SFB 823

Multiple break detection in the correlation structure of financial returns

Pedro Galeano, Dominik Wied

Discussion Paper

Nr. 31/2011

Multiple break detection in the correlation structure of financial returns

Pedro Galeano, Dominik Wied*

Universidad Carlos III de Madrid and TU Dortmund

This Version: August 30, 2011

Abstract

Correlations between asset returns plays an important role in financial analysis. More precisely, accurate estimates of the correlation between financial returns are crucial in portfolio management. In particular, in periods of financial crisis, extreme movements in asset prices are found to be more highly correlated than small movements. It is precisely under these conditions that investors are extremely concerned about changes on correlations. We propose a sequential procedure to detect the number and position of multiple change points in the correlation structure of financial returns. It is shown analytically that the proposed algorithm asymptotically gives the correct number of change points and the change points are consistently estimated. It is also shown by simulation studies and by an empirical application that the algorithm yields reasonable results.

*Universidad Carlos III de Madrid, Departamento de Estadística, E-28903 Getafe, Madrid, Spain. Email: pedro.galeano@uc3m.es (P. Galeano). TU Dortmund, Fakultät Statistik, D-44221 Dortmund, Germany. Email: wied@statistik.tu-dortmund.de, Phone: +49 231 755 3869 (D. Wied). Financial support by MCI grant MTM2008-03010 and Deutsche Forschungsgemeinschaft (SFB 823, project A1) is gratefully acknowledged.

Keywords: Correlations; CUSUM statistics; Financial returns; Multiple change point detection; Sequential procedure.

JEL Classification: C12, C14, C63, G12

1. Introduction and Summary

There are many empirical hints that the correlation structure of financial returns of all sorts cannot be assumed to be constant over time, see e.g. Longin and Solnik (1995) and Krishan et al. (2009). Especially in times of crisis, correlation often increases, a phenomenon which is referred to as "Diversification Meltdown" (Campbell et al., 2008). Recently, Wied, Krämer and Dehling (2011) proposed a CUSUM type procedure adapting ideas by Ploberger et al. (1989) to formally test if correlations remain constant over time. Given a bivariate time series of returns, denoted by (X_t, Y_t), the fluctuation test uses the test statistic

$$Q_T(X,Y) = \hat{D} \max_{2 \le j \le T} \frac{j}{\sqrt{T}} \left| \hat{\rho}_j - \hat{\rho}_T \right|,$$

where $\hat{\rho}_j$ is the empirical correlation up to time j, for $j = 2, \ldots, T$ and $\hat{D}$ is a normalizing constant which is described in Wied et al. (2011). The null hypothesis of constant correlation is rejected when the test statistic becomes too large, i.e. when the estimated correlations fluctuate too much over time. However, with this approach the practitioner is just able to see if there is a change or not; he cannot determine where a possible change occurs or how many changes we have.

The present paper fills this gap by proposing an algorithm based on the correlation constancy test to estimate the number of change points and the time of the changes as a fraction in the interval $[0, 1]$. For this purpose, we adapt a method for sequential estimation of multiple breaks which was dealt with or implemented in various problems by Vostrikova (1981), Inclán and Tiao (1994), Bai (1997), Bai and Perron (1998), Andreou and Ghysels (2002) and Galeano and Tsay (2010), among others. The segmentation algorithm proceeds as follows: First, we want to find the (what we will later call) dominating change point and decide if this point is statistically significant. Second, we split the return series in two parts and look for possible change points again in each part of the series. The procedure stops if we do not find any new change point any more. In this paper, we will analytically show that the algorithm asymptotically gives the correct number of change points and the change points are consistently estimated, assuming that there exists a finite number of change points. Furthermore, we will show that the algorithm gives reasonable results in simulations with finite samples and in an empirical application.

The rest of the paper is organized as follows. Section 2 introduces the proposed procedure. In Section 3, we derive the asymptotic properties of the procedure. In Sections 4 and 5, we present some simulation studies and a real data application, respectively. Section 6 provides some conclusions. All proofs are presented in the Appendix.

2. Algorithm for the detection of change points

In this section, we present the algorithm for detection of change points in financial returns. It is made for a bivariate series of random variables and can be applied to general sequences of random variables with some serial dependency.

To be more precisely, let $(X_t, Y_t), t \in \mathbb{Z}$, be a sequence of bivariate random variables with finite first four moments and let $1, \ldots, T$ be the observation period. The series (X_t, Y_t) are assumed to be near-epoch dependent on a strong mixing or uniform mixing sequence. Therefore, variations of the variances are also permitted to a certain extent and for example GARCH-effects are covered by our assumptions. For more details about technical assumptions see Wied et al. (2011). Denoting the correlation between X_t and Y_t with

$$\rho_t = \frac{Cov(X_t, Y_t)}{\sqrt{Var(X_t)}\sqrt{Var(Y_t)}},$$

the test problem is

$$H_0 : \rho_t = \rho_0 \ \forall t \in \{1, \ldots, T\} \text{ vs. } H_1 : \exists t \in \{1, \ldots, T-1\} : \rho_t \neq \rho_{t+1}$$

for a constant ρ_0. In this paper, we basically focus on the test's behavior under the alternative.

We assume that there is a finite number of change points. However, as usual in practice, the number, location and size of the change points are unknown. The formal assumption is:

Assumption 1. *Under the alternative, expectations and variances are constant and equal to μ_x, μ_y, σ_x^2 and σ_y^2, the second cross moment changes from $\mathsf{E}(X_iY_i) = m_{xy}$ to $\mathsf{E}(X_iY_i) = m_{xy} + g\left(\frac{i}{T}\right)$. The function $g(z), z \in [0,1]$ is a step function with a finite number of steps b, i.e. there is a partition $0 = s_0 < s_1 < \ldots < s_b < s_{b+1} = 1$ such that*

$$g(z) = \sum_{i=0}^{b} a_i \mathbf{1}_{\{z \in [s_i, s_{i+1})\}}$$

and $g(1) = a_b$.

The function g is the function that gives information about the time and height of the correlation jumps. Using this expression, which is similar to the expression in the local power analysis of Wied et al. (2011), we can describe the jumps in an elegant way. For ease of exposition, we assume m_{xy} to be 0. Then, for instance, assuming that $m_x = m_y = 0$ and $\sigma_x^2 = \sigma_y^2 = 1$, if, say, the correlation is equal to 0.5 in the first half of the sample, jumps to 0.7 in the middle of the sample and falls down to 0.6 after the third quarter of the sample, the function g would be

$$g(z) = 0.5 \cdot \mathbf{1}_{\{z \in [0,0.5)\}} + 0.7 \cdot \mathbf{1}_{\{z \in [0.5,0.75)\}} + 0.6 \cdot \mathbf{1}_{\{z \in [0.75,1]\}}. \tag{1}$$

The condition on the finiteness of the steps of g ensures that there is a finite number of change points b under the alternative. Our goal is to estimate b, the values $s_1, \ldots, s_b$

and the values $a_0, \dots, a_b$. For that, we propose a binary segmentation type algorithm for estimating the number, location and size of multiple correlation changes. Thereby, we concentrate on the "change point fractions" in the interval $[0,1]$. If e.g. a break occurs in the middle of the sample, the value of interest would be 0.5 and not $T/2$.

The main idea behind the procedure is to isolate each change point in different time intervals by splitting the two series into two parts once a change point is found. Then, the search of a new change point is initialized in both pieces. The proposed procedure for detecting correlation changes essentially bases on the intuitive estimator of the change point fraction. To obtain it, note that we can write,

$$Q_T(X,Y) = \sup_{z \in [0,1]} \hat{D} \frac{\tau(z)}{\sqrt{T}} \left| \hat{\rho}_{\tau(z)} - \hat{\rho}_T \right|$$

with $\tau(z) = [2 + z(T-2)]$. Then the estimator is the value of z which maximizes the function,

$$B_T(z) := \hat{D} \frac{\tau(z)}{\sqrt{T}} \left| \hat{\rho}_{\tau(z)} - \hat{\rho}_T \right| .$$

The algorithm proceeds as follows:

Change detection procedure

1. Let X_t and Y_t be the observed series. Obtain the test statistic $Q_T(X,Y)$. There are two possibilities:

   (a) If the test statistic is statistically significant, i.e., if $Q_T(X,Y) > c_{T,\alpha}$, where $c_{T,\alpha}$ is the asymptotic critical value for a given upper tail probability, then a correlation change is detected. Let z_1 be the point at which the function $B_T(z)$ attains its maximum value and go to step 2.

