APPLIED STOCHASTIC MODELS IN BUSINESS AND INDUSTRY
*Appl. Stochastic Models Bus. Ind.*, 2006; **22**:573–586
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/asmb.653

# Interaction between stock indices *via* changepoint analysis

Martin J. Lenardon and Anna Amirdjanova*,†

*Department of Statistics, University of Michigan, Ann Arbor, MI 48109, U.S.A.*

SUMMARY

Stock market indices from several countries are modelled as discretely sampled diffusions whose parameters change at certain times. To estimate these times of parameter changes we employ both a sequential likelihood-ratio test and a non-parametric, spectral algorithm designed specifically for time series with multiple changepoints. Finally, we use point-process techniques to model relationships between changepoints of different financial time series. Copyright © 2006 John Wiley & Sons, Ltd.

Received 30 June 2005; Revised 3 April 2006; Accepted 15 July 2006

KEY WORDS: changepoint analysis; financial time series models; likelihood and non-parametric estimation; diffusion

## 1. INTRODUCTION

The world-wide stock market crash in October of 1987 caused a surge of interest in the interaction between national stock markets. Since then, technological advances such as the spread of the Internet have made notions of a 'global marketplace' ubiquitous; however, there is much uncertainty as to how exactly marketplaces are global. In [1] it is suggested that stock markets are more closely correlated following an increase in volatility. Further evidence of this is found in [2], where the authors examine the U.S. and Japanese stock markets. Using instrumental variables, the authors in [3] find evidence of relations between national markets but only with the U.S. affecting other markets. Examining impulse responses from a vector autoregression leads to a similar result in [4]. Hypothesis tests for 'common features' are used in [5] on quarterly output time series for G7 countries and the authors find that the observed international comovement, the term used in econometric literature to refer to the relationship

*Correspondence to: Anna Amirdjanova, University of Michigan, Department of Statistics, 439 West Hall, 1085 South University Ave., Ann Arbor, MI 48109, U.S.A.
† E-mail: anutka@umich.edu

Contract/grant sponsor: National Security Agency

between stock markets or between national output across countries, is consistent with a common synchronized business cycle. Additionally, there is the possibility that many situations in which we see long-range dependence are actually instances of structural change, as investigated by Diebold and Inoue [6].

In this paper we explore the notion that changes, resulting from shocks to the fundamental nature of the countries or markets, i.e. technology, politics or natural disasters, in the stock market of one country are transmitted, over time, to others. In contrast to the above papers, we attempt to capture this effect by using changepoint detection methods on underlying continuous-time diffusion models. In Section 2, we review likelihood-based estimation procedures for continuous-time diffusions and discretely observed diffusions. These are of interest in their own right but are also an important part of the changepoint detection algorithms employed here. In Section 3.1 we review classical likelihood-based changepoint detection schemes for single changepoints as well as their modifications designed to detect the presence of multiple changepoints. In Section 3.2, to avoid making the parametric assumption of geometric Brownian motion (GBM), we present some non-parametric changepoint detection algorithms. In Section 4, we employ point-process techniques to model the relationship between the estimated changepoints in different national markets. Finally, in Section 5, we present our findings.

## 2. LIKELIHOOD ESTIMATION

We begin by modelling each index as a discretely observed diffusion solving the following stochastic differential equation:

$$\mathrm{d}X_t = a(t, X_t, \theta_1)\,\mathrm{d}t + b(t, X_t, \theta_2)\,\mathrm{d}B_t \tag{1}$$

observed at times $t_0, t_1, \ldots, t_n$, where $\theta_1$ and $\theta_2$ are unknown parameters and $B_t$ is a standard, one-dimensional Brownian motion, and where $a(\cdot)$ and $b(\cdot)$ may be different for different indices. This is a very general set-up which includes, as special cases, many commonly proposed models in finance such as the Vasicek, Cox–Ingersoll–Ross, Brennan–Schwartz and GBM. There are numerous approaches to estimating $\theta_1$ and $\theta_2$. One can form a continuous-time log-likelihood

$$\log L(\Theta) = \int_0^T \frac{a(t, X_t, \theta_1)}{[b(t, X_t, \theta_2)]^2}\,\mathrm{d}X_t - \frac{1}{2}\int_0^T \left(\frac{a(t, X_t, \theta_1)}{b(t, X_t, \theta_2)}\right)^2 \mathrm{d}t \tag{2}$$

where we use $\Theta$ to denote the vector $(\theta_1, \theta_2)$. This can be maximized analytically to obtain continuous-time maximum likelihood estimates (MLEs) for $\theta_1$ and $\theta_2$, which will be stochastic integrals with respect to the semi-martingale $X_t$. These integrals will have to be approximated by sums with evaluation at the observed time points. Alternatively, one could approximate (2) with

$$\sum_{t_i \leqslant T} \frac{a(t_{i-1}, X_{t_{i-1}}, \theta_1)}{[b(t_{i-1}, X_{t_{i-1}}, \theta_2)]^2}(X_{t_i} - X_{t_{i-1}}) - \frac{1}{2}\sum_{t_i \leqslant T}\left(\frac{a(t_{i-1}, X_{t_{i-1}}, \theta_1)}{b(t_{i-1}, X_{t_{i-1}}, \theta_2)}\right)^2 \Delta t_i \tag{3}$$

$$\Delta t_i = t_i - t_{i-1} \tag{4}$$

and maximize numerically. Examples and comparisons of these techniques can be found in [7, 8]. Another way of constructing the likelihood for discrete observations utilizes the Markov



DOI: 10.1002/asmb

property of solutions to (1):

$$\log L(\Theta) = \sum_{t_i \leqslant T} \log p(X_{t_i} \mid X_{t_{i-1}}) \tag{5}$$

where $p(X_{t_i} \mid X_{t_{i-1}})$ is the density of the conditional distribution of $X_{t_i}$ given $X_{t_{i-1}}$, called the transition density of $X_t$. There are numerous methods for finding or approximating the transition density, e.g. see [9, 10].

