

Numerical Analysis of Volatility Change Point Estimators for Discretely Sampled Stochastic Differential Equations

STEFANO M. IACUS* – NAKAHIRO YOSHIDA†

In this paper, we review recent advances on change point estimation for the volatility component of stochastic differential equations under different discrete sampling schemes. We consider both ergodic and non-ergodic cases, and present a Monte Carlo study on the change point estimator to compare the three methods under different setups.

(J.E.L.: C58, C13).

1. Introduction

In finance, the volatility of the market or of the asset prices plays a crucial role in many aspects. For example, in option pricing, although the very basic Black and Scholes (1973) and Merton (1974) model assumes a constant volatility, it is well known that this assumption is unrealistic when one works with real financial data. This fact causes well-known effects like implied volatility and volatility smiles. Change point analysis was initially introduced in the framework of independent and identically distributed data (see, e.g., Hinkley, 1971; Inlan and Tiao, 1994; Bai, 1994, 1997; Csörgö and Horváth, 1997), and quickly applied to the analysis of time series (see, e.g., Kim *et al.*, 2000; Lee *et al.*, 2003; Chen *et al.*, 2005). Kutoyants (1994, 2004) and Lee *et al.* (2006) studied structural change point problems for the drift term for continuous observations from ergodic diffusion processes. Due to the fact that volatility can be estimated without error in continuous time, the change point analysis in this setup is not very interesting.

*Corresponding author: Stefano M. Iacus, Department of Economics, Business and Statistics, University of Milan, Via Conservatorio 7, 20122 Milan, Italy. E-mail: stefano.iacus@unimi.it

†University of Tokyo, and Japan Science and Technology Agency, Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan

We thank two anonymous referees and the two Editors for the valuable comments to improve the paper.

© 2010 The Authors

Economic Notes © 2010 Banca Monte dei Paschi di Siena SpA. Published by Blackwell Publishing Ltd, 9600 Garsington Road, Oxford, OX4 2DQ, UK and 350 Main Street, Malden, MA 02148, USA.

In this paper, we review recent advances for the problem of time-change point for the volatility component of a process satisfying a stochastic differential equation based on discrete observations. We consider the case of both ergodic and non-ergodic processes under different sampling schemes and compare the different approaches via Monte Carlo experiments. This review is based on the recent works of De Gregorio and Iacus (2008), Song and Lee (2009) and Iacus and Yoshida (2009). To the authors' knowledge the above references represent the state of the art on change point estimation for the volatility term of continuous time processes solutions to stochastic differential equations from discrete time observations.

The paper is organized as follows. Section 2 reviews the least squares approach for diffusions as in De Gregorio and Iacus (2008); Section 3 reviews the results based on the CUSUM approach as in Song and Lee (2009); Section 4 reviews the quasi-maximum likelihood approach from Iacus and Yoshida (2009). All the proofs and technical conditions are contained in the original papers of the respective authors. Section 5 concludes the paper with a numerical analysis on the performance of the different methods to identify the change point.

2. Least Squares Approach

This section reviews the results of De Gregorio and Iacus (2008) without proofs. We denote by $X = \{X_t, 0 \leq t \leq T\}$ the diffusion process, with state space $\mathcal{X} = (l, r)$, $-\infty \leq l \leq r \leq +\infty$, such that

$$(1) \quad X_t = \begin{cases} X_0 + \int_0^t b(X_s) ds + \int_0^t \sqrt{\theta_1} \sigma(X_s) dW_s, & 0 \leq t \leq \tau^* \\ X_{\tau^*} + \int_{\tau^*}^t b(X_s) ds + \int_{\tau^*}^t \sqrt{\theta_2} \sigma(X_s) dW_s, & \tau^* < t \leq T \end{cases}$$

with $X_0 = x_0$, $0 < \theta_1, \theta_2 < \infty$ and $\{W_t, t \geq 0\}$ a standard Brownian motion. The value $\tau^* \in (0, T)$ is the change point instant. The parameters θ_1 and θ_2 belong to Θ , a compact set of \mathbb{R}^+ . The coefficients $b : \mathcal{X} \rightarrow \mathbb{R}$ and $\sigma : \mathcal{X} \rightarrow (0, \infty)$ are supposed to be known, continuous with continuous derivatives and regular so that (1) is well defined and the process X is unique and such that the process possesses the ergodic property. Let $s(x) = \exp\{-\int_{x_*}^x 2b(u)/\sigma^2(u)du\}$ be the scale function (where x_* is an arbitrary point inside \mathcal{X}). The following condition will be required throughout this section:

$$A1. \quad \lim_{\substack{x_1 \rightarrow l \\ x_2 < r}} \int_{x_1}^x s(u)du = +\infty, \quad \lim_{\substack{x_2 \rightarrow r \\ x < x_1}} \int_x^{x_2} s(u)du = +\infty, \quad \text{where } l < x_1 < x < x_2 < r$$

Condition A1 guarantees that the exit time from \mathcal{X} is infinite (see Karatzas and Shreve, 1991). In Section 2.2, we will also consider the case where $b(\cdot)$ is unknown and estimated non-parametrically.

The process X is observed at $n + 1$ equidistant discrete times $0 = t_0 < t_1 < \dots < t_n = T$, with $t_i = i \Delta_n$. For the sake of simplicity we assume $T = 1$ and with little abuse of notation, we will write X_i instead of X_{t_i} and W_i instead of W_{t_i} . The asymptotic framework is an high frequency scheme: $n \rightarrow \infty, \Delta_n \rightarrow 0$ with $n \Delta_n = T$ with T fixed. Given the observations $X_i, i = 0, 1, \dots, n$, the aim of this work is to estimate the change time τ^* as well as the two parameters θ_1, θ_2 .

In order to apply the least-squares approach of Bai (1994), the standardized residuals are constructed using Euler’s approximation, that is,

$$Z_i = \frac{X_{i+1} - X_i - b(X_i)\Delta_n}{\sqrt{\Delta_n}\sigma(X_i)}, \quad i = 1, \dots, n$$

We denote by $k_0 = [n\tau^*]$ and $k = [n\tau], \tau, \tau^* \in (0, 1)$, where $[x]$ is the integer part of the real value x . The least squares estimator of the change point is given by

$$\begin{aligned} \hat{k}_0 &= \arg \min_k \left(\min_{\theta_1, \theta_2} \left\{ \sum_{i=1}^k (Z_i^2 - \theta_1)^2 + \sum_{i=k+1}^n (Z_i^2 - \theta_2)^2 \right\} \right) \\ (2) \quad &= \arg \min_k \left\{ \sum_{i=1}^k (Z_i^2 - \bar{\theta}_1)^2 + \sum_{i=k+1}^n (Z_i^2 - \bar{\theta}_2)^2 \right\} \end{aligned}$$

where

$$\bar{\theta}_1 = \frac{1}{k} \sum_{i=1}^k Z_i^2 =: \frac{S_k}{k} \quad \text{and} \quad \bar{\theta}_2 = \frac{1}{n-k} \sum_{i=k+1}^n Z_i^2 =: \frac{S_{n-k}^*}{n-k}$$