   (b) Otherwise, there are no correlation changes and the procedure ends.

2. Let $z_1, \dots, z_\ell$ be the ℓ change points already detected in previous iterations sorted

in increasing order. Repeat this step until,

$$\max_k \left\{ Q_T^k(X,Y),\ k=1,\ldots,\ell+1 \right\} < c_{T,\alpha},$$

where $Q_T^k(X,Y)$ is the value of the statistic $Q_T(X,Y)$ obtained for values of $\tau(z)$ with z in the interval $z_{k-1}+1/T \leq z \leq z_k$, for $k=1,\ldots,\ell+1$, taking $z_0=0$ and $z_{\ell+1}=1$.

3. Let $(z_1,\ldots,z_\ell)$ be the vector of detected change points sorted in increasing order. If $\ell>1$, refine the estimate of the location of the change points by calculating the statistic $Q_T(X,Y)$ for values of $\tau(z)$ with z in the subintervals $z_{k-1}+1/T \leq z \leq z_{k+1}$, for $k=1,\ldots,\ell$, where $z_0=0$ and $z_{\ell+1}=1$. If any of the change points is not statistically significant, delete it from the list, and repeat this step.

4. Finally, the correlation between X_t and Y_t is estimated in the intervals between change points.

Some comments on the proposed procedure are in order. First, the key point of the proposed procedure is that it detects a single change point in each iteration, which may not be the most efficient way to detect correlation changes when multiple changes exist. However, our theoretical results show that the procedure consistently detects the true change points. Moreover, the proposed procedure works well in small samples in terms of detection of the true number of changes as shown in the Monte Carlo experiments of Section 4. Second, step 3 is included to refine the estimation of the change points. Note that in this step, the procedure computes the value of the $Q_T(X,Y)$ statistics in intervals that are only affected by the presence of one change point, something not achieved in step 2. Third, the main objective of the proposed procedure is to identify points at which further attention is needed. Thus, if the number of change points detected is large compared with the sample size, then a piecewise constant correlation may not be a good way to describe the correlation between the two series. Fourth, although our theoretical results are shown assuming that the critical value used in the procedure tends to infinity with the sample size, in practice, we use different critical values in each step of the procedure. Using the same critical level in steps 2 and 3 may lead to over-estimation of the number of change points, because more tests are performed in each iteration as the number of detected change points increases and the type-I errors accumulate. Then, to avoid this multiple-test problem we assume that the type-I errors used depend on the number of change points already detected by the algorithm. In particular, in step 1 we use an initial type-I error such as $\alpha_0 = 0.05$.[1] Then, after detecting the $(\ell - 1)$-th change point, we use the critical value c_{T,α_ℓ}, where α_ℓ is given by $1 - \alpha_0 = (1 - \alpha_\ell)^{\ell+1}$. This choice of α_ℓ is taken to maintain the same significance level for all tests. Finally, we use the quantiles of the distribution of the supremum of the absolute value of a Brownian Bridge in order to apply the procedure in practice, see Wied et al. (2011). The explicit form of this distribution function can be found in Billingsley (1968).

3. Asymptotic results

In this section, we show that the algorithm proposed in Section 2 works. To this end, we maintain another assumption which guarantees that we do not have two or more change points with "equal form", i.e. we assume that there are always change points which dominate others.

Assumption 2. *Let $0 \leq l_1 < l_2 \leq 1$ be arbitrary. The function g from Assumption 1 is such that the function*

$$A^*(z) := \left| \int_{l_1}^{z} g(t)dt - z \int_{l_1}^{l_2} g(t)dt \right|, z \in [l_1, l_2],$$

is either constant or has a unique maximum.

A dominating change point is then defined as the argmax of $A^*(z)$ in a given interval $[l_1, l_2]$. We illustrate Assumption 2 in the case of example (1). On the interval $[0, 1]$ for

[1] See also the discussion on the error levels before Theorem 2.

example, the function $A^*(z)$ then looks like

$$A^*(z) = \begin{cases} 0.075 \cdot z & \text{when } z < 0.5, \\ 0.1 - 0.125 \cdot z & \text{when } 0.5 \leq z < 0.75, \\ 0.025 - 0.025 \cdot z & \text{when } 0.75 \leq z \leq 1 \end{cases}$$

and has a unique maximum at $z = 0.5$, i.e. the point with the "strongest" correlation change, see also Figure 1.

Figure 1 around here

In general, the height and the position of the change point decides if it is dominant or not. Assumption 2 is violated if the correlation jumps at equal sizes at symmetric time points, e.g. in the case

$$g(z) = 0.5 \cdot \mathbf{1}_{\{z\in[0,0.25)\}} + 0.7 \cdot \mathbf{1}_{\{z\in[0.25,0.75)\}} + 0.5 \cdot \mathbf{1}_{\{z\in[0.75,1]\}}, \tag{2}$$

because here,

$$A^*(z) = \begin{cases} 0.1 \cdot z & \text{when } z < 0.25, \\ 0.05 - 0.1 \cdot z & \text{when } 0.25 \leq z < 0.5, \\ 0.1 \cdot z - 0.05 & \text{when } 0.5 \leq z < 0.75, \\ 0.1 - 0.1 \cdot z & \text{when } 0.75 \leq z \leq 1 \end{cases}$$

has two non-unique maxima in $z = 0.25$ and $z = 0.75$, see also Figure 2.

Figure 2 around here

Finally, we need a rather technical assumption regarding the normalizing constant $\hat{D}$.

Assumption 3. *Under the alternative, $\hat{D}$ converges to a real number $D_1 \in (0, \infty)$.*

The first theorem shows that the change point estimator is consistent if it is known a priori that there is a change point in a given interval.

Theorem 1. *Let Assumptions 1, 2 and 3 be true and let there be one or several break points in a given interval* $[l_1, l_2] \subseteq [0, 1]$ *with* $l_1 < l_2$. *Then the change point estimator is consistent for the dominating change point.*

While also of interest in its own, the preceding theorem is mainly needed for the next theorem yielding the convergence of the algorithm. We require one additional assumption on the critical value $c_{T,\alpha}$. While we argued in the preceding section that we have to adjust the value for finite T due to multiple testing problems, we need another kind of assumption for the asymptotics as $T \to \infty$.

Assumption 4. *The critical value* $c_{T,\alpha}$ *used in the algorithm fulfills the condition* $\lim_{T\to\infty} c_{T,\alpha} = \infty$ *and* $c_{T,\alpha} = o(\sqrt{T})$.

The assumption rules e.g. out a choice of initial type-I error such as $\alpha_0 = 0.05$ because the initial type-I error must converge to 0. However, it is legitimate using a fixed type-I error in finite samples if we consider an upper bound for T.

Theorem 2. *Under Assumptions 1, 2, 3 and 4, the change point algorithm asymptotically gives the correct number of change points* b *and the change points are consistently estimated.*

4. Monte Carlo experiments

In this section, we carry out several Monte Carlo experiments to gain insight into the finite sample performance of the proposed procedure. In particular, we study several aspects, including the size (type-I error) of the procedure, the power of the procedure in correct detection of the changes and the ability of the procedure to accurately identify the location of the change points.