A third method for constructing the likelihood is analogous to the deterministic case; a recursive scheme for approximating a solution to (1) is given by Euler's method:

$$\mathring{X}_{ih} = \mathring{X}_{(i-1)h} + a((i-1)h, \mathring{X}_{(i-1)h}, \theta_1)h + b((i-1)h, \mathring{X}_{(i-1)h}, \theta_2)\sqrt{h}Z_i$$

for $i = 1, \ldots, t_n/h$ with stepsize $h$ and $Z_i$ i.i.d. standard normal random variables.

This implies

$$\mathring{X}_{t+h} \mid \mathring{X}_t = x_t \sim N(x_t + a(t, x_t, \theta_1)h, [b(t, x_t, \theta_2)]^2 h) \tag{6}$$

which, in the context of (5), is a Gaussian approximation to the transition density of the Markov process $X_t$. From this we get the approximate log-likelihood conditional on $X_{t_1}$:

$$\log L(\Theta) = \sum_{i=1}^{n} \left[ -\frac{1}{2}\log 2\pi - \log\sqrt{\Delta t_i}\, b(t_{i-1}, X_{t_{i-1}}, \theta_2) - \frac{1}{2}\frac{(X_{t_i} - (X_{t_{i-1}} + a(t_{i-1}, X_{t_{i-1}}, \theta_1)\Delta t_i))^2}{[b(t_{i-1}, X_{t_{i-1}}, \theta_2)]^2 \Delta t_i} \right] \tag{7}$$

We can improve the performance of the Euler scheme by transforming $X_t$ into a process with constant diffusion coefficient; this makes the transition density 'more Gaussian' [11]. Consider the transformation

$$Y_t = \phi(t, X_t)$$

Then by Itô's Lemma

$$\mathrm{d}Y_t = \left( \frac{\partial \phi}{\partial t} + \frac{1}{2}\frac{\partial^2 \phi}{\partial x^2}[b(t, X_t, \theta_2)]^2 + \frac{\partial \phi}{\partial x} a(t, X_t, \theta_1) \right) \mathrm{d}t + \frac{\partial \phi}{\partial x} b(t, X_t, \theta_2)\, \mathrm{d}B_t$$

Thus, if $\phi$ does not depend on $t$ and

$$\phi'(x) b(t, x, \theta_2) = \theta_2 \tag{8}$$

then

$$\mathrm{d}Y_t = (\tfrac{1}{2}\phi''(X_t)[b(t, X_t, \theta_2)]^2 + \phi'(X_t) a(t, X_t, \theta_1))\, \mathrm{d}t + \theta_2\, \mathrm{d}B_t$$

In the case of GBM, $a(t, x_t, \theta_1) = \mu x_t$ and $b(t, x_t, \theta_2) = \sigma x_t$, so (8) gives us

$$\phi'(x)\sigma x = \sigma$$

which results in

$$\phi(x) = \log(x)$$

and

$$\mathrm{d}Y_t = (\mu - \tfrac{1}{2}\sigma^2)\, \mathrm{d}t + \sigma\, \mathrm{d}B_t \tag{9}$$



DOI: 10.1002/asmb

So the Euler scheme gives us

$$\begin{aligned} \overset{\circ}{Y}_{t+1} &= \overset{\circ}{Y}_t + \mu - \tfrac{1}{2}\sigma^2 + \sigma Z_t \\ \overset{\circ}{Y}_{t+1} - \overset{\circ}{Y}_t &\sim N(\mu - \tfrac{1}{2}\sigma^2, \sigma^2) \end{aligned} \tag{10}$$

which is the typical difference of logs transformation. Looking at (9) we see that in this case the Euler scheme will give us the exact likelihood. This is possible because the transformed drift does not depend on $Y_t$. From (10) we get MLEs conditional on the first observation:

$$\begin{aligned} \hat{\mu} &= \left[\frac{1}{N-1}\sum_{t=1}^{N-1}(Y_{t+1} - Y_t)\right] + \frac{1}{2}\hat{\sigma}^2 = \frac{1}{N-1}(Y_N - Y_1) + \frac{1}{2}\hat{\sigma}^2 \\ \hat{\sigma}^2 &= \frac{1}{N-1}\sum_{t=1}^{N-1}\left(Y_{t+1} - Y_t - \frac{1}{N-1}(Y_N - Y_1)\right)^2 \end{aligned} \tag{11}$$

For the non-transformed process the Euler scheme would give us conditional MLEs:

$$\hat{\mu} = \frac{1}{N-1}\sum_{t=1}^{N-1}\frac{X_{t+1}}{X_t} - 1$$

$$\hat{\sigma}^2 = \frac{1}{N-1}\sum_{t=1}^{N-1}\left(\frac{X_{t+1}}{X_t} - (1+\hat{\mu})\right)^2$$

The estimates obtained using the MLEs (11) for the financial time series studied in this paper can be found in Table II. Note that for $b(t, X_t, \theta_2) = \theta_2 X_t^{\gamma}$ (which includes the Vasicek, Cox–Ingersoll–Ross and Brennan–Schwartz models) the appropriate transformation is

$$\phi(x) = \frac{x^{1-\gamma}}{1-\gamma}$$

which is analogous to the so-called Box–Cox transformation of linear regression.