It is easy to show that the problem (2) is equivalent to the following

$$(3) \quad \hat{k}_0 = \arg \max_k |D_k|$$

where

$$D_k = \frac{k}{n} - \frac{S_k}{S_n}$$

Once \hat{k}_0 has been obtained, the following estimator of the parameters θ_1 and θ_2 can be used

$$(4) \quad \hat{\theta}_1 = \frac{S_{\hat{k}_0}}{\hat{k}_0}, \quad \hat{\theta}_2 = \frac{S_{n-\hat{k}_0}^*}{n - \hat{k}_0}$$

The quantity D_k also serves as a test statistics to verify the presence/absence of a change point in the usual way. Indeed, it is possible to show

that, under no change point, H_0 : ' $\theta_1 = \theta_2$ for all $\tau \in (0, 1)$ ', for any $\delta \in (0, 1/2)$, the test statistic $\sqrt{\frac{n}{2}} \sup_{\delta n \leq k \leq (1-\delta)n} |D_k|$ converges in distribution to $\sup_{\delta \leq \tau \leq (1-\delta)} |B_0(\tau)|$, where $\{B_0(\tau), 0 \leq \tau \leq 1\}$ is a Brownian bridge. The distribution of the test under the alternative hypothesis H_1 : ' $\theta_1 \neq \theta_2$ for one $\tau^* \in (0, 1)$ ', was not considered in the original paper. The motivation is that, even in the simple i.i.d. case, this distribution depends on the change point instant τ^* and the distance between the two values of θ . In general such results require additional regularity conditions on the model (see, e.g. section 1.5. in Csörgö and Horváth, 1997). The same issue arises for the CUSUM test statistic presented in Section 3. Nevertheless, one can easily see that the proposed test is consistent. In order to prove consistency, we just need to show that under H_1 , for at least one k the statistic $\sqrt{\frac{n}{2}} |D_k|$ diverges. Looking at D_k and the definitions of S_k and S_n we can see that

$$(5) \quad \sqrt{n} \left(\frac{k}{n} - \frac{S_k}{S_n} \right) = \sqrt{n} \frac{k}{n} \left[1 - \frac{1}{k} \sum_{i=1}^{[n\tau]} Z_i^2 \left(\frac{1}{n} \sum_{i=1}^n Z_i^2 \right)^{-1} \right] \\ \simeq \sqrt{n\tau} \frac{(1-\tau)(\theta_2 - \theta_1)}{\tau\theta_1 + (1-\tau)\theta_2}$$

for $k = [n\tau]$. Now, obviously, $\sup_{\delta n \leq k \leq (1-\delta)n} \sqrt{n} |D_k|$ diverges.

2.1. Contiguous Alternatives

Denote by $\mathcal{W}(v)$ the two-sided Brownian motion, that is,

$$(6) \quad \mathcal{W}(u) = \begin{cases} W_1(-u), & u < 0 \\ W_2(u), & u \geq 0 \end{cases}$$

where W_1, W_2 are two independent Brownian motions. Let $\vartheta_n = |\theta_2(n) - \theta_1(n)| \neq 0$ for finite n . Under the additional condition that $\vartheta_n \rightarrow 0$ and $\frac{\sqrt{n}\vartheta_n}{\sqrt{\log n}} \rightarrow \infty$, the change point estimator $\hat{\tau}_n^* = \hat{k}_0/n$ is consistent and such that

$$(7) \quad \frac{n\vartheta_n^2(\hat{\tau}_n^* - \tau^*)}{2\tilde{\theta}^2} \xrightarrow{d} \arg \max_v \left\{ \mathcal{W}(v) - \frac{|v|}{2} \right\}$$

for any consistent estimator $\tilde{\theta}$ for the common limiting value θ_0 of $\theta_1(n)$ and $\theta_2(n)$. The condition $\vartheta_n \rightarrow 0$ corresponds to the setup of contiguous alternatives (Roussas, 1972), that is, the two parameters $\theta_1 = \theta_1(n)$ and $\theta_2 = \theta_2(n)$ are allowed to be closer and closer as the sample size increases. Thus, in order to discriminate the two regimes a sufficiently large number n of observations (or rate of convergence ϑ_n) is required. In other words, ϑ_n is the smallest distance between θ_1 and θ_2 such that the two parameters

can be distinguished from a sample of at least n observations. Under the conditions above, the estimators $\hat{\theta}_1, \hat{\theta}_2$ are \sqrt{n} -consistent and such that

$$(8) \quad \sqrt{n} \begin{pmatrix} \hat{\theta}_1 - \theta_1 \\ \hat{\theta}_2 - \theta_2 \end{pmatrix} \xrightarrow{d} N(0, \Sigma), \quad \text{where} \quad \Sigma = \begin{pmatrix} 2\frac{\theta_0^2}{\tau^*} & 0 \\ 0 & 2\frac{\theta_0^2}{1 - \tau^*} \end{pmatrix}$$

The above results hold in the high frequency case $\Delta_n \rightarrow 0$ with $n \rightarrow \infty$ and $n\Delta = T$ fixed, but also in the case where $n\Delta_n = T \rightarrow \infty$ with rapidly increasing design satisfying $n\Delta_n^2 \rightarrow 0$.

The distribution of the change point estimator \hat{k}_0 in case of non-contiguous alternatives has not been obtained in the original paper while consistency still holds. The estimators $\hat{\theta}_k, k = 1, 2$ have the same properties as in (8) with limiting variance-covariance matrix with θ_0 replaced by $\theta_k, k = 1, 2$. The same considerations apply to the results in the next section.

2.2. Change Point Estimation with Non-Parametric Estimation of Drift

We assume now to observe a diffusion process that is a solution to the reduced stochastic differential equation

$$dX_t = b(X_t)dt + \sqrt{\theta}dW_t$$

where $b(\cdot)$ is unknown and estimated using non-parametric methods. Let $K \geq 0$ be a kernel function, that is, K is symmetric and continuously differentiable, with $\int_{\mathbb{R}} uK(u)du = 0, \int_{\mathbb{R}} K^2(u)du < \infty$ and such that $\int_{\mathbb{R}} K(u)du = 1$. We consider the new standardized residuals

$$(9) \quad \hat{Z}_i = \frac{X_{i+1} - X_i - \hat{b}(X_i)\Delta_n}{\sqrt{\Delta_n}}$$

where

$$(10) \quad \hat{b}(x) = \frac{\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right) \frac{X_{i+1} - X_i}{\Delta_n}}{\sum_{i=1}^n K\left(\frac{X_i - x}{h_n}\right)}$$

is a non-parametric estimator of the drift constructed using the full sample and h_n is the bandwidth defined as in Silverman (1986). The estimator (10) used in De Gregorio and Iacus (2008) is a particular case of the non-parametric estimator for the drift in Bandi and Phillips (2003) although similar non-parametric estimators can be found in the earlier papers by Pham (1981), Florens-Zmirou (1993) or Stanton (1997). Unfortunately, for fixed T the drift coefficient cannot be estimated consistently (for a

recent account, see also, Wang *et al.*, 2010). So it is assumed that $n\Delta_n = T \rightarrow \infty$, $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$. In addition, some relationship between the bandwidth of the kernel and the mesh Δ_n should be required. Let

$$\bar{L}_X(t, x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} \int_0^t \mathbf{1}_{[x, x+\epsilon)}(X_s) ds$$

be the chronological local time of X . Bandi and Phillips (2003) have shown that under the following additional assumption

$$\frac{\bar{L}_X(T, x)}{h_n} \sqrt{\Delta_n \log \left(\frac{1}{\Delta_n} \right)} = o(1) \quad \text{a.s.}$$

and $h_n \bar{L}_X(T, x) \xrightarrow{\text{a.s.}} \infty$, $\hat{b}(\cdot)$ is a consistent estimator of $b(\cdot)$. In the stationary case, the above condition may be replaced by

$$\frac{T}{h_n} \sqrt{\Delta_n \log \left(\frac{1}{\Delta_n} \right)} = o(1) \quad \text{a.s.}$$

Now, substituting the new residuals \hat{Z}_i from (9) in place of the Z_i 's in (2), we obtain all the results in the previous section. This mixed result of parametric and non-parametric estimation is quite useful in applications because in practice there is no need to fully specify the data generating model for the observed data. The change point analysis for this reduced model identifies a change in the scale (or intensity) in the volatility levels. De Gregorio and Iacus (2008) discuss also about the choice of the bandwidth selection problem for h_n .