First, we check the size (type-I error) of the procedure. For that, we consider a vector autoregressive of order 1 given by:

$$\begin{pmatrix} X_t \\ Y_t \end{pmatrix} = \begin{pmatrix} \phi & 0 \\ 0 & \phi \end{pmatrix} \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_t^1 \\ \epsilon_t^2 \end{pmatrix},$$

where $(\epsilon_t^1, \epsilon_t^2)'$ are iid bivariate t_5 distributed with correlation ρ. The Student-t distribution is considered to make our data better resemble financial time series. Three values of the correlation parameter ρ are considered, $\rho = -.5$, 0 and .5. Two values of the parameter ϕ are considered, $\phi = 0$, that represent the case in which X_t and Y_t are iid observations, and $\phi = .5$, which represents the case of temporal dependency. The sample sizes considered in the experiments are $T = 200$, 500, 1000, 2000 and 3000, that are usual in financial returns. Table 1 gives the results based on 1000 replications and an initial nominal significant level of $\alpha_0 = 0.05$. From this table, it seems that the type-I error of the proposed procedure is very close to the initial nominal level even with the smallest sample size. Therefore, overdetection does not appear to be an issue for the proposed procedure if there are no changes in the correlation.

Table 1 around here

Next, we analyze the power of the proposed detection procedure when there is a single change point in the series. The Monte Carlo setup is similar to the one described previously, but the series are generated with a single change point in the correlation. Three locations of the change point are considered, $z_1 = 0.25$, 0.50 and 0.75. Thus, for each sample size T, the time points of the change are $t_1 = [0.25T]$, $[0.50T]$, and $[0.75T]$, respectively. The change is such that the correlation of the series before the change point is $\rho_0 = .5$ and then changes to $\rho_1 = 0$, .25 or .75. The value of $\rho_1 = 0$ represents a big change in correlation while the values $\rho_1 = .25$ and .75 represent moderate changes. The results are shown in Table 2. Some comments on the table apply. First, the procedure performs quite well in detecting a single change point, with many cases over 90% correct detection. Second, as the sample size increases and the size of the change gets larger, the procedure works better. However, the magnitudes of the exception are small in general. Third, when the sample size of the change is small, the probability of under-detection may be large. Fourth, the power of the procedure is larger if the correlation coefficient increases than if the correlation coefficient decreases, see the second and third part of the table for the cases $\rho_1 = .25$ or $.75$. This is particularly so when the sample size is small. Fifth, the location of the change point does not strongly affect the detection frequency of the procedure when the sample size is large. However, if the sample size is small then the procedure detects more frequently the change point at the middle of the series. Finally, in most cases, the percentage of false detection is smaller than the nominal 5%. In particular, the frequency of over-detection is small for all cases. On the other hand, Table 3 shows the median and mean absolute deviation of the change point estimators in each case. The median of the estimates are quite close to the true change point locations. Note that the larger is the size of the change, the better is estimated its location.

Table 2 around here

Table 3 around here

Next, we conduct another Monte Carlo experiment to study the power of the proposed procedure for detecting two change points. In this case, the location of the change points considered are $z_1 = 0.33$ and $z_2 = 0.66$. Thus, the time point of the changes are $t_1 = [0.33T]$ and $t_2 = [0.66T]$, respectively, for each sample size T. Three situations are considered. First, the changes are such that the correlation of the series before the first change point is $\rho_0 = .5$, then changes to $\rho_1 = 0$ at the second change point, and, finally, changes to $\rho_2 = .25$. Second, the correlation of the series before the first change point is $\rho_0 = .5$, then changes to $\rho_1 = .25$ at the second change point, and, finally, changes to $\rho_2 = .75$. Third, the correlation of the series before the first change point is $\rho_0 = .5$, then changes to $\rho_1 = .75$ at the second change point, and, finally, changes to $\rho_2 = .25$. The results are shown in Table 4. As in the case of a single change point, the proposed procedure works reasonably well, especially when the sample size is large or the size of the correlation change is large. In addition, the procedure does not overdetect the number of change points. It may underestimate the number of change points, however. The underestimation can be serious when the sample size is small, say $T = 200$. Finally, the percentage of false change points detected in both cases, one and two change points, is smaller than the nominal 5% in almost all the cases. On the other hand, Table 5 shows the median and mean absolute deviation of the estimates of the change point locations. Note that the medians of the estimates are quite close to the true ones. Also, it appears that the larger is the size of the change, the better is estimated its location.

Table 4 around here

Table 5 around here

5. Application

In this section, we look for changes in the correlation structure of the log-return series of two U.S. asset indexes: the Standard & Poors 500 Index and the IBM stock Index from January 2, 1997 to December 31, 2010 consisting of $T = 3524$ data points. Both log-returns series are plotted in Figure 3, which shows different volatility periods. The empirical correlation of both log-returns is given by 0.6225. The autocorrelation functions of the log-returns show some minor serial dependence, while the autocorrelation functions of the squared log-returns shows serial dependency, as usual in stock market returns.

Figure 3 around here

Next, we apply the proposed segmentation procedure of Section 2 to detect correlation changes for the log-returns of the S&P 500 and IBM stock indexes. Table 6 and Figures 4, 5, 6 and 7 show the iterations taken by the procedure. Similar to the simulation experiments of Section 4, we start with the asymptotic critical value at the 5% significance level. In the first iteration, the procedure detects a change in the correlation at time point $t = 988$ (November 29, 2000). Indeed, as shown in Figure 4 there are two local modes of the CUSUM statistic. The value of the test statistic (1) is 1.5699, which is significant at the 5% level. Following the proposed procedure, we split the series into two subperiods and look for changes in the subintervals $[1, 988]$ and $[989, 3524]$, respectively. In the first subinterval (see Figure 5), the procedure detects a change at time point $t = 664$ (August 19, 1999). The value of the test statistic (1) is 2.1009. Then, we split the subinterval $[1, 988]$ into two subintervals and look for changes in the subintervals $[1, 664]$, $[665, 988]$ and $[989, 3524]$ (see Figure 6). No more changes were found in these three subintervals. Then, we pass to step 3 and refine the search. For that we estimate the location of the change points in the intervals $[1, 988]$ and $[665, 3524]$, respectively. In the first subinterval (see Figure 7), as in the previous step, the procedure detects a change at time point $t = 664$ (August 19, 1999) and the value of the test statistic (1) is 2.1009. On the other hand, in the second subinterval (see Figure 7), as in the previous step, the procedure detects a change at time point $t = 2734$ (November 12, 2007) and the value of the test statistic (1) is 1.6193. These are the finally estimated change points.

Table 6 around here

Figure 4 around here

Figure 5 around here

Figure 6 around here

Figure 7 around here

The empirical correlation coefficients in the three subintervals are 0.6284, 0.5785 and 0.7823, respectively, indicating that the correlation shifted to a smaller value after the first change point and to a higher value after the second change point. Figure 8 shows the scatterplots of the two log-returns indexes at three different subperiods. It is interesting to see that the dates of the detected change points fare well with well known financial facts. The period starting at 1994 till the end of 1999 is a period of economic growth in the U.S. economy in which the inflation was under control and the unemployment rate dropped to below 5%. This is a period with high increases in the stock markets. However, the collapse of the dot-com bubble started at the end of the 1990s and the beginning of the 2000s, and the market gave back around the 75% of the growth obtained in the 1990s. However, note that, contrarily to the diversification meltdown theory, the correlation did not increase during the dot-com bubble crisis. The third estimated change point roughly corresponds to the beginning of the Global Financial Crisis around the end of 2007, which is considered by many economists the worst financial crisis since the Great Depression of the 1930s. The reduction of interest rates leads to several consequent issues starting with the easiness of obtaining credit, leading to sub-prime lending, so that a increased debt burden, and finally a liquidity shortfall in the banking system. This resulted in the collapse of well known financial institutions such as Lehman Brothers, Merrill Lynch, Washington Mutual, Wachovia, and AIG, amon others, the bailout of banks by national governments such as Bear Stearns, Citigroup, Bank of America and Northern Rock, among others, and great loses in stock markets around the world. In this case, the Global Financial Crisis produced an increase in the correlation between both log-returns. Of course, it is important to note that these economic interpretations are mere speculations. These comments only point out that, for this particular example, the proposed detection procedure in Section 2 identifies changes in the correlation structure that fare well with well known events affecting the U.S. financial market.