## 3. CHANGEPOINT DETECTION

Typically, the motivation for changepoint detection is to decide how much of the data is relevant: it is very easy to get financial data going back 20 or more years but it is not clear that behaviour 20 years ago has anything to do with behaviour today. Indeed, the terms 'changepoint detection' and 'segmentation' are often used interchangeably and this is precisely the idea: to divide the data into segments which are 'similar'. In this paper, however, our attention is focused on the changepoints themselves, since we are interested in the transmission of shocks from one market to another.

### *3.1. Likelihood-based changepoint detection*

Suppose for simplicity we have observations of $X_t$ at times $t = 1, \ldots, n$. Our goal will be to detect times at which the parameters $\theta_1$ and $\theta_2$ change. Intuitively, to decide if there is a change at time $k$ we can maximize the log likelihood on $t = 1, \ldots, k-1$ and on $t = k, \ldots, n$ and



DOI: 10.1002/asmb

compare this to the maximized log likelihood on $t = 1, \ldots, n$. If allowing a change at time $k$ increases the likelihood 'a lot' then we suspect there is a change. Indeed, this sort of procedure enjoys optimality properties under a variety of circumstances, see [12, 13]; however, these optimality results are all for the case of a single changepoint. To deal with multiple changepoints, modifications need to be made: typically multiple local changepoint tests or stepwise procedures are used. For example, one could test for changepoints in a sliding window, or test in a small interval and then slowly expand the interval until a change is detected as in [14], or sequentially test the hypotheses of no change, one or more changes, etc. as in [15].

Let $l_{i,j}(\Theta)$ denote the log likelihood of observations $i$ through $j$. Then for employing the sliding window technique with window size $2w+1$ we compute for each time $t$ the statistic

$$T_t = \max_{\Theta} l_{t-w,t-1}(\Theta) + \max_{\Theta} l_{t,t+w}(\Theta) - \max_{\Theta} l_{t-w,t+w}(\Theta) \tag{12}$$

Alternatively, the slowly expanding window algorithm works as follows. Select a minimum interval size $m$ and a critical value $c$. Compute

$$S_{1,m} = \max_{1 \leqslant k \leqslant m} \left\{ \max_{\Theta} l_{1,k-1}(\Theta) + \max_{\Theta} l_{k,m}(\Theta) - \max_{\Theta} l_{1,m}(\Theta) \right\} \tag{13}$$

If $S_{1,m} \geqslant c$ then a changepoint is found at the time $k_1$ which maximizes (13), in which case move on and compute $S_{k_1,k_1+m}$. If $S_{1,m} < c$ then expand the interval, i.e. compute $S_{1,m+1}$. Continue in this way until reaching the end of the data. In contrast to the sliding window method, which provides an entire time series of the test statistic, the latter algorithm requires a critical value to be set at the beginning and provides only locations of estimated changepoints.

### 3.2. *Non-parametric changepoint detection*

We consider two non-parametric methods of changepoint detection, both based on spectral methods. Both methods work as follows: choose a window size, $n$. For each point in the series, compute the spectrum based on the previous $n$ data points, $\hat{f}_0$, and also on the subsequent $n$ data points, $\hat{f}_1$. Finally, construct a 'metric' to decide if the spectra are different enough to signify that a change has occurred, where the most popular choices are given by the Kolmogorov–Smirnov statistic, based on the cumulative periodogram (see e.g. [16]):

$$\sup_{\lambda \in [0,\pi]} \left| \int_0^{\lambda} \hat{f}_0(u)\,\mathrm{d}u - \int_0^{\lambda} \hat{f}_1(u)\,\mathrm{d}u \right|$$

and the metric based on the Kullback–Leibler discrimination information (see e.g. [17]):

$$\frac{1}{n} \sum_{\lambda} \left( \frac{\hat{f}_0(\lambda)}{\hat{f}_1(\lambda)} + \frac{\hat{f}_1(\lambda)}{\hat{f}_0(\lambda)} \right)$$

On the data in question the two seem to indicate the same changepoints and so specific examples will only be given for the latter case.

## 4. CHANGEPOINT COMPARISON

Both the likelihood-based and the spectral-based sliding window methods of changepoint detection provide an entire time series of test statistics. To decide to what extent changes in one



DOI: 10.1002/asmb

series lead or lag changes in another, a natural idea is to use the cross-correlation function (CCF) between the two series of test statistics. The CCF is simply the correlation between one series and lags of the other series. The expanding window algorithm provides only estimated

Table I. Indices investigated.

| Index | Symbol | Country |
|---|---|---|
| Dow Jones | DJI | U.S. |
| Standard and Poor | GSPC | U.S. |
| NASDAQ composite | IXIC | U.S. |
| CAC 40 | FCHI | France |
| DAX | GDAXI | Germany |
| FTSE 100 | FTSE | Great Britain |
| Hang Seng | HSI | Hong Kong |
| Straits Times | STI | Singapore |
| Nikkei 225 | N225 | Japan |

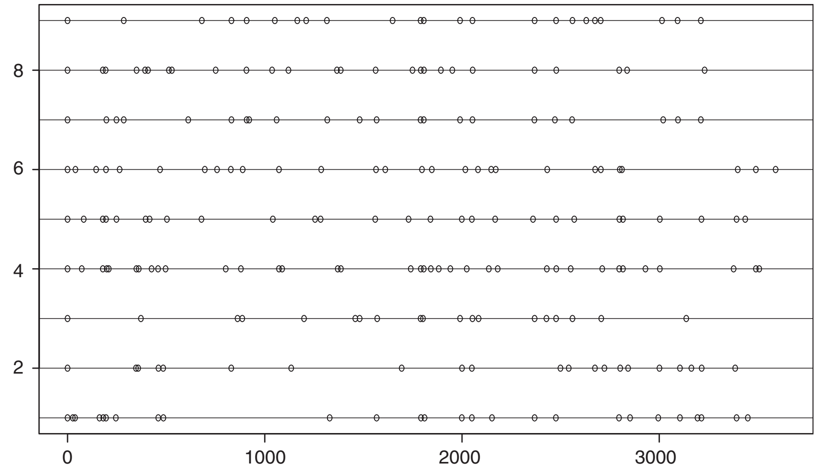


Figure 1. All estimated changepoints using the likelihood-based method from (13).