Under additional mild regularity conditions, it is also possible to obtain the same results as in the above in the case of a data generating model of this form $dX_t = b(X_t)dt + \sqrt{\theta}\sigma(X_t)dW_t$ and the diffusion coefficient $\sigma(x) > 0$ is known, although the original paper does not contain such result. Note that the point of the approach presented in this section is to use it in real applications without assuming too much information about the data generating model. This means that, such an instrument can be used to check where the scaling factor θ in the volatility component has changed. A drift may exist, but is considered as nuisance in this setup and, in this view, a specification of $\sigma(x)$ will be a little strange.

An extensive application of this method to the analysis of financial markets has been carried out in Smaldone (2009) both on daily and intra-daily data. In particular, the analysis of the global financial crisis of 2008 was reread via change point analysis using data from different markets. It emerged that the global structural change in 2008 could have been predicted through the analysis of some assets and market indexes.

Figure 1 reports a summary of the empirical analysis performed in Smaldone (2009) using the method presented in this section. This analysis

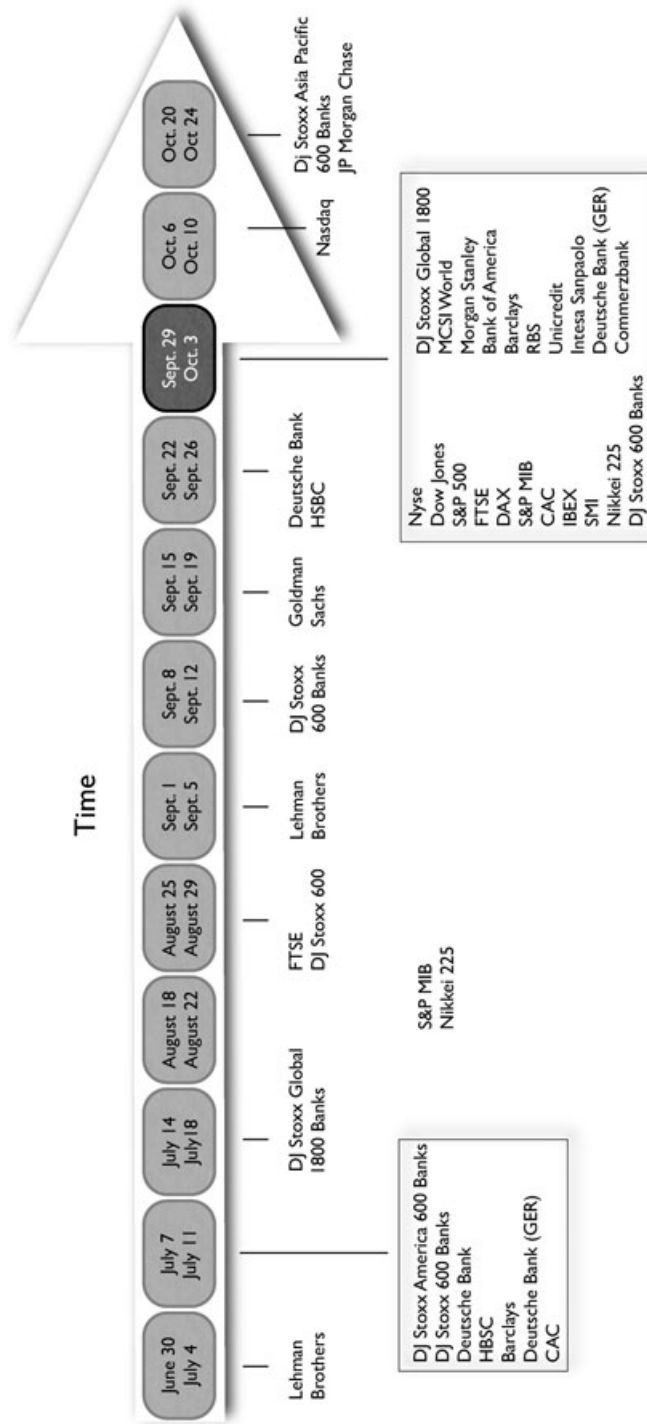


Figure 1: Summary of a simple structural change point analysis in the volatility of several financial market indexes and assets. Source Smaildone (2009).

involved the change point detection in the volatility of several market indexes including Nyse, Dow Jones, S&P 500 and Nasdaq; Dow Jones Stoxx 600, Nikkei, Dow Jones Global 1800, MSCI World, FTSE (UK), DAX (Germany), S&P/Mib (Italy), CAC (France), Ibex (Spain) and SMI (Switzerland). The analysis also considered several indexes for the bank compartment like Dow Jones Stoxx Global 1800 Banks (worldwide), Dow Jones Stoxx Americas 600 Banks (USA), Dow Jones Stoxx Asia-Pacific 600 Banks (Asia), Dow Jones Stoxx 600 Banks (Europe). Individual stock prices from the USA, UK, Italy, France, Germany, Spain and Japan stock exchanges were also considered. The evidence from this analysis is interesting as it shows how the crisis (measured here by uncertainty) affects the different markets at different dates (e.g., more protected markets, like the Italian case, reacts slowly to the crisis). Figure 1 reports only the largest volatility changes for some of the above time series. For more details, see Smaldone (2009).

3. The CUSUM Approach

Song and Lee (2009) considered the following data generating model

$$dX_t = b(X_t) dt + \sigma(X_t, \theta) dW_t, \quad 0 \leq t \leq T, \quad X_0 = x_0$$

under the sampling scheme $n\Delta_n = T \rightarrow \infty$, $\Delta_n \rightarrow 0$ as $n \rightarrow \infty$. The function $b(\cdot)$ and $\sigma(\cdot, \theta)$ are supposed to be known up to $\theta \in \Theta \subset R$. The object is to consider only the problem of hypotheses testing of a change point rather than estimating the change point itself. In particular, they considered the test $H_0 : \theta$ does not change in $[0, T]$ versus $H_1 : \theta_1 \neq \theta_2$ for one time instant τ^*T , $\tau^* \in (0, 1)$. So the problem is equivalent to the test based on the statistics D_k of the previous section, although in this setup the change point moves with T as τT . The process X is assumed to be ergodic and the model is supposed to satisfy the regularity conditions such that the minimum contrast estimation procedure described in Kessler (1997) is valid. Let θ_0 be the true value of θ under the null hypothesis H_0 . Consider the following CUSUM test statistics