Figure 8 goes around here

6. Conclusions

In this paper, we have proposed a sequential detection procedure for change points in the correlation structure of financial returns. As far as we know, this is the first procedure for solving such a problem. The procedure is based on a CUSUM test statistic proposed by Wied et al. (2011). The asymptotic distribution of the test coincides with the one of the supremum of the absolute value of a Brownian Bridge in the interval $[0, 1]$. We have shown that the proposed procedure is consistent to detect the true number of location of the change points. Also, the finite sample properties of the procedure have been analyzed by the analysis of several simulation studies and the application of the procedure to a real data example. The empirical findings in the real data example suggest that the procedure detects changes in situations in which the relationship between financial returns may change due to financial crisis.

References

Andreou, E. and E. Ghysels (2002): "Detecting multiple breaks in financial market volatility dynamics," *Journal of Applied Econometrics*, 17, 579–600.

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

Bai, J. (1997): "Estimating multiple breaks one at a time," *Econometric Theory*, 13, 315–352.

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

Billingsley, P. (1968): *Convergence of probability measures*, Wiley, New York.

Campbell, R., C. Forbes, K. Koedijk, and P. Kofman (2008): "Increasing correlations or just fat tails?" *Journal of Empirical Finance*, 15, 287–309.

Galeano, P. and R. Tsay (2010): "Shifts in individual parameters of a GARCH model," *Journal of Financial Econometrics*, 8, 122–153.

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

Krishan, C., R. Petkova, and P. Ritchken (2009): "Correlation risk," *Journal of Empirical Finance*, 16, 353–367.

Longin, F. and B. Solnik (1995): "Is the correlation in international equity returns constant: 1960-1990?" *International Money and Finance*, 14(1), 3–26.

Ploberger, W., W. Krämer, and K. Kontrus (1989): "A new test for structural stability in the linear regression model," *Journal of Econometrics*, 40, 307–318.

Vostrikova, L. (1981): "Detecting 'disorder' in multidimensional random processes," *Soviet Mathematics Doklady*, 24, 55–59.

Wied, D., W. Krämer, and H. Dehling (2011): "Testing for a change in correlation at an unknown point in time using an extended functional delta method," *appears in Econometric Theory.*

7. Appendix section

Proof of Theorem 1

Obtaining the maximum of $B_T(z)$ is equivalent to obtaining the maximum of $|A_T(z)|$ with,

$$A_T(z) := \hat{D}\frac{\tau(z)}{T}\left(\hat{\rho}_{\tau(z)} - \hat{\rho}_T\right).$$

We first show that $A_T(z)$ converges in distribution to,

$$A(z) := C_A\left(\int_{l_1}^{z} g(t)dt - z\int_{l_1}^{l_2} g(t)dt\right)$$

uniformly in $z \in [l_1, l_2]$ with a constant C_A. For this purpose, write,

$$A_T(z) := \hat{D}\frac{\tau(z)}{T}\left(\hat{\rho}_{\tau(z)} - \rho_0\right) - \hat{D}\frac{\tau(z)}{T}\left(\hat{\rho}_T - \rho_0\right)$$

with $\rho_0 = \frac{m_{xy} - \mu_x\mu_y}{\sigma_x\sigma_y}$ and consider first the case $l_1 > 0$. We thus have,

$$A_T(z) := \hat{D}\frac{\tau(z)}{T}\left(\frac{\overline{XY}_{\tau(z)} - \overline{X}_{\tau(z)}\overline{Y}_{\tau(z)}}{\sqrt{\overline{[\mathsf{Var}X]}_{\tau(z)}\overline{[\mathsf{Var}Y]}_{\tau(z)}}} - \frac{m_{xy} - \mu_x\mu_y}{\sigma_x\sigma_y}\right) - \hat{D}\frac{\tau(z)}{T}\left(\frac{\overline{XY}_T - \overline{X}_T\overline{Y}_T}{\sqrt{\overline{[\mathsf{Var}X]}_T\overline{[\mathsf{Var}Y]}_T}} - \frac{m_{xy} - \mu_x\mu_y}{\sigma_x\sigma_y}\right)$$

Straightforward calculations using the strong law of large numbers, Slutzky's theorem, the fact that,

$$\sup_{z\in[l_1,l_2]}\left|\frac{1}{T}\sum_{t=1}^{\tau(z)}(X_tY_t - m_{xy}) - \int_{l_1}^{z} g(t)dt\right| = \sup_{z\in[l_1,l_2]}\left|\frac{1}{T}\sum_{t=1}^{\tau(z)} g\left(\frac{t}{T}\right) - \int_{l_1}^{z} g(t)dt\right| \to 0$$

and the fact that $\sup_{z\in[l_1,l_2]}\tau(z) \to \infty$ yield,

$$A_T(z) \to_{a.s.} A(z)$$

and,

$$|A_T(z)| \to_{a.s.} |A(z)|$$

uniformly on $[l_1, l_2]$.

Consider now the case $l_1 = 0$. By the preceding calculations, we immediately get,

$$A_T(z) \to_{a.s.} A(z)$$

uniformly on $[\epsilon, l_2]$ for a fixed $\epsilon > 0$ with $\epsilon < l_2$. Consider now the following functions:

$$A_T^{\epsilon}(z) = \begin{cases} A_T(z), & z \geq \epsilon \\ 0 & z < \epsilon \end{cases},$$

$$A^{\epsilon}(z) = \begin{cases} A(z), & z \geq \epsilon \\ 0 & z < \epsilon \end{cases}.$$

The previous results then imply that,

$$A_T^{\epsilon}(\cdot) \to_d A^{\epsilon}(\cdot)$$

for $T \to \infty$ on $[\epsilon, l_2]$ and also

$$A^{\epsilon}(\cdot) \to_d A(\cdot)$$

for rational $\epsilon \to 0$. The convergence of $A_T(\cdot)$ on $[0, l_2]$ then follows from Theorem 4.2 in Billingsley (1968) if we can show that

$$\lim_{\epsilon \to 0} \limsup_{T \to \infty} \mathbb{P}(\sup_{z \in [0, l_2]} |A_T^{\epsilon}(z) - A_T(z)| \geq \eta) = \lim_{\epsilon \to 0} \limsup_{T \to \infty} \mathbb{P}(\sup_{z \in [0, \epsilon]} |A_T(z)| \geq \eta) = 0$$

for all $\eta > 0$. Note that the separability condition of this theorem is not necessary in our case, because for each interval $I \subset [0, 1]$, $\sup_{z \in I} |A(z)|$ is always a random variable when $A(\cdot)$ is a right-continuous random function.

Since

$$\limsup_{T \to \infty} \mathbb{P}(\sup_{z \in [0, \epsilon]} |A_T(z)| \geq \eta) \leq \limsup_{T \to \infty} \mathbb{P}(\sup_{z \in [0, \epsilon]} \left| \hat{D} 2 \frac{\tau(z)}{T} \right| \geq \eta) = \mathbb{P}(\sup_{z \in [0, \epsilon]} D_A 2\epsilon \geq \eta)$$

we get

$$A_T(z) \to_{a.s.} A(z)$$

and

$$|A_T(z)| \rightarrow_{a.s.} |A(z)|$$

uniformly on $[0, l_2]$.

With Assumptions 2 and 3, $|A(z)|$ has a unique maximum m in the change point fraction. Let $\hat{F}$ the maximum of $|A_T(z)|$ for $z \in [0,1]$. Since $|A_T(\hat{F})| \geq |A_T(m)|$ we get stochastic convergence of $\hat{F}$ to m (compare the argument in Bai and Perron, 1998, p.77).

■

Proof of Theorem 2

With Theorem 1, we get

$$A_T(z) \rightarrow_{a.s.} A(z)$$

under the alternative. Since

$$Q_T(X,Y) = \sqrt{T} \sup_{z \in [0,1]} |A_T(z)|,$$

we have

$$\frac{1}{a_T} Q_T(X,Y) \rightarrow_p \infty \tag{3}$$

for $a_T = o\left(\sqrt{T}\right)$ under the alternative. With this argument (which is partially similar to Corollary 2 in Andrews, 1993), one can adapt the proof of Proposition 11 of Bai (1997).