Table II. Parameter estimates using the entire series.

| Index | $\hat{\mu}$ | $\hat{\sigma}$ |
|---|---|---|
| FCHI | 0.0003300 | 0.01342 |
| FTSE | 0.0002739 | 0.01033 |
| IXIC | 0.0006155 | 0.01559 |
| STI | 0.0002447 | 0.01258 |
| GDAXI | 0.0003944 | 0.01425 |
| N225 | −0.0001083 | 0.01412 |
| DJI | 0.0004415 | 0.00989 |
| HSI | 0.0005498 | 0.01606 |
| GSPC | 0.0004132 | 0.01007 |



DOI: 10.1002/asmb

times of changes; thus we turn to point-process methods. Intuitively, a point process is just an increasing sequence of non-negative random variables, typically representing times of events. A central tool in the study of point processes is the intensity function, $\lambda(t)$ defined by

$$E[\text{number of events in } (a,b]] = \int_a^b \lambda(t)\,\mathrm{d}t \tag{14}$$

If the point process is stationary and events cannot occur simultaneously then the intensity function is given by

$$\lambda(t) = \lim_{h \downarrow 0} \frac{P(\text{event in } (t, t+h])}{h} \tag{15}$$

In this paper, we are particularly interested in how the intensity changes conditional on a past change: if there was a changepoint yesterday are we more or less likely to observe another changepoint this week? More importantly, if there was a changepoint in the U.S. market, are we

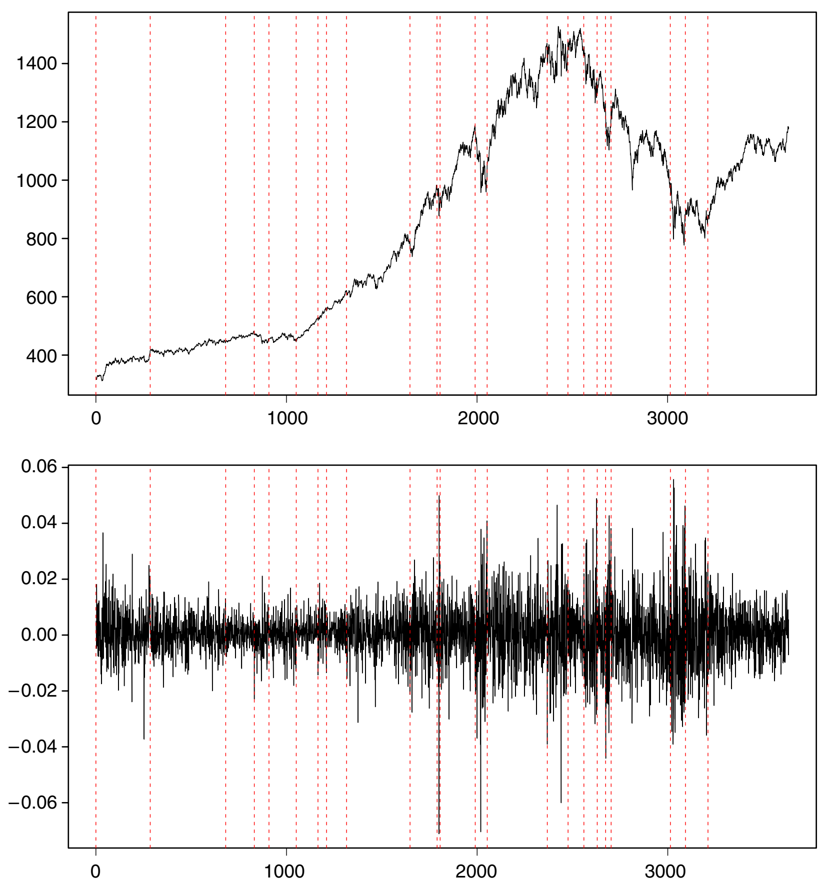


Figure 2. Raw S&P500 data (top) and difference of logs transformed S&P500 (bottom) both with likelihood-based changepoints marked with vertical lines.



DOI: 10.1002/asmb

more or less likely to observe a changepoint this week in Japan? To address these questions consider two point processes, $X$ and $Y$, with event times $\tau_1 \dots \tau_n$ and $\sigma_1 \dots \sigma_m$. Then the auto-intensity function, which measures the intensity given an event observed in the past, is defined as

$$m_{YY}(u) = \lim_{h \to 0} \frac{P(Y \text{ event in } (t, t+h] \mid Y \text{ event at } t-u)}{h} \tag{16}$$

and can be estimated, as suggested in [18], by the smoothed-histogram-like estimator:

$$\hat{m}_{YY}(u) = \frac{\#(\{j \neq k : u - h/2 < \tau_k - \tau_j < u + h/2\})}{hN_Y} \tag{17}$$

where $N_Y$ is the total number of events for the process $Y$. The cross-intensity function of $X$ and $Y$, which measures the intensity of $Y$ given an event of $X$ observed in the past, is defined as

$$m_{XY}(u) = \lim_{h \to 0} \frac{P(Y \text{ event in } (t, t+h] \mid X \text{ event at } t-u)}{h} \tag{18}$$