$$(11) \quad T_n = \max_{1 \leq k \leq n} \frac{k^2}{n} (\hat{\theta}_{n,k} - \hat{\theta}_{n,n})^2 \hat{\mathcal{J}}_n$$

where $\hat{\theta}_{n,k}$ is the minimum contrast estimator based on the first k observations and $\hat{\theta}_{n,n}$ the one based on all the observations up to n . These estimators are obtained using the contrast function proposed in Kessler (1997). In this setup θ_1 and θ_2 are not contiguous alternatives. The quantity

$$\hat{\mathcal{J}}_n = \frac{2}{n} \sum_{i=1}^n \left\{ \frac{\partial_{\theta} \sigma(X_{i-1}, \hat{\theta}_{n,k})}{\sigma(X_{i-1}, \hat{\theta}_{n,k})} \right\}^2$$

is a consistent estimator of

$$\mathcal{J} = 2 \int \left(\frac{\partial_\theta \sigma(x, \theta_0)}{\sigma(x, \theta_0)} \right)^2 d\mu_0(x)$$

where μ_0 is the invariant law of the process X and \mathcal{J} is assumed to be positive. The first result concerns the estimation of θ . Indeed, suppose that $n\Delta_n^p \rightarrow 0$ and $n\Delta_n^q \rightarrow \infty$ for some $p > q > 1$. Then, under H_0 (no change point), $\max_{[n^{1/q}] \leq k \leq n} |\hat{\theta}_{n,k} - \theta_0| \rightarrow 0$ with probability 1, so consistency of estimators is established under H_0 . Further, under slightly stronger conditions, provided that $n\Delta_n^p \rightarrow 0$ and $n\Delta_n^q \rightarrow \infty$ for some $p > q > 4$, under H_0 ,

$$\mathcal{J}^{\frac{1}{2}} \frac{[ns]}{\sqrt{n}} (\hat{\theta}_{n,[ns]} - \theta_0) \xrightarrow{d} W_s \quad \text{in } \mathbb{D}[0, 1]$$

where W_s is a Wiener process. Therefore, $T_n \xrightarrow{d} \sup_{0 \leq s \leq 1} B_0^2(s)$.

As Song and Lee (2009) explained, the condition $q > 4$ is essential in order to get the asymptotic results, but this indicates a very sparse sampling and, at the same time, a large number of observations in order to have $n\Delta_n$ to diverge. In particular, they show that this method is not feasible for daily data.

Note that, for the reasons explained in the previous and next sections, the assumptions of a completely specified drift function is a bit unrealistic in applications. There is some possibility to apply a mixed parametric/non-parametric approach as in Section 2.2 although the original authors did not consider such situation.

As a final remark we note that the proof of the consistency of the CUSUM test statistics under the alternative hypothesis H_1 is much more involved due to misspecification. Indeed, in the presence of the change point, for $k < n\tau$, $\hat{\theta}_{n,k}$ is a consistent estimator of θ_1 while $\hat{\theta}_{n,n}$ cannot estimate θ_2 because it uses the complete set of observations from 1 to n . So a misspecification problem arises.

4. Quasi-maximum Likelihood Approach

De Gregorio and Iacus (2008) and Song and Lee (2009) paved the way for the volatility change point analysis for discretely observed diffusion processes. Both papers considered one-dimensional diffusion processes and both approaches have benefits and limits. For example, in the first paper the limit comes from the fact that the model is very simple though this is compensated by the non-parametric approach and by the fact that the regularity conditions are very mild. The second paper considers a more general setup, but the CUSUM test statistics does not allow study of the properties of a change point estimator because it is just a tool to detect

changes but does not provide an estimator of the change point. The sparsity of the data collection may also be a limit in some applications as well as the complete specification of the drift function.

Iacus and Yoshida (2009) considered a general model which includes as special cases the ones presented in the previous sections. Consider a d -dimensional Itô process described by the stochastic differential equation

$$(12) \quad dY_t = b_t dt + \sigma(X_t, \theta) dW_t, \quad t \in [0, T],$$

where W_t is an r -dimensional standard Wiener process, on a stochastic basis, b_t and X_t are vector valued progressively measurable processes and $\sigma(x, \theta)$ is a matrix valued function.

As in the previous sections, it is assumed that there is a time $\tau^* \in (0, T)$ at which the diffusion coefficient changes from $\sigma(x, \theta_1)$ to $\sigma(x, \theta_2)$. More precisely, (Y, X) satisfy the following stochastic integral equation

$$Y_t = \begin{cases} Y_0 + \int_0^t b_s ds + \int_0^t \sigma(X_s, \theta_1^*) dW_s & \text{for } t \in [0, \tau^*] \\ Y_{\tau^*} + \int_{\tau^*}^t b_s ds + \int_{\tau^*}^t \sigma(X_s, \theta_2^*) dW_s & \text{for } t \in [\tau^*, T] \end{cases}$$

The change point $\tau^* \in (0, T)$ is unknown and is to be estimated from the observations sampled from the path of (X, Y) . We denote by \mathcal{X} the state space of X . The coefficient $\sigma(x, \theta)$ is assumed to be known up to the parameter θ , while b_t is completely unknown and unobservable, therefore possibly depending on θ and τ^* . In this framework the interest is purely on τ^* and the estimation of θ is secondary. In the first part, it is assumed that consistent estimators exist for θ_k and then some possible construction of the estimators are provided. Note that diffusion models are included in this framework by simply taking $Y = X$ in equation (12).

The sample consists of $(X_{t_i}, Y_{t_i}), i = 0, 1, \dots, n$, where $t_i = i \Delta_n$ for $\Delta = \Delta_n = T/n$. Time T is fixed, so this scheme is purely high frequency and asymptotic results are of mixed normal type. The parameter θ belongs to Θ which is a bounded domain in \mathbb{R}^{d_0} , $d_0 \geq 1$. As already mentioned, the parameter θ is a nuisance in estimation of τ^* . Denote by θ_k^* the true value of θ_k for $k = 1, 2$ and let $\vartheta_n = |\theta_1^* - \theta_2^*|$. In order to obtain the desired results some conditions should be further assumed on the regularity of the coefficient σ , on the Hölder continuity of the trajectories and the behaviour of the process b_t , that is, the process itself cannot be too irregular in order to discover changes in the structure. The process b_t can have jumps. Before introducing formal mathematical conditions we need additional notation. We denote the modulus of continuity of a function $f : I \rightarrow \mathbb{R}^{d_0}$ by

$$w_I(\delta, f) = \sup_{s, t \in I, |s-t| \leq \delta} |f(s) - f(t)|$$

For a matrix A , we denote by A' the transpose of A , $\text{Tr}(A)$ the trace of A , $A^{\otimes 2} = AA'$ and by A^{-1} the inverse of A , further we use the notation $A[B] = \sum_{ij} a_{ij}b_{ij} = \text{Tr}(AB')$. Let $S(x, \theta) = \sigma(x, \theta)^{\otimes 2} \cdot \lambda_1(M)$ denotes the minimum eigenvalue of a symmetric matrix M . We now introduce the regularity conditions.

