breaks hypothesis we find that both tests share good power properties especially the BIC criterion in the RSS of Lavielle and Moulines. We also suggest the application of these change-point tests to more precise measures of volatility, including the high frequency data-driven processes studied by Andersen *et al.* (2001), Andreou and Ghysels (2002), Barndorff-Nielsen and Shephard (2002), among others.

The empirical analysis examines various financial series, including equity index returns for several financial markets in Hong Kong, Japan, the UK and the USA. The data series are similar to several prior studies, particularly Granger and Hyung (1999) who consider a longer but less recent sample. The applications of the Kokoszka and Leipus as well as the Lavielle and Moulines tests detect change-points in the volatility dynamics which are associated with the Asian and Russian financial crises. The empirical analysis is also performed using high frequency data from the FX markets. The above tests are applied to the Yen/US$ class of data driven volatility filters in an attempt to provide more efficient approximations of the quadratic variation of the process for which we also detect change-points.

[12] The IGARCH type of models violate the assumption of finite fourth moments required by the Kokoszka and Leipus tests.

REFERENCES

Aggarwal R, Inclán C, Leal R. 1999. Volatility in emerging stock markets. *Journal of Financial and Quantitative Analysis* **34**: 33–55.

Andersen T, Bollerslev T. 1998. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review* **39**: 885–905.

Andersen T, Bollerslev T, Diebold FX, Labys P. 2001. The distribution of exchange rate volatility. *Journal of American Statistical Association* **96**: 42–55.

Andreou E. 2002. Testing for change-points in long memory volatility models for stock returns. Mimeo, University of Cyprus.

Andreou E, Ghysels E. 2002. Rolling sample volatility estimators: some new theoretical, simulation and empirical results. *Journal of Business and Economic Statistics* **20**: 363–376.

Andrews D. 1993. Tests for parameter instability and structural change with unknown change point. *Econometrica* **61**(4): 821–856.

Bai J. 1994. Least squares estimation of a shift in linear process. *Journal of Time Series Analysis* **15**: 453–472.

Bai J. 1997. Estimating multiple breaks one at a time. *Econometric Theory* **13**: 315–351.

Bai J, Perron P. 1998. Estimating and testing linear models with multiple structural changes. *Econometrica* **66**: 47–78.

Bai J, Perron P. 2002. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* (forthcoming).

Barndorff-Nielsen OE, Shephard N. 2002. Econometric analysis of realised volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* **64**: 253–280.

Bellman R, Roth R. 1969. Curve fitting by segmented straight lines. *Journal of American Statistical Association* **64**: 1079–1084.

Bollerslev T, Engle RF, Nelson DB. 1994. ARCH models. In *Handbook of Econometrics*, Volume IV, Engle RF, McFadden Dan (eds). North-Holland; Amsterdam: 2959–3038.

Carrasco M, Chen X. 2001. Mixing and moment properties of various GARCH and stochastic volatility models. *Econometric Theory* **18**: 17–39.

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

Chu C-S. 1995. Detecting parameter shift in GARCH models. *Econometric Reviews* **14**: 241–266.

Csörgo M, Horváth L. 1997. *Limit Theorems in Change-Point Analysis*. Wiley: New York.

den Haan WJ, Levin A. 1997. A Practioner's guide to robust covariance matrix estimation. In *Handbook of Statistics—Vol. 15*, Rao CR, Maddala GS (eds). 291–341.

Diebold FX. 1986. Modeling the persistence of conditional variances: A comment. *Econometric Reviews* **5**: 51–56.

Diebold FX, Inoue A. 2001. Long memory and structural change. *Journal of Econometrics* **105**: 131–159.

Ding Z, Granger CWJ, Engle RF. 1993. Long-memory properties of stock market returns and a new model. *Journal of Empirical Finance* **1**: 83–106.

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

Engle RF, Smith A. 1999. Stochastic permanent breaks. *Review of Economics and Statistics* **81**: 553–574.

Francq C, Zakoïan J-M. 2000a. Estimating stochastic volatility models: A new approach based on ARMA representations. CREST Discussion Paper No. 2000-47.

Francq C, Zakoïan J-M. 2000b. Estimating weak GARCH representations. *Econometric Theory* **16**: 692–728.

French KR, Schwert GW, Stambaugh RF. 1987. Expected stock returns and volatility. *Journal of Financial Economics* **19**: 3–29.

Foster D, Nelson D. 1996. Continuous record asymptotics for rolling sample estimators. *Econometrica* **64**: 139–174.

Granger CWJ, Ding Z. 1996. Varieties of long memory models. *Journal of Econometrics* **73**(1): 61–77.

Granger CWJ, Hyung N. 1999. Occasional structural breaks and long memory. Discussion Paper 99-14, Department of Economics, University of California, San Diego.

Hendry DF. 1986. An excursion into conditional variance land. *Econometric Reviews* **5**: 63–69.

Hsieh DA. 1991. Chaos and nonlinear dynamics: application to financial markets. *Journal of Finance* **46**: 1839–1877.

Inclán C, Tiao G. 1994. Use of cumulative sums of squares for retrospective detection of change of variance. *Journal of the American Statistical Association* **89**: 913–923.

Kim S, Cho S, Lee S. 2000. On the CUSUM test for parameter changes in GARCH(1,1) models. *Communications in Statistics* **29**: 445–462.

Kokoszka P, Leipus R. 1998. Change-point in the mean of dependent observations. *Statistics and Probability Letters* **40**: 385–393.

Kokoszka P, Leipus R. 1999. Testing for parameter changes in ARCH models. *Lithuanian Mathematical Journal* **39**: 231–247.

Kokoszka P, Leipus R. 2000. Change-point estimation in ARCH models. *Bernoulli* **6**: 1–28.

Lamoureux CG, Lastrapes WD. 1990. Persistence in variance, structural change and the GARCH model. *Journal of Business and Economic Statistics* **8**: 225–234.

Lavielle M, Moulines E. 2000. Least-squares estimation of an unknown number of shifts in time series. *Journal of Time Series Analysis* **20**: 33–60.

Liu J, Wu S, Zidek JV. 1997. On segmented multivariate regressions. *Statistica Sinica* **7**: 497–525.

Lundbergh S, Terasvirta T. 1998. Evaluating GARCH models. Technical Report No. 292, Working paper, Stockholm School of Economics.

Merton RC. 1980. On estimating the expected return on the market: an exploratory investigation. *Journal of Financial Economics* **8**: 323–361.

Mikosch T, Starica C. 1999. Change of structure in financial time series, long range dependence and the GARCH model. Technical Report, preprint available at http,//www.cs.nl/ eke/iwi/preprints.

Poterba JM, Summers LH. 1986. The persistence of volatility and stock market fluctuations. *American Economic Review* **76**: 1142–1151.

van Dijk D, Franses PhH, Lucas A. 1999. Testing for ARCH in the presence of additive outliers. *Journal of Applied Econometrics* **14**: 539–562.

White H. 1984. *Asymptotic Theory for Econometricians*. Academic Press: New York.

Yao Y. 1988. Estimating the number of change-points via Schwarz Criterion. *Statistics and Probability Letters* **6**: 181–189.