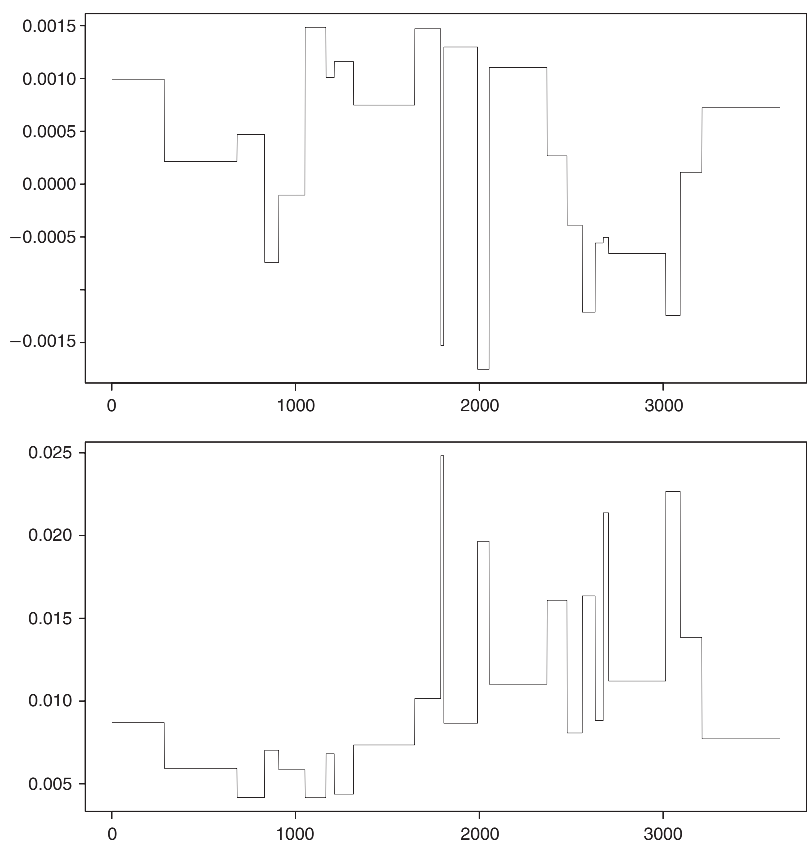


Figure 3. Estimated values of $\mu$ (top) and $\sigma$ (bottom) for the S&P500 index. The estimates are computed separately between the changepoints, which are detected by a version of the expanding window algorithm (similar to [14]).



DOI: 10.1002/asmb

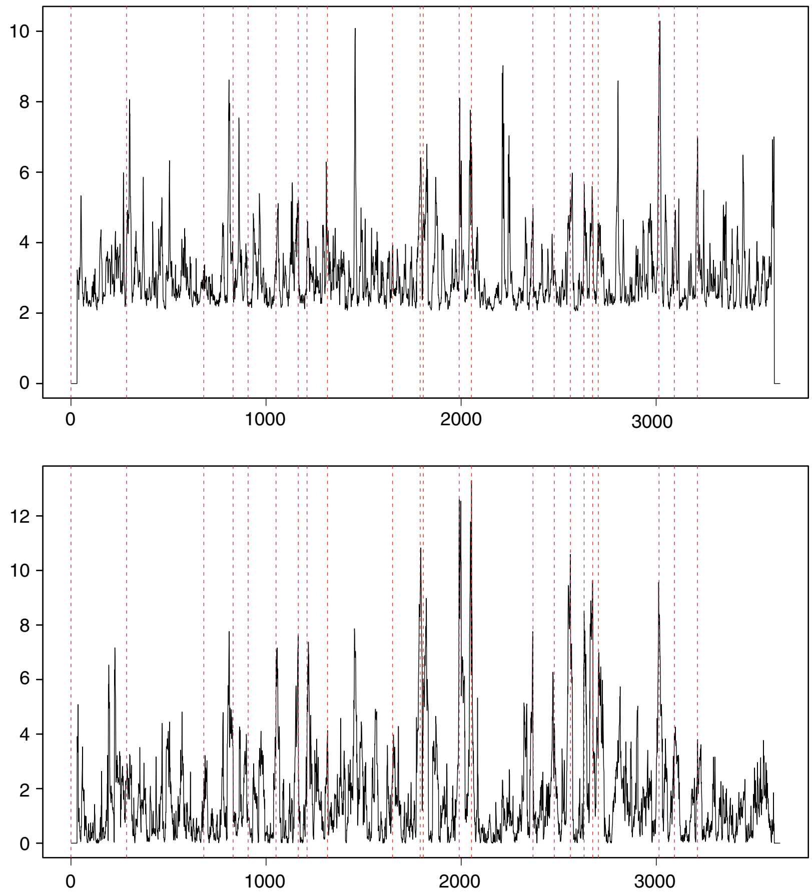


Figure 4. Spectral (top) and windowed likelihood (bottom) estimated changepoints for S&P500.

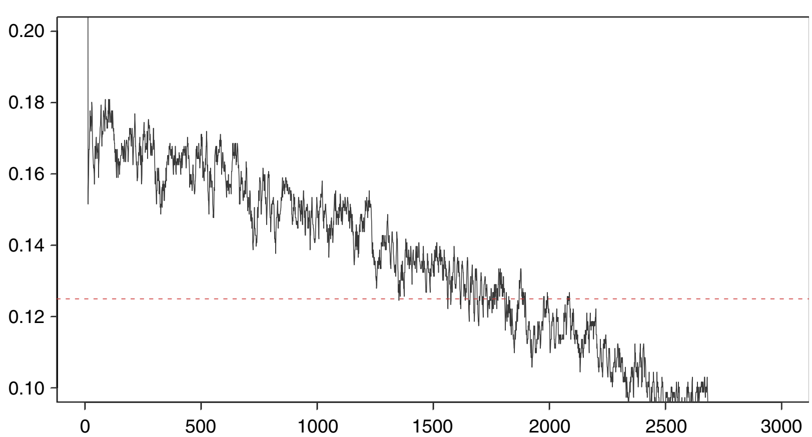


Figure 5. Estimated auto-intensity of S&P500 index with upper confidence limit using the estimator given in (17).