[H]_j (i) $\sigma(x, t)$ is a measurable function defined on $\mathcal{X} \times [0, T]$ satisfying

- (a) $\inf_{(x, \theta) \in \mathcal{X} \times \Theta} \lambda_1(S(x, \theta)) > 0$,
- (b) derivatives $\partial_{\theta}^{\ell} \sigma(0 \leq \ell \leq j + [d_0/2])$ exist and those functions are continuous on $\mathcal{X} \times \Theta$,
- (c) there exists a locally bounded function $L : \mathcal{X} \times \mathcal{X} \times \Theta \rightarrow \mathbb{R}_+$ such that

$$|\sigma(x_1, \theta) - \sigma(x_2, \theta)| \leq L(x_1, x_2, \theta) |x_1 - x_2|^{\alpha} \quad (x_1, x_2 \in \mathcal{X}, \theta \in \Theta)$$

for some constant $\alpha > 0$.

(ii) $(X_t)_{t \in [0, T]}$ is a progressively measurable process taking values in \mathcal{X} such that

$$w_{[0, T]} \left(\frac{1}{n}, X \right) = o_p(\vartheta_n^{1/\alpha})$$

as $n \rightarrow \infty$.

(iii) $(b_t)_{t \in [0, T]}$ is a progressively measurable process taking values in \mathbb{R}^d such that $(b_t - b_0)_{t \in [0, T]}$ is locally bounded.

The term ‘locally bounded’ in **[H]_j** (i) (c) means, as usual, being bounded on every compact set. The case where the drift b_t changes its structure at time τ^* , or at any other time, is included in our context because b_t admits jumps. The following two different asymptotics are considered:

- (A) θ_1^* and θ_2^* are fixed and do not depend on n .
- (B) θ_1^* and θ_2^* depend on n , and as $n \rightarrow \infty$, $\theta_1^* \rightarrow \theta_0^* \in \Theta$, $\vartheta_n \rightarrow 0$ and $n\vartheta_n^2 \rightarrow \infty$.

In Case **(A)**, ϑ_n is a constant ϑ_0 independent of n . Usual identifiability conditions should be assumed in addition to **[H]₁**. In Case **(B)** condition **[H]₂** is required.

As θ is considered as a nuisance parameter at this stage, the assumption that $|\hat{\theta}_k - \theta_k^*| = o_p(\vartheta_n)$ as $n \rightarrow \infty$ for $k = 1, 2$ is necessary to validate the estimating procedure for the change point. In case the parameters are known, $\hat{\theta}_k$ should be read as θ_k^* , and this condition is not required. Section 4.2 presents an example of estimator for θ_k which satisfies this condition.

4.1. Estimator of the Change Point

Let $\Delta_i Y = Y_{t_i} - Y_{t_{i-1}}$ and define

$$(13) \quad \Phi_n(t; \theta_1, \theta_2) = \sum_{i=1}^{\lfloor nt/T \rfloor} G_i(\theta_1) + \sum_{i=\lfloor nt/T \rfloor + 1}^n G_i(\theta_2)$$

with

$$(14) \quad G_i(\theta) = \log \det S(X_{t_{i-1}}, \theta) + \Delta_n^{-1} S(X_{t_{i-1}}, \theta)^{-1} [(\Delta_i Y)^{\otimes 2}]$$

The contrast function in (13) is a version of the one in Genon-Catalot and Jacod (1993). Suppose that there exists an estimator $\hat{\theta}_k$ for each θ_k , $k = 1, 2$. Each estimator is based on $(X_{t_i}, Y_{t_i})_{i=0,1,\dots,n}$. The change point estimator of τ^* is

$$\hat{\tau}_n = \arg \min_{t \in [0, T]} \Phi_n(t; \hat{\theta}_1, \hat{\theta}_2)$$

The estimator $\hat{\tau}_n$ has a structure similar to the estimator \hat{k}_0 in equation (2). It is possible to prove that, in both cases **(A)** and **(B)**, the change point estimator $\hat{\tau}_n$ is consistent for τ^* at the rate of convergence of order $n\vartheta_n^2$. More precisely, suppose that, in case **(B)**, the limit $\eta = \lim_{n \rightarrow \infty} \vartheta_n^{-1}(\theta_2^* - \theta_1^*)$ exists. Let Ξ be the positive-definite matrix

$$\Xi(x, \theta) = \left(\text{Tr}((\partial_{\theta^{i_1}} S) S^{-1} (\partial_{\theta^{i_2}} S) S^{-1})(x, \theta) \right)_{i_1, i_2=1}^{d_0} \quad \theta = (\theta^{(i)}).$$

Define further,

$$\mathbb{H}(v) = -2 \left(\Gamma_{\eta}^{\frac{1}{2}} \mathcal{W}(v) - \frac{1}{2} \Gamma_{\eta} |v| \right)$$

for $\Gamma_{\eta} = (2T)^{-1} \Xi(X_{\tau^*}, \theta_0^*)[\eta^{\otimes 2}]$, where \mathcal{W} is a two-sided standard Wiener process independent of X_{τ^*} . Then,

$$n\vartheta_n^2(\hat{\tau}_n - \tau^*) \rightarrow^{d_s(\mathcal{F}_T)} \rightarrow^d \underset{v \in \mathbb{R}}{\text{argmin}} \mathbb{H}(v)$$

as $n \rightarrow \infty$. This result is the analogue of (7) although in the present case, the double sided Brownian motion \mathcal{W} is pre-multiplied by the random Fisher information Γ_{η} , so this results involves mixed normal limit. In order to use it in practice a studentization procedure is required, that is, the quantity $n\vartheta_n^2(\hat{\tau}_n - \tau^*)$ has to be normalized with Γ_{η} evaluated at the change point estimator $\hat{\tau}_n$ and at $\hat{\theta}_1$. In this case, the limit above is similar to the one of (7). The joint convergence of the normalized $\hat{\tau}_n$ and $X_{\hat{\tau}_n}$ can also be proved.

A similar result can be obtained in case **(A)** but the limit involves Chi-squared random variables instead of Gaussians and the limit distribution is not known in closed form. For details see the original paper Iacus and Yoshida (2009).

4.2. *How to Construct Estimators of θ*

We will briefly discuss the construction of the initial estimators for θ_1 and θ_2 (more details can be found in Iacus and Yoshida, 2009). We construct the estimators $\hat{\theta}_1, \hat{\theta}_2$ using the data over the time interval $[0, a_n], [T - a_n, T]$, respectively, for some sequences a_n tending to zero. We assume that there exists a constant $\beta \in (0, 1/2)$ such that $a_n \geq 1/(n\vartheta_n^{1/\beta})$ and that $|\hat{\theta}_k - \theta_k^*| = o_p((na_n)^{-\beta})$ for $k = 1, 2$. In case (A), clearly $n\vartheta_n^2 \rightarrow \infty$. We further assume that $na_n \rightarrow \infty$, which is a natural requirement because it requires that the number of data has to be proportional to na_n . To obtain $\hat{\theta}_k$, we need the identifiability condition that $\sigma(\theta, x) = \sigma(\theta', x)$ implies $\theta = \theta'$; it is a strong condition like monotonicity of $\sigma(\theta, x)$ in θ . The first stage estimator $\hat{\theta}_1$ (resp. $\hat{\theta}_2$) of θ_1^* (resp. θ_2^*) is obtained using the first na_n observations from the left (resp. na_n from the right) via quasi-maximum likelihood estimation, that is,

$$\hat{\theta}_1 = \arg \min_{\theta} \sum_{i=1}^{[na_n]} G_i(\theta), \quad \hat{\theta}_2 = \arg \min_{\theta} \sum_{i=[n(1-a_n)]}^n G_i(\theta)$$

where $G_i(\theta)$ is given in (14). Once the first stage estimators $\hat{\theta}_1$ and $\hat{\theta}_2$ are available, the first stage estimator $\hat{\tau}_n$ of τ^* is obtained so as

$$\Phi_n(\hat{\tau}_n; \hat{\theta}_1, \hat{\theta}_2) = \min_{t \in [0, T]} \Phi_n(t; \hat{\theta}_1, \hat{\theta}_2)$$