Consider the event $\{\hat{b} < b\}$. If the estimated number of change points $\hat{b}$ is smaller than b, there is at least one segment $[\hat{s}_k, \hat{s}_l]$ with $\hat{s}_k \rightarrow_p s_k$ and $\hat{s}_l \rightarrow_p s_l$ such that there is another change point $s_m \in [s_k, s_l]$. Since $\mathsf{P}(Q_T(X,Y) > a_T) \rightarrow 1$ as $T \rightarrow \infty$ with (3), we have $\mathsf{P}(\hat{b} < b) \rightarrow 0$ as $T \rightarrow \infty$. Consider the event $\{\hat{b} > b\}$. For this event to be true, there must be a false rejection of the null hypothesis at a certain stage in the sequential estimation. If $(s_i, i = 0, \ldots, b)$ are the true change points and $(\hat{s}_i, i = 0, \ldots, b)$ are the corresponding consistent estimates, it holds

$$\begin{aligned}\mathsf{P}(\hat{b} > b) &\leq \mathsf{P}(\exists i: \text{ the test based on data for } \tau(z) \text{ with } z \in [\hat{s}_i, \hat{s}_{i+1}] \text{ rejects}) \\ &\leq \sum_{i=0}^{b} \mathsf{P}(\text{ the test based on data for } \tau(z) \text{ with } z \in [\hat{s}_i, \hat{s}_{i+1}] \text{ rejects}).\end{aligned}$$

Since under the null hypothesis $\mathsf{P}(Q_T(X,Y) > a_T) \to 0$, it holds with the test statistic computed for $\tau(z)$ with $z \in [\hat{s}_i, \hat{s}_{i+1}]$, $Q_T^i(X,Y)$,

$$\mathsf{P}(\hat{b} > b) \leq (b+1) \max_{0 \leq i \leq b} \mathsf{P}(Q_T^i(X,Y) > a_T) \to 0.$$

Consequently, $\mathsf{P}(\hat{b} => b) \to 1$ for $T \to \infty$.

Combining the argumentation for the event $\{\hat{b} < b\}$ with Theorem 1 yields the proposed consistency results and the proof is completed. ■

Figure 1: Function $A^*(z)$ in example (1) for $z \in [0, 1]$

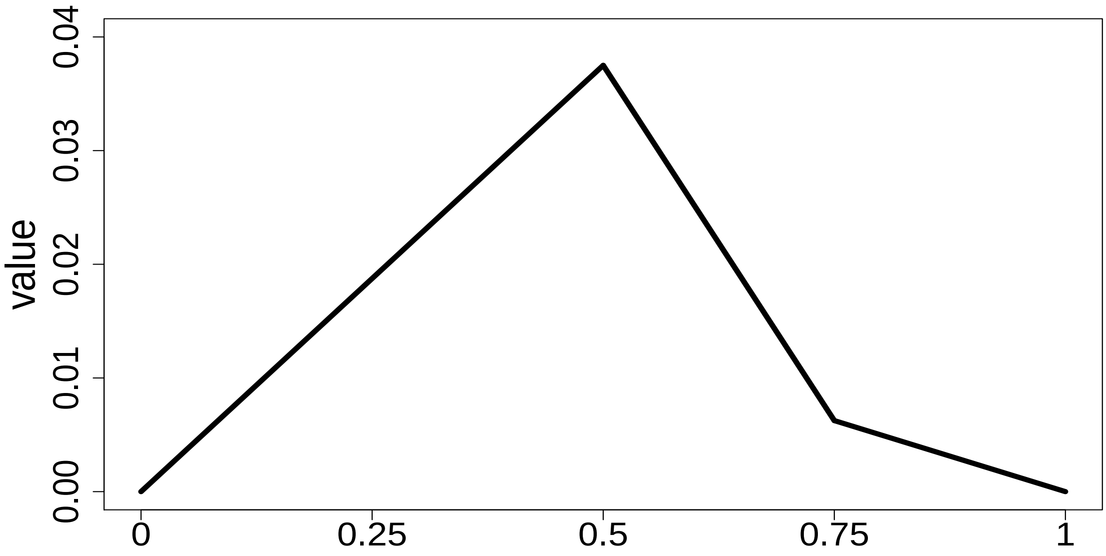

Table 1: Results for type I errors.

	$\phi = 0$						$\phi = 0.5$					
	$\rho_0 = -.5$		$\rho_0 = 0$		$\rho_0 = .5$		$\rho_0 = -.5$		$\rho_0 = 0$		$\rho_0 = .5$	
	Rel. freq.		Rel. freq.		Rel. freq.		Rel. freq.		Rel. freq.		Rel. freq.	
T	0	≥ 1	0	≥ 1	0	≥ 1	0	≥ 1	0	≥ 1	0	≥ 1
200	.961	.039	.966	.034	.946	.054	.934	.066	.930	.070	.928	.072
500	.961	.039	.968	.032	.968	.032	.936	.064	.952	.048	.930	.070
1000	.960	.040	.970	.030	.957	.043	.942	.058	.942	.058	.947	.053
2000	.961	.039	.964	.036	.963	.037	.947	.053	.960	.040	.951	.049
3000	.961	.039	.965	.035	.968	.032	.943	.057	.952	.048	.949	.051