DOI: 10.1002/asmb

and can be similarly estimated by

$$\hat{m}_{XY}(u) = \frac{\#(\{(j,k) : u - h/2 < \tau_k - \sigma_j < u + h/2\})}{hN_Y} \tag{19}$$

Note that for both of these estimators, just as with kernel density estimators, one needs to specify a bin-size or bandwidth, $h$.

## 5. DATA ANALYSIS

We examine the closing prices of the stock indices found in Table I over the period of 3635 days from 26 November 1990 to 17 November 2004. Figure 1 shows all the estimated changepoints. Table II shows the estimates for $\mu$ and $\sigma$ under the GBM model for each index using the entire series. Figure 2 shows the raw S&P500 data and the difference of logs transformed data, as given by Equation (10), each with vertical lines at the estimated changepoints

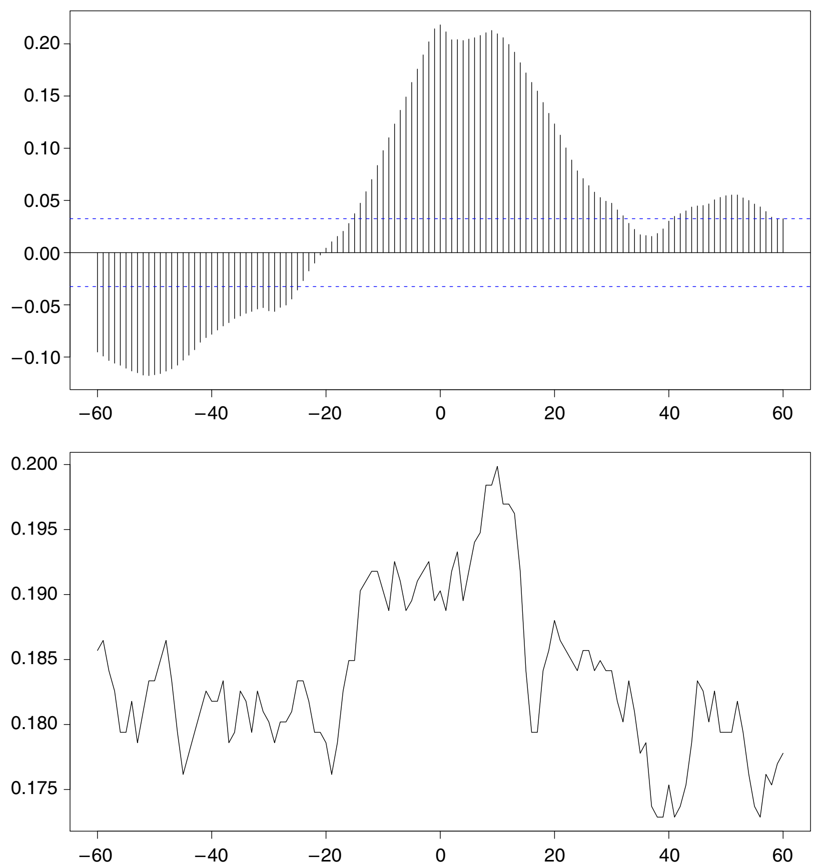


Figure 6. Estimated cross-correlation function (top) and cross-intensity (bottom) for the series of test statistics and estimated changepoints for S&P500 and DAX.



DOI: 10.1002/asmb

from the expanding window algorithm. Figure 3 shows the estimated $\mu$ and $\sigma$ for the S&P500 series where the estimates are computed separately between the changepoints, with the latter also being estimated by the expanding window algorithm. Figure 4 is quite interesting; for the S&P500 series the changepoints estimated by the expanding window algorithm (vertical lines) seem to coincide with the spikes in the series of test statistics from both the likelihood and spectral sliding window algorithms. This holds for all the other series as well.

For visual inspection of estimated cross-intensities between $X$ and $Y$ it is suggested in [18] to plot the square root of the cross-intensities and use $N_Y/T \pm (hN_Y)^{-1/2}$ as an approximate 95% confidence band for independent Poisson cross-intensities, where $N_Y$ is the total number of $Y$ events and $T$ is the total time. This is to say that intensities within the confidence band are consistent with independent Poisson processes. In all the plots that follow, the cross-intensity lies so far above the upper confidence limit that the limit is not visible on the plot. Figure 5 shows a plot of the estimated auto-intensity function for S&P500. This plot is typical of the