Then, with the first stage estimator of τ^* in hands, we calculate the second stage estimators of θ_i^* using observations in the interval $[0, \hat{\tau}_n - b_n]$ for θ_1^* and observations in the interval $[\hat{\tau}_n + b_n, T]$ for θ_2^* , where (b_n) is a certain sequence of positive numbers tending to 0. We denote the second stage estimators of θ_i^* by $\check{\theta}_i$. Finally, the second stage estimator of τ^* , that is, $\check{\tau}_n$, is obtained as

$$(15) \quad \Phi_n(\check{\tau}_n; \check{\theta}_1, \check{\theta}_2) = \min_{t \in [0, T]} \Phi_n(t; \check{\theta}_1, \check{\theta}_2)$$

Under certain assumptions, $\check{\theta}_k$ are consistent and the asymptotic properties of $\check{\tau}_n$ are the same as for $\hat{\tau}_n$. Both the estimators $\hat{\tau}_n$ and $\check{\tau}_n$ are consistent, but the two stage procedure is suggested only to increase the performance of the change point estimator in finite samples.

4.3. *Modified Quasi-maximum Likelihood*

The quasi-maximum likelihood approach does not take into account the estimation of the drift term b_i . This approach can be widely applied because there is no need to specify the drift term, however, for small sample sizes or not sufficiently small Δ , the presence of the drift may

affect the estimation. Note also that, in this high-frequency setup, the drift term cannot be estimated consistently for the reasons already mentioned in Section 2.2.

To take small samples effects into account, it is possible to consider a new contrast function

$$(16) \quad \Phi'_n(t; \theta_1, \theta_2) = \sum_{i=1}^{[nt/2T]} G'_i(\theta_1) + \sum_{i=[nt/2T]+1}^{[n/2]} G'_i(\theta_2)$$

where the term $\Delta_i Y$ in $G_i(\theta)$ of (14) is replaced by

$$\tilde{\Delta}_i Y = \frac{Y_{2i+1} - 2Y_{2i} + Y_{2i-1}}{\sqrt{2}}$$

to define

$$(17) \quad G'_i(\theta) = \log \det S(X_{t_{2i-1}}, \theta) + \Delta_n^{-1} S(X_{t_{2i-1}}, \theta)^{-1} [(\tilde{\Delta}_i Y)^{\otimes 2}]$$

It can be proved that the first and second stage estimators obtained by this new contrast function have the same properties of the estimators in the previous sections. Intuitively, the use of $\tilde{\Delta}_i Y$ has the effect of compensating for the unknown drift. Indeed, the term $\tilde{\Delta}_i Y$ reduces to

$$\tilde{\Delta}_i Y = \frac{\left(\int_{t_{2i}}^{t_{2i+1}} b_s ds - \int_{t_{2i-1}}^{t_{2i}} b_s ds \right) + \left(\int_{t_{2i}}^{t_{2i+1}} \sigma(X_s, \theta) dW_s - \int_{t_{2i-1}}^{t_{2i}} \sigma(X_s, \theta) dW_s \right)}{\sqrt{2}}$$

We will compare this estimator with the one obtained on the original contrast function (13).

5. Numerical Study

In this section, we give simulation studies to assess the quality of the estimators of the change point. We compare the change point estimated by: (i) the least squares methods via non-parametric estimation of the drift as in Section 2.2 ($\hat{\tau}_n$ in the tables); (ii) the change point ($\tilde{\tau}_n$) obtained via the CUSUM statistics as the point at which the statistics T_n in (11) reaches its maximum; (iii) the quasi-likelihood maximum approach via the second stage estimator ($\check{\tau}_n$) as in (15) and the second stage estimator ($\check{\tau}'_n$ in the tables) based on the modified contrast function (16).

We first consider the following diffusion model without drift

$$X_t = \begin{cases} X_0 + \int_0^t (1 + X_s^2)^{\theta_1^*} dW_s & \text{for } t \in [0, \tau^*] \\ X_{\tau^*} + \int_{\tau^*}^t (1 + X_s^2)^{\theta_2^*} dW_s & \text{for } t \in [\tau^*, T] \end{cases} \quad (\text{Mod1})$$

and the same model (Mod1) with drift

$$X_t = \begin{cases} X_0 + \int_0^t X_s ds + \int_0^t (1 + X_s^2)^{\theta_1^*} dW_s & \text{for } t \in [0, \tau^*) \\ X_{\tau^*} + \int_{\tau^*}^t X_s ds + \int_{\tau^*}^t (1 + X_s^2)^{\theta_2^*} dW_s & \text{for } t \in [\tau^*, T] \end{cases} \quad (\text{Mod1B})$$

Models (Mod1) and (Mod1B) are particular cases of the short-term interest rates models used in, for example, Ait-Sahalia (1996). Further, we consider also the famous linear Vasicek's (1977) model

$$X_t = \begin{cases} X_0 + \int_0^t (2 - X_s) ds + \theta_1^* dW_s & \text{for } t \in [0, \tau^*) \\ X_{\tau^*} + \int_{\tau^*}^t (2 - X_s) ds + \theta_2^* dW_s & \text{for } t \in [\tau^*, T] \end{cases} \quad (\text{Mod2})$$

We also consider this model with cyclic volatility

$$X_t = \begin{cases} X_0 + \int_0^t \sqrt{2 + \theta_1^* \sin(4\pi s)} dW_s & \text{for } t \in [0, \tau^*) \\ X_{\tau^*} + \int_{\tau^*}^t \sqrt{2 + \theta_2^* \sin(4\pi s)} dW_s & \text{for } t \in [\tau^*, T]. \end{cases} \quad (\text{Mod3})$$

In all models τ^* is the true change point assumed to be $\tau^* = 0.6 \cdot T$. The true value of the parameters are $\theta_1^* = 0.2$ and $\theta_2^* = \theta_1^* + n^{-\gamma}$, with $\gamma = \frac{1}{4}$, n is the sample size and $\Delta_n = T/n$. The initial value for all simulation experiments is $X_0 = 5$. We construct different combinations of the sample size $n = 500, 1000, 2000$ and time horizon $T = 1, 2, 5, 10$ as reported in Tables 1–3 which correspond to the different $\Delta_n = 0.01, 0.001$. Clearly this setup also produces different values of $\vartheta_n = |\theta_2^* - \theta_1^*|$. For (Mod3) the initial value is $X_0 = 1$ and the true change point τ^* is set to $\tau^* = 0.5 \cdot T$ to ensure completeness of the volatility cycle.