Figure 2: Function $A^*(z)$ in example (2) for $z \in [0, 1]$

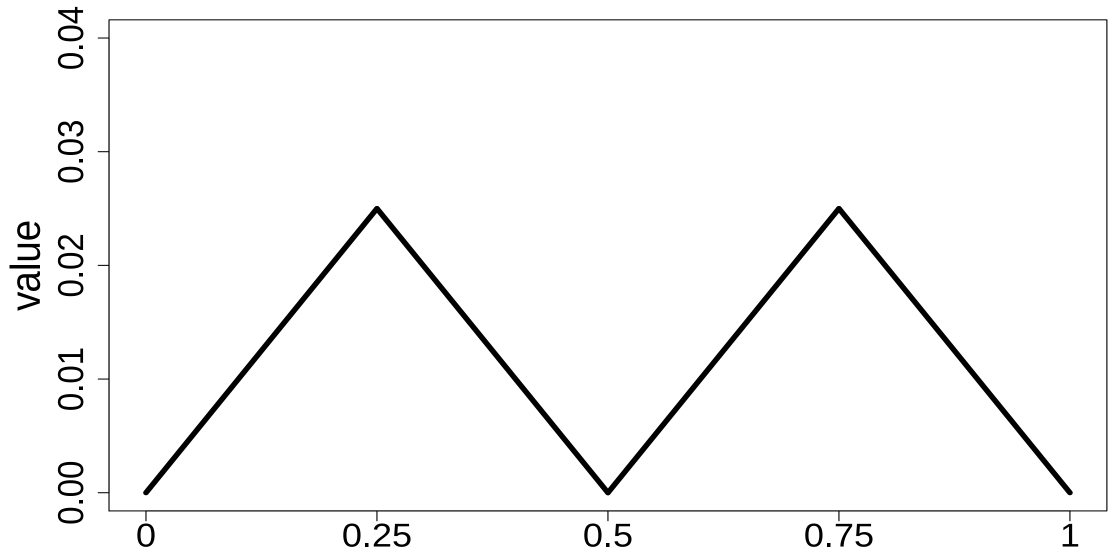

Figure 3: Log-retuns of S&P 500 and IBM indexes

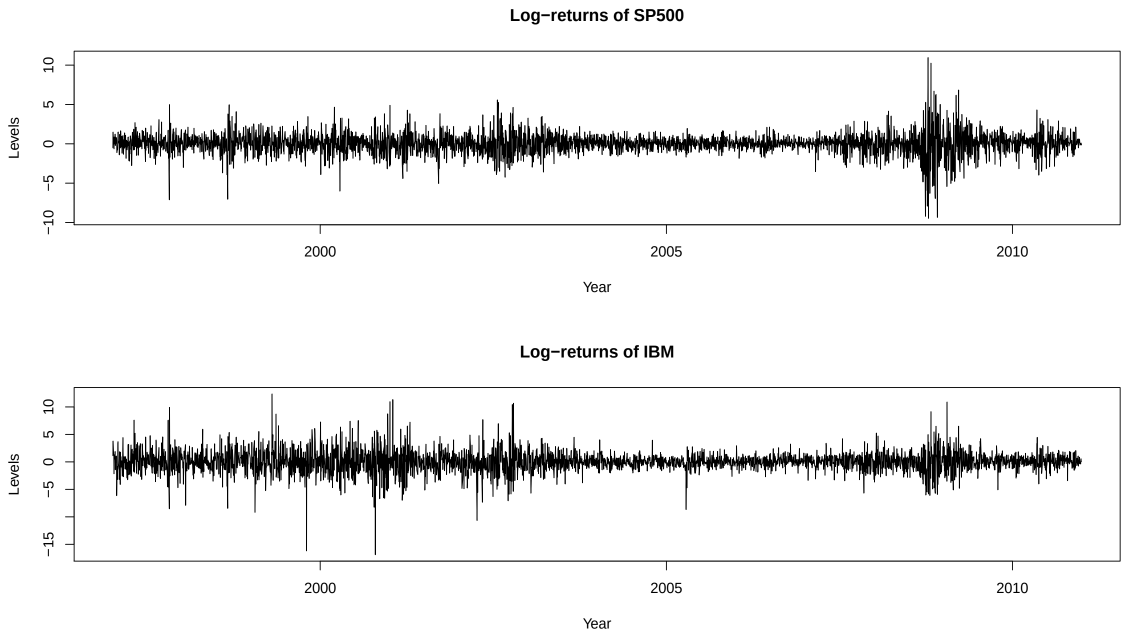

Figure 4: First step of the procedure

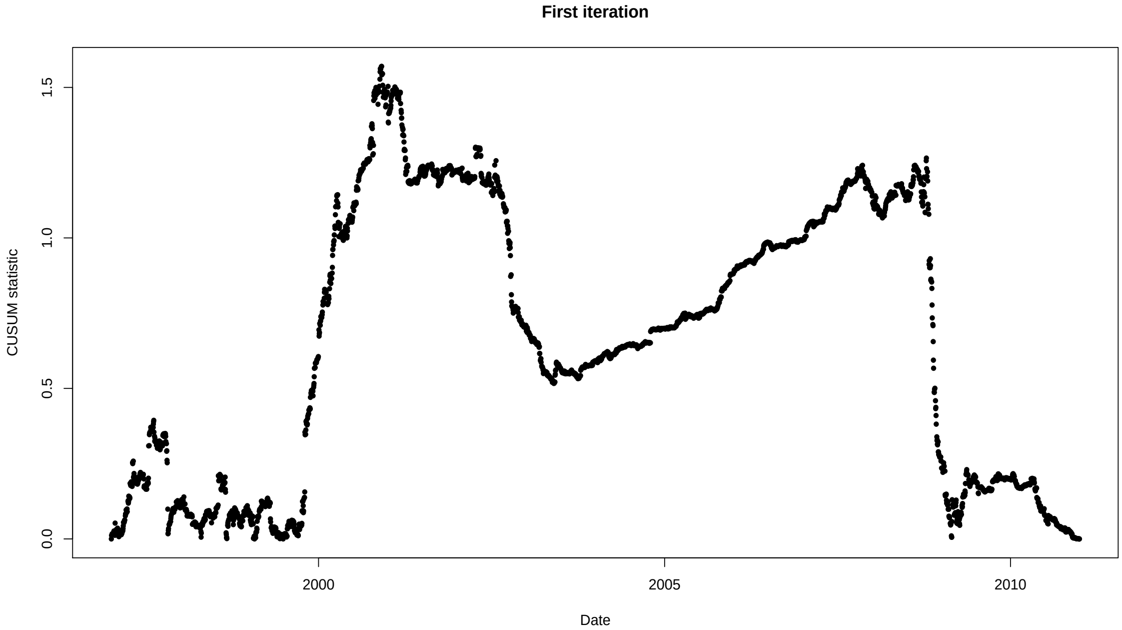

Figure 5: Second step of the procedure (first iteration)

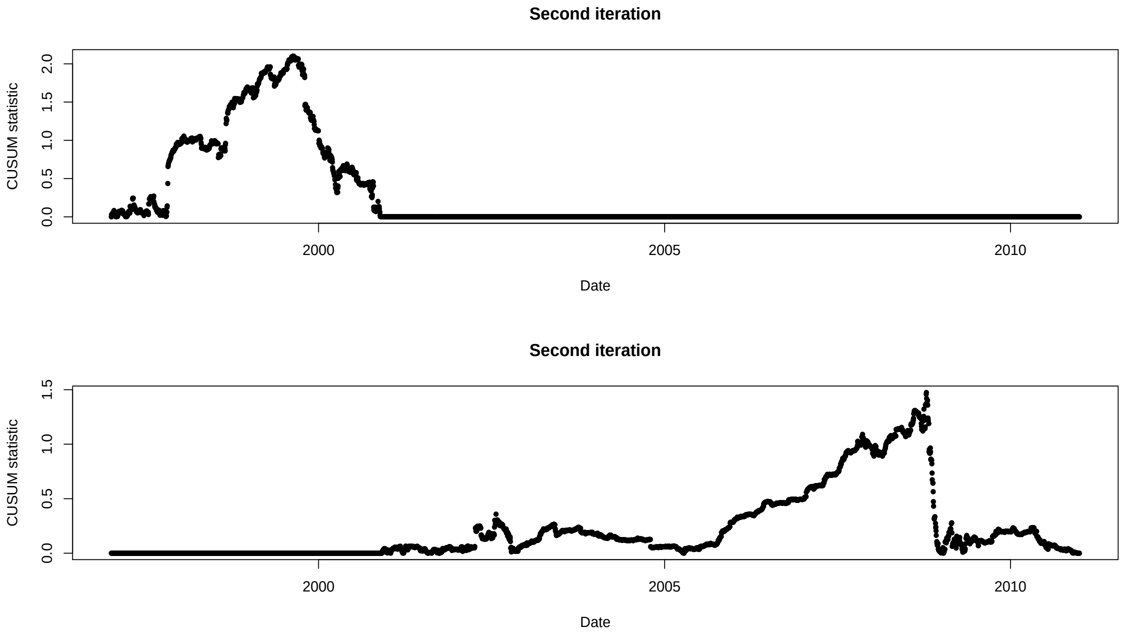

Figure 6: Second step of the procedure (second iteration)