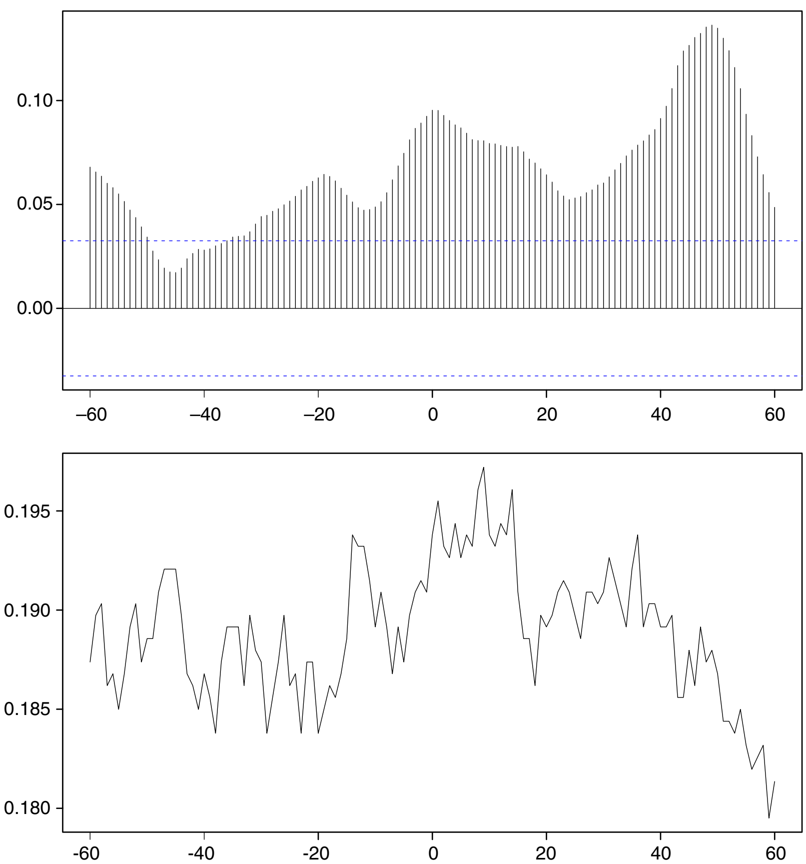


Figure 7. Estimated cross-correlation function (top) and cross-intensity (bottom) for the series of test statistics and estimated changepoints from Nikkei and Hang Seng.



DOI: 10.1002/asmb

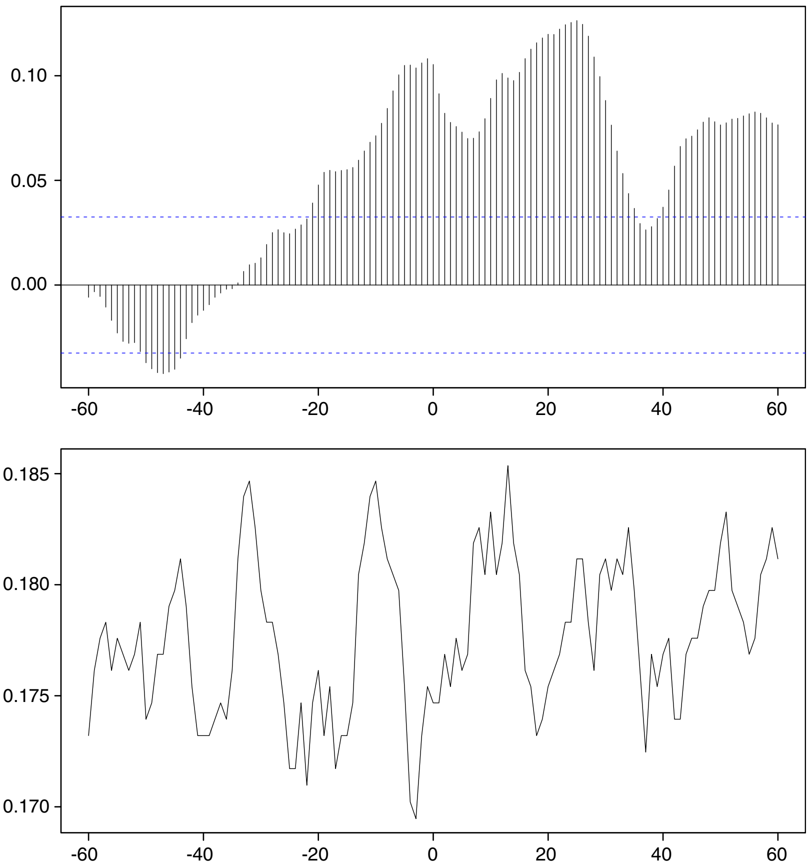


Figure 8. Estimated cross-correlation function (top) and cross-intensity (bottom) for the series of test statistics and estimated changepoints from S&P500 and Nikkei.

indices under consideration: the intensities seem to have a long memory, in the sense that changes affect the intensity of changes far into the future, and are clearly not Poisson. Figures 6 and 7 show the CCF of the time series of test statistics from (12) and the cross-intensity; Figures 8 and 9 show the same for the statistic from (13). While it is not totally clear that the sample CCF is the appropriate tool to use here, its 'naive' interpretation is in apparent agreement with the cross-intensity analysis and together with the lack of an obvious alternative make it worthwhile investigating. In Figure 6 we see a typical example of the relationship between U.S. and European indices: changes seem to occur in Europe up to 20 days after the U.S. Figures 7 and 8 show the somewhat mercurial behaviour of Nikkei: despite the fact that all changepoint methods seem to give the same changepoints, when comparing Nikkei to other indices, the choice of method seems to change the apparent relationship. Nonetheless there does appear to be some evidence of shock transmissions from the U.S. to Japan. Figure 9 shows a typical relationship between European indices: there does not appear to be any clear lag between change transmissions.