Table 1: Average Results After 1000 Replications. In Parentheses Monte Carlo Standard Errors. Results for (Mod1)

n	Δ_n	$T = n\Delta_n$	τ^*	$\hat{\tau}_n$ L.S.	$\check{\tau}_n$ Q.MLE	$\tilde{\tau}_n$ CUSUM	$\check{\tau}'_n$ Modified Q.MLE.
1000	0.001	1	0.6	0.606 (0.061)	0.602 (0.012)	0.609 (0.020)	0.604 (0.031)
2000	0.001	2	1.2	1.202 (0.196)	1.205 (0.095)	1.208 (0.082)	1.214 (0.109)
500	0.01	5	3	3.020 (0.698)	3.024 (0.512)	3.038 (0.533)	3.075 (0.623)
1000	0.01	10	6	5.981 (1.122)	6.087 (1.617)	6.207 (1.182)	6.129 (1.296)

Table 2: Average Results After 1000 Replications. In Parentheses Monte Carlo Standard Errors. Results for (Mod1B)

n	Δ_n	$T = n\Delta_n$	τ^*	$\hat{\tau}_n$ L.S.	$\check{\tau}_n$ Q.MLE	$\tilde{\tau}_n$ CUSUM	$\check{\tau}'_n$ Modified Q.MLE
1000	0.001	1	0.6	0.609 (0.018)	0.600 (0.005)	0.607 (0.010)	0.602 (0.011)
2000	0.001	2	1.2	1.230 (0.049)	1.200 (0.005)	1.204 (0.007)	1.202 (0.010)
500	0.01	5	3	3.644 (0.320)	2.977 (0.020)	3.017 (0.067)	3.002 (0.052)
1000	0.01	10	6	8.179 (0.221)	3.607 (0.375)	6.004 (0.007)	6.001 (0.011)

Table 3: Average Results After 1000 Replications. In Parentheses Monte Carlo Standard Errors. Results for (Mod2).

n	Δ_n	$T = n\Delta_n$	τ^*	$\hat{\tau}_n$ L.S.	$\check{\tau}_n$ Q.MLE	$\tilde{\tau}_n$ CUSUM	$\check{\tau}'_n$ Modified Q.MLE
1000	0.001	1	0.6	0.604 (0.008)	0.601 (0.007)	0.601 (0.007)	0.602 (0.015)
2000	0.001	2	1.2	1.204 (0.009)	1.202 (0.009)	1.201 (0.009)	1.204 (0.017)
500	0.01	5	3	3.030 (0.051)	3.033 (0.084)	3.001 (0.055)	3.013 (0.145)
1000	0.01	10	6	6.036 (0.076)	6.027 (0.078)	6.001 (0.078)	6.022 (0.146)

Model (Mod1) and model (Mod1B) have been chosen in order to show the performance of change point estimators for highly nonlinear parametric models. Model (Mod1) differs from Model (Mod1B) in that the latter has a drift term. This choice has been made to show that the modified quasi-maximum likelihood estimator $\check{\tau}'_n$ is able to remove the drift effect for small sample size as explained in Section 4.3. As explained below, both models (Mod1) and (Mod1B) do not satisfy the assumptions of the least squares estimator $\hat{\tau}_n$. For this reason, in order to make a fair comparison between all estimators we also introduced model (Mod2), which is a mean reverting process with constant diffusion coefficient. Finally, (Mod3) fits the assumptions of both quasi-maximum likelihood and CUSUM approaches only.

While to be rigorous for model (Mod1) and (Mod1B) we can only apply the quasi-maximum likelihood approach of Section 4, model (Mod2)

Table 4: Average Results After 1000 Replications. In Parentheses Monte Carlo Standard Errors. Results for (Mod3).

n	Δ_n	$T = n\Delta_n$	τ^*	$\hat{\tau}_n$ L.S.	$\check{\tau}_n$ Q.MLE	$\tilde{\tau}_n$ CUSUM	$\check{\tau}'_n$ Modified Q.MLE
1000	0.001	1	0.5	0.588 (0.210)	0.479 (0.259)	0.597 (0.233)	0.497 (0.300)
2000	0.001	2	1.0	1.076 (0.448)	0.947 (0.582)	1.137 (0.389)	0.967 (0.615)
500	0.01	5	2.5	2.516 (1.056)	2.483 (1.356)	2.850 (1.292)	2.574 (1.520)
1000	0.01	10	5	5.102 (2.216)	4.940 (2.848)	5.721 (2.332)	4.995 (3.108)

is a simple Ornstein–Uhlenbeck model which fits also the setup of Sections 2 and 3. Model (Mod1) is a misspecified case for the assumptions in Section 2.2 on the least-squares approach: $dX_t = b(X_t) dt + \sqrt{\theta} dW_t$. In this case, the estimators $\bar{\theta}_i$ in Section 2 estimate the quadratic variation rather than the true value of θ , but the change in the quadratic variation is sometimes enough to identify the change point.

Tables 1–4 report the results of 1000 replications of the experiments under the different conditions. The average values of the estimators $\hat{\tau}_n$, $\check{\tau}_n$, $\tilde{\tau}_n$ and $\check{\tau}'_n$ are reported along with their Monte Carlo standard errors (in parentheses). Although the third decimal point is not significant we report it just for comparison purposes.

The behaviour of the CUSUM test statistics has been discussed already in Song and Lee (2009), so we do not consider this problem in our experiments but we report the value of the CUSUM change point estimator $\tilde{\tau}_n$. For the estimators of the θ 's in the CUSUM test statistics, we used the contrast function in Kessler (1997) using derivatives up to order 2 like in the original paper of Song and Lee (2009). Note that the application of the CUSUM statistic requires the additional knowledge of the drift function, while all other methods do not use this information. This is a small advantage for the CUSUM test statistic in a simulation study, although unrealistic in practical daily analysis of financial data.

Our numerical analysis mainly focuses on the problem of change point estimation rather than on the tests for the existence of a change point. Indeed, all methods presented here correctly identify the presence of a change point in all the simulations performed. Differences emerge (and are usually of interest in applications) in the identification (estimation) of the change point time instant. In fact, the estimation of the parameters in the volatility term of the stochastic differential equation can be done separately

for the two sub-series identified by the change point estimate and estimators of these parameters are consistent and asymptotically normal at the usual rate \sqrt{n} , with n the number of observations as shown in the previous sections.

5.1. Numerical Evidence

From the Monte Carlo experiments summarized in Tables 1–4 it emerges that the CUSUM estimator behaves well only in the case of Model (Mod2), that is, when the model is mean reverting and when the drift can be completely specified in the construction of the Kessler's likelihood (see Table 3).

The least squares estimator $\hat{\tau}_n$ seems to have a good performance in term of bias and variability for models with constant or bounded drift (Tables 1 and 3), while it behaves badly in the presence of unbounded drift when time T grows (see Table 2). CUSUM statistics do not work well in the cyclic case of Model (Mod3) at all frequencies (see Table 4).

Quasi-maximum likelihood estimator $\check{\tau}_n$ and its modified version $\check{\tau}'_n$ both behave well, although $\check{\tau}_n$ has lower variance than $\check{\tau}'_n$ due to the fact that the latter uses less (actually half) observations to construct the contrast function. In some cases, and when the drift plays a role in the stochastic differential equation, the estimator $\check{\tau}'_n$ based on the modified quasi-likelihood has better properties as expected. For example, in Table 2 for sample size $n = 500$ and low frequency $\Delta_n = 0.01$ the estimator $\check{\tau}'_n$ is preferable to $\check{\tau}_n$.