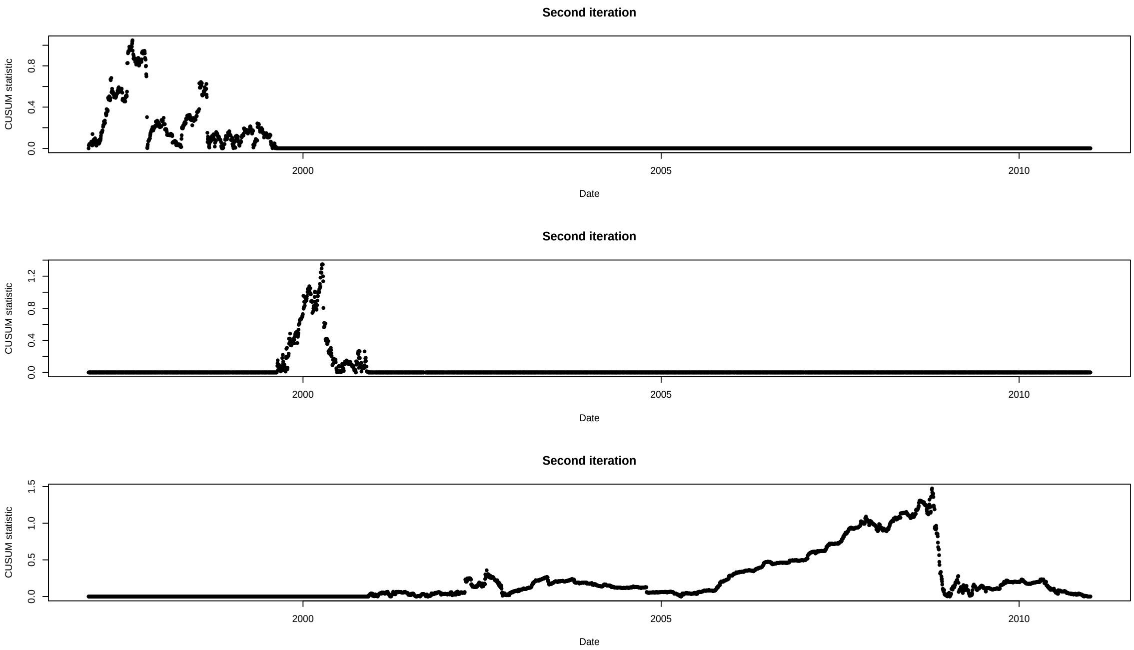

Figure 7: Third step of the algorithm

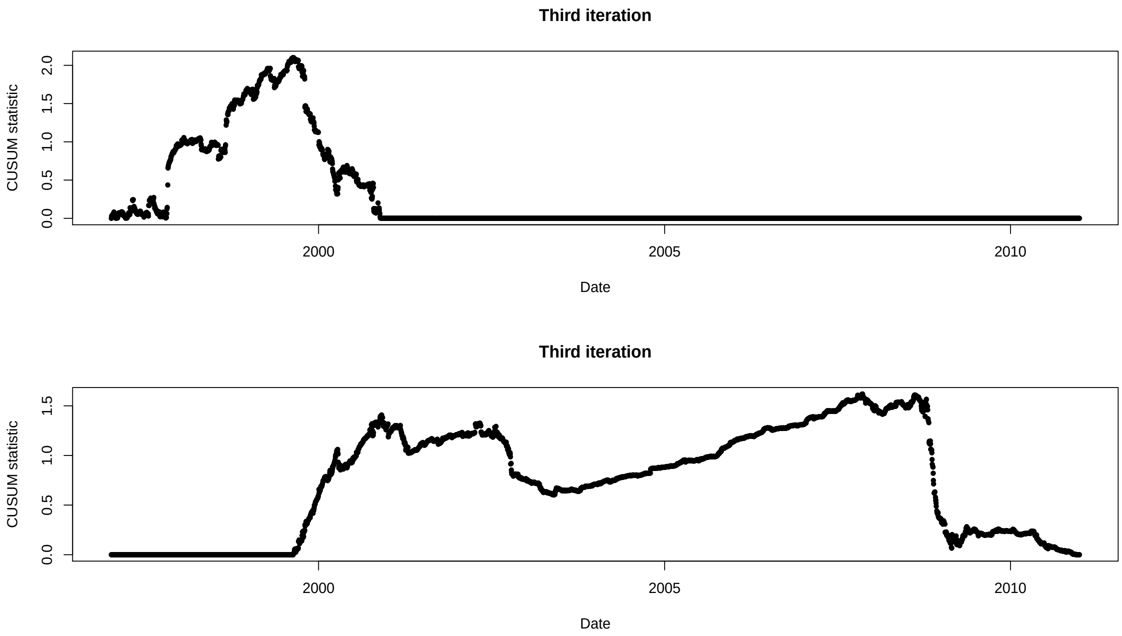

Figure 8: Scatterplots of the two indexes at three different subperiods

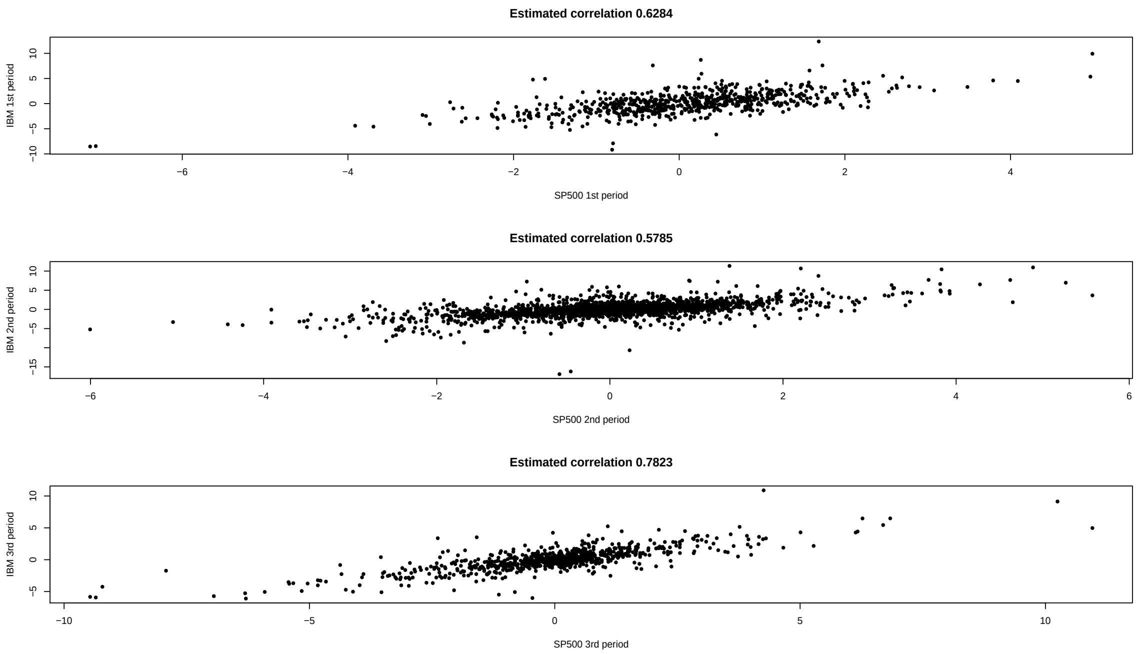

Table 2: Results for one change point.

			$\phi = 0$									$\phi = 0.5$								
			$z_1 = 1/4$			$z_1 = 1/2$			$z_1 = 3/4$			$z_1 = 1/4$			$z_1 = 1/2$			$z_1 = 3/4$		
			Rel. freq.			Rel. freq.			Rel. freq.			Rel. freq.			Rel. freq.			Rel. freq.		
ρ_0	ρ_1	T	0	1	≥ 2	0	1	≥ 2	0	1	≥ 2	0	1	≥ 2	0	1	≥ 2	0	1	≥ 2
		200	.709	.274	.017	.431	.536	.033	.660	.312	.028	.742	.244	.014	.483	.475	.042	.689	.290	.021
		500	.262	.720	.018	.072	.886	.042	.205	.746	.049	.360	.623	.017	.090	.868	.042	.281	.671	.048
.5	0	1000	.046	.928	.026	.007	.960	.033	.040	.914	.046	.071	.876	.053	.014	.938	.048	.055	.886	.059
		2000	.008	.947	.045	.002	.963	.035	.007	.937	.056	.012	.940	.048	.001	.952	.047	.005	.933	.062
		3000	.002	.957	.041	.001	.964	.035	.002	.945	.053	.005	.934	.061	.001	.956	.043	.004	.942	.054
		200	.926	.071	.003	.841	.147	.012	.910	.082	.008	.922	.065	.013	.844	.129	.027	.875	.099	.026
		500	.815	.175	.010	.583	.399	.018	.735	.248	.017	.832	.152	.016	.642	.336	.022	.776	.200	.024
.5	.25	1000	.539	.445	.016	.234	.731	.035	.454	.521	.025	.607	.375	.018	.312	.637	.051	.532	.437	.031
		2000	.206	.768	.026	.053	.918	.029	.176	.785	.039	.295	.657	.048	.071	.882	.047	.223	.725	.052
		3000	.079	.890	.031	.015	.952	.033	.061	.892	.047	.122	.830	.048	.017	.936	.047	.082	.872	.046
		200	.580	.408	.012	.554	.431	.015	.747	.253	.000	.608	.350	.036	.444	.504	.049	.757	.236	.007
		500	.326	.666	.008	.235	.752	.013	.568	.427	.005	.378	.598	.024	.308	.671	.021	.591	.403	.003
.5	.75	1000	.120	.861	.019	.056	.918	.026	.277	.708	.015	.163	.805	.032	.080	.884	.036	.314	.665	.021
		2000	.014	.946	.040	.016	.954	.030	.049	.922	.029	.033	.931	.036	.006	.949	.045	.086	.887	.027
		3000	.005	.945	.050	.005	.965	.030	.016	.936	.048	.011	.946	.043	.003	.965	.032	.021	.922	.057

Table 3: Estimation of the change point.