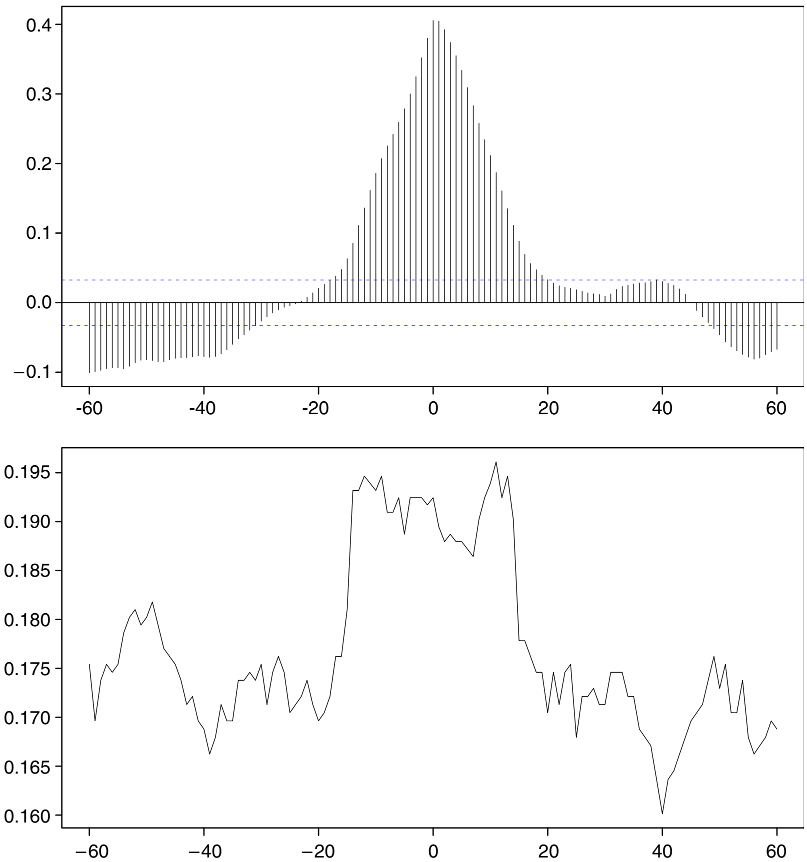


Figure 9. Estimated cross-correlation function (top) and cross-intensity (bottom) for the series of test statistics and estimated changepoints from FTSE and CAC.

## 6. CONCLUSIONS

Changepoint analysis seems to be an under-used tool in financial time series; there is a fairly unanimously held opinion that markets are intimately connected and propagation of changes provides one method of quantifying this relationship. As shown in Section 5, there appears to be some delay in the transmission of shocks from U.S. indices to European indices and some delays between the Asian indices. The European indices seem to change together, as do the U.S. indices. A particularly attractive feature of the analysis is the lack of assumptions on the process generating shocks to the markets. If the changes were due only to external events, one might suspect the changepoints to be Poisson; our analysis suggests that this is quite far from the case. Additionally, using the non-parametric changepoint techniques presented, no assumptions on the indices themselves were necessary. Interestingly, the changepoints detected by methods assuming GBM seem to coincide with those detected by methods identifying only changes in spectrum.

ACKNOWLEDGEMENTS

The authors would like to thank the anonymous referee for a number of helpful comments and suggestions. The second author is very grateful to Vijay Nair for the invitation to present parts of this research at the Fourth International Symposium on Business and Industrial Statistics and for providing incidental inspiration for this work. Grant support from the National Security Agency is also acknowledged.

REFERENCES

1. King MA, Wadhwani S. Transmission of volatility between stock markets. *The Review of Financial Studies* 1990; **3**(1):5–33.
2. Karolyi GA, Stulz RM. Why do markets move together? An investigation of U.S.–Japan stock return comovements. *The Journal of Finance* 1996; **51**(3):951–986.
3. Cheung Y-W, He J, Ng LK. Common predictable components in regional stock markets. *Journal of Business and Economic Statistics* 1997; **15**(1):35–42.
4. Eun CS, Shim S. International transmission of stock market movements. *The Journal of Financial and Quantitative Analysis* 1989; **24**(2):241–256.
5. Engle RF, Kozicki S. Testing for common features. *Journal of Business and Economic Statistics* 1993; **11**(4):369–380.
6. Diebold FX, Inoue A. Long memory and regime switching. *Journal of Econometrics* 2001; **105**:131–159.
7. Kloeden PE, Platen E, Schurz H, Sørensen M. On effects of discretization on estimators of drift parameters for diffusion processes. *Journal of Applied Probability* 1996; **33**(4):1061–1076.
8. Cleur EM. Maximum likelihood estimates of a class of one-dimensional stochastic differential equation models from discrete data. *Journal of Time Series Analysis* 2000; **22**(5):505–515.
9. Aït-Sahalia Y. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica* 2002; **70**(1):223–262.
10. Chib S, Pitt MK, Shephard N. Likelihood based inference for diffusion driven models. *OFRC Working Papers Series* 2004fe17, Oxford Financial Research Centre, 2004.
11. Durham GB, Gallant AR. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *Journal of Business and Economic Statistics* 2002; **20**(3):297–338.
12. Basseville M, Nikiforov IV. *Detection of Abrupt Changes: Theory and Application*. Prentice-Hall: Englewood Cliffs, NJ, 1993. Now available online at http://www.irisa.fr/sisthem/kniga/kniga.pdf
13. Moustakides GV. Quickest detection of abrupt changes for a class of random processes. *IEEE Transactions on Information Theory* 1998; **44**(5):1965–1968.
14. Mercurio D. Adaptive estimation for financial time series. *Ph.D. Thesis*, Humboldt-Universität zu Berlin, 2004.
15. Altissimo F, Corradi V. Strong rules for detecting the number of breaks in a time series. *Journal of Econometrics* 2003; **117**:207–244.
16. Basseville M, Benveniste A. *Detection of Abrupt Changes in Signals and Dynamical Systems*. Springer: Berlin, 1986. Papers from a conference held in Paris, March 21–22, 1984.
17. Dahlhaus R. On the Kullback–Leibler information divergence of a locally stationary process. *Stochastic Processes and their Applications* 1996; **62**:139–168.
18. Brillinger DR. Estimation of the second-order intensities of a bivariate stationary point process. *Journal of the Royal Statistical Society, Series B (Methodological)* 1976; **38**(1):60–66.