The conclusion from this analysis is that the estimator $\check{\tau}_n$ and in particular its modified version $\check{\tau}'_n$ behave almost uniformly better or equally well than other methods proposed in the literature so far. Moreover, it applies to a wide class of models (one or multi-dimensional) and does not require the unrealistic specification of the drift term.

All the analysis has been done using the R statistical environment (R Development Core Team, 2009) and the package *sde* (see Iacus, 2008) and *Yuima* (see, Yuima Project Team, 2010). Code is available on request to the authors.

REFERENCES

- Y. AIT-SAHALIA (1996), "Testing Continuous-Time Models of the Spot Interest Rate," *Review of Financial Studies*, 9, pp. 385–426.
- J. BAI (1994), "Least Squares Estimation of a Shift in Linear Processes," *Journal of Times Series Analysis*, 15, pp. 453–72.
- J. BAI (1997), "Estimation of a Change Point in Multiple Regression Models," *The Review of Economics and Statistics*, 79, pp. 551–63.
- F.M. BANDI – P.C.B. PHILLIPS (2003), "Fully Nonparametric Estimation of Scalar Diffusion Models," *Econometrica*, 71, pp. 241–83.
- F. BLACK – M.S. SCHOLES (1973), "The Pricing of Options and Corporate Liabilities," *Journal of Political Economy*, 81, pp. 637–54.
- G. CHEN – Y.K. CHOI – Y. ZHOU (2005), "Nonparametric Estimation of Structural Change Points in Volatility Models for Time Series," *Journal of Econometrics*, 126, pp. 79–144.
- M. CSÖRGÖ – L. HORVÁTH (1997), *Limit Theorems in Change-point Analysis*. New York: Wiley.
- A. DE GREGORIO – S.M. IACUS (2008), "Least Squares Volatility Change Point Estimation for Partially Observed Diffusion Processes," *Communications in Statistics, Theory and Methods*, 37, pp. 2342–57.
- D. FLORENS-ZMIROU (1993), "On Estimating the Diffusion Coefficient from Discrete Observations," *Journal of Applied Probability*, 30, pp. 790–804.
- V. GENON-CATALOT – J. JACOD (1993), "On the Estimation of the Diffusion Coefficient for Multidimensional Diffusion Processes," *Annales de l'Institut Henri Poincaré*, 29, pp. 119–51.
- D.V. HINKLEY (1971), "Inference about the change-point from cumulative sum tests," *Biometrika*, 58, pp. 509–23.
- S.M. IACUS (2008), *Simulation and Inference for Stochastic Differential Equations: with R Examples*, Springer Series in Statistics, New York: Springer.
- S.M. IACUS – N. YOSHIDA (2009), "Estimation for the Change Point of the Volatility in a Stochastic Differential Equation," available at <http://arxiv.org/abs/0906.3108>.
- C. INCLAN – G.C. TIAO (1994), "Use of Cumulative Sums of Squares for Retrospective Detection of Change of Variance," *Journal of the American Statistical Association*, 89, pp. 913–23.
- I. KARATZAS – S.E. SHREVE (1991), *Brownian Motion and Stochastic Calculus*, New York: Springer-Verlag.
- M. KESSLER (1997), "Estimation of an Ergodic Diffusion from Discrete Observations," *Scandinavian Journal of Statistics*, 24, pp. 211–29.

- S. KIM – S. CHO – S. LEE (2000), “On the Cusum Test for Parameter Changes in GARCH(1,1) Models,” *Communications in Statistics - Theory and Methods*, 29, pp. 445–62.
- Y. KUTOYANTS (1994), *Identification of Dynamical Systems with Small Noise*, Dordrecht: Kluwer.
- Y. KUTOYANTS (2004), *Statistical Inference for Ergodic Diffusion Processes*, London: Springer-Verlag.
- S. LEE – J. HA – O. NA – S. NA (2003), “The Cusum Test for Parameter Change in Time Series Models,” *Scandinavian Journal of Statistics*, 30, pp. 781–96.
- S. LEE – Y. NISHIYAMA – N. YOSHIDA (2006), “Test for Parameter Change in Diffusion Processes by Cusum Statistics Based on One-Step Estimators,” *Annals of the Institute of Statistical Mathematics*, 58, pp. 211–22.
- R.C. MERTON (1974), “Theory of Rational Option Pricing,” *Bell Journal of Economics and Management Science*, 41, pp. 141–83.
- D.T. PHAM (1981), “Nonparametric Estimation of the drift Coefficient in the Diffusion Equation,” *Statistics*, 12, pp. 61–73.
- R Development Core Team (2009), *R: A Language and Environment for Statistical Computing*, Vienna, Austria: R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL: <http://www.R-project.org>.
- G. ROUSSAS (1972), *Contiguity of Probability Measures. Some Applications in Statistics*, Cambridge: Cambridge University Press.
- B.W. SILVERMAN (1986), *Density Estimation*. London: Chapman and Hall.
- L. SMALDONE (2009), *Analysis of the Change Point of the Volatility in Financial Markets*, in Italian, Master Thesis, University of Milan.
- J. SONG – S. LEE (2009), “Test for Parameter Change in Discretely Observed Diffusion Processes,” *Statistical Inference for Stochastic Processes*, 12, pp. 165–83.
- R. STANTON (1997), “A Non-parametric Model of Term Structure Dynamics and the Market Price of Interest Rate Risk,” *Journal of Finance*, 52, pp. 1973–2002.
- O. VASICEK (1977), “An Equilibrium Characterization of the Term Structure,” *Journal of Financial Economics*, 5, pp. 177–88.
- X. WANG – P.C.B. PHILLIPS – J. YU (2010), “Bias in Estimating Linear Multivariate Diffusions,” Working Paper, available at www.mysmu.edu/faculty/yujun/Research/multidiffusion13.pdf.
- YUIMA Project Team (2010), “yuima: The YUIMA Project Package,” available at <http://R-Forge.R-project.org/projects/yuima>.

Non-technical Summary

In financial markets it is crucial to have an accurate description of the volatility of the market and/or the different financial products. All pricing formulas make use of the historical values of the volatility as a fundamental ingredient. It is well known, however, that volatility is not constant over time, even short time, and thus the monitoring of the volatility is one of the primary tasks in empirical finance.

The statistical way to monitor structural changes is called change point analysis or change point estimation. Stochastic differential equations are among the most used stochastic models to describe continuous time financial time series. Although data are collected in discrete time, the underlying structure of the continuous model allows for very detailed analysis of these data.

This paper reviews recent advances on change point estimation for the volatility component of stochastic differential equations under different discrete sampling schemes, namely the rapidly increasing design and the pure high frequency sampling.

Both ergodic and non-stationary models are considered and different techniques have been developed to handle the different cases. In particular, the methods presented are the least squares, the CUSUM and the quasi-maximum likelihood methods.

To corroborate the theoretical results, an extensive Monte Carlo analysis under different sampling schemes and stochastic models is performed with the aim of comparing the different statistical approaches to the identification of the change point instant.

An analysis of the recent financial crisis has been performed retrospectively using the above techniques. The results show that monitoring of the volatility, although it cannot provide an explanation, can conversely provide interesting insights about instability of the market.