			$\phi = 0$			$\phi = 0.5$		
			$z_1 = 1/4$	$z_1 = 1/2$	$z_1 = 3/4$	$z_1 = 1/4$	$z_1 = 1/2$	$z_1 = 3/4$
ρ_0	ρ_1	T	$\widehat{z}_1$	$\widehat{z}_1$	$\widehat{z}_1$	$\widehat{z}_1$	$\widehat{z}_1$	$\widehat{z}_1$
		200	.355 .080	.520 .030	.730 .035	.405 .090	.525 .035	.735 .035
		500	.300 .046	.508 .018	.742 .022	.312 .052	.510 .018	.740 .028
.5	0	1000	.272 .022	.504 .009	.746 .010	.279 .028	.506 .012	.744 .014
		2000	.263 .013	.501 .004	.747 .005	.268 .017	.503 .006	.747 .008
		3000	.257 .007	.501 .003	.748 .003	.259 .009	.502 .004	.748 .005
		200	.485 .085	.540 .060	.707 .062	.520 .115	.540 .085	.670 .115
		500	.388 .094	.514 .034	.719 .047	.406 .093	.520 .040	.720 .048
.5	.25	1000	.322 .066	.506 .023	.728 .035	.342 .080	.507 .027	.721 .042
		2000	.283 .033	.502 .014	.737 .023	.295 .044	.503 .017	.733 .026
		3000	.271 .022	.502 .010	.739 .016	.278 .028	.503 .013	.739 .018
		200	.290 .075	.485 .060	.605 .125	.300 .090	.457 .080	.535 .085
		500	.276 .040	.498 .030	.690 .062	.287 .057	.498 .042	.652 .096
.5	.75	1000	.267 .027	.500 .021	.718 .034	.273 .035	.498 .026	.697 .054
		2000	.262 .015	.499 .011	.736 .016	.262 .019	.499 .014	.726 .026
		3000	.258 .012	.499 .007	.742 .010	.259 .015	.499 .009	.738 .014

Table 4: Results for two change points.

				$\phi = 0$				$\phi = 0.5$			
				$(z_1, z_2) = (1/3, 2/3)$				$(z_1, z_2) = (1/3, 2/3)$			
				Rel. freq.				Rel. freq.			
ρ_0	ρ_1	ρ_2	T	0	1	2	≥ 3	0	1	2	≥ 3
			200	.852	.134	.014	.000	.843	.131	.025	.001
			500	.458	.422	.117	.003	.529	.359	.107	.005
.5	0	.25	1000	.130	.525	.339	.006	.169	.507	.306	.018
			2000	.018	.271	.702	.009	.017	.337	.627	.019
			3000	.003	.133	.844	.020	.007	.205	.761	.027
			200	.622	.352	.026	.000	.660	.286	.052	.002
			500	.240	.554	.199	.007	.314	.523	.147	.016
.5	.25	.75	1000	.042	.456	.490	.012	.074	.500	.404	.022
			2000	.006	.159	.802	.033	.012	.217	.734	.037
			3000	.004	.062	.891	.043	.002	.091	.850	.057
			200	.699	.223	.078	.000	.711	.204	.080	.005
			500	.253	.313	.424	.010	.345	.328	.318	.009
.5	.75	.25	1000	.061	.170	.753	.016	.089	.210	.670	.031
			2000	.010	.036	.932	.022	.007	.047	.903	.043
			3000	.001	.012	.956	.031	.005	.011	.940	.044

Table 5: Estimation of two change points.

				$\phi = 0$	$\phi = 0.5$
				$(z_1, z_2) = (1/3, 2/3)$	$(z_1, z_2) = (1/3, 2/3)$
ρ_0	ρ_1	ρ_2	T	$(\widehat{z}_1, \widehat{z}_2)$	$(\widehat{z}_1, \widehat{z}_2)$
			200	$\left(\underset{.037}{.355}, \underset{.260}{.652}\right)$	$\left(\underset{.035}{.337}, \underset{.075}{.635}\right)$
			500	$\left(\underset{.018}{.340}, \underset{.076}{.670}\right)$	$\left(\underset{.022}{.342}, \underset{.088}{.672}\right)$
.5	0	.25	1000	$\left(\underset{.008}{.339}, \underset{.052}{.666}\right)$	$\left(\underset{.009}{.339}, \underset{.064}{.668}\right)$
			2000	$\left(\underset{.004}{.335}, \underset{.039}{.666}\right)$	$\left(\underset{.006}{.336}, \underset{.043}{.667}\right)$
			3000	$\left(\underset{.003}{.334}, \underset{.022}{.666}\right)$	$\left(\underset{.004}{.335}, \underset{.040}{.668}\right)$
			200	$\left(\underset{.032}{.362}, \underset{.050}{.662}\right)$	$\left(\underset{.035}{.365}, \underset{.070}{.67}\right)$
			500	$\left(\underset{.028}{.350}, \underset{.028}{.666}\right)$	$\left(\underset{.030}{.356}, \underset{.040}{.664}\right)$
.5	.25	.75	1000	$\left(\underset{.020}{.340}, \underset{.016}{.666}\right)$	$\left(\underset{.023}{.344}, \underset{.020}{.667}\right)$
			2000	$\left(\underset{.013}{.336}, \underset{.009}{.666}\right)$	$\left(\underset{.018}{.337}, \underset{.013}{.666}\right)$
			3000	$\left(\underset{.009}{.335}, \underset{.006}{.667}\right)$	$\left(\underset{.013}{.336}, \underset{.007}{.667}\right)$
			200	$\left(\underset{.057}{.307}, \underset{.030}{.68}\right)$	$\left(\underset{.045}{.320}, \underset{.040}{.685}\right)$
			500	$\left(\underset{.032}{.326}, \underset{.016}{.672}\right)$	$\left(\underset{.039}{.326}, \underset{.020}{.676}\right)$
.5	.75	.25	1000	$\left(\underset{.019}{.332}, \underset{.010}{.670}\right)$	$\left(\underset{.025}{.331}, \underset{.014}{.672}\right)$
			2000	$\left(\underset{.011}{.332}, \underset{.005}{.668}\right)$	$\left(\underset{.013}{.332}, \underset{.007}{.669}\right)$
			3000	$\left(\underset{.007}{.333}, \underset{.003}{.668}\right)$	$\left(\underset{.009}{.332}, \underset{.004}{.668}\right)$

Table 6: Iterations taken by the procedure in the real data example, (*) means significant change point.

Interval	$Q_T(X,Y)$	Change point	Time point	Date
Step 1				
Interval	$Q_T(X,Y)$	Change point	Time point	Date
$[1, 3524]$	1.5699 (*)	988	0.2803	November 29, 2000
Step 2				
Interval	$Q_T(X,Y)$	Change point	Time point	Date
$[1, 988]$	2.1009 (*)	664	0.1884	August 19, 1999
$[989, 3524]$	1.4744	2966	0.8416	October 14, 2008
$[1, 664]$	1.0482	157	0.0445	August 14, 1997
$[665, 988]$	1.3470	825	0.2341	April 7, 2000
$[989, 3524]$	1.4744	2966	0.8416	October 14, 2008
Step 3				
Interval	$Q_T(X,Y)$	Change point	Time point	Date
$[1, 988]$	2.1009 (*)	664	0.1884	August 19, 1999
$[665, 3524]$	1.6193 (*)	2734	0.7758	November 12, 2007