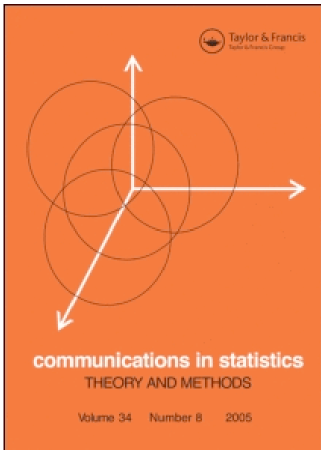

Communications in Statistics - Theory and Methods
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713597238

Least Squares Volatility Change Point Estimation for Partially Observed Diffusion Processes

Alessandro De Gregorio [a]; Stefano M. Iacus [a]
[a] Dipartimento di Scienze Economiche, Aziendali e Statistiche, Milan, Italy

Online Publication Date: 01 September 2008

To cite this Article: Gregorio, Alessandro De and Iacus, Stefano M. (2008) 'Least Squares Volatility Change Point Estimation for Partially Observed Diffusion Processes', Communications in Statistics - Theory and Methods, 37:15, 2342 — 2357

To link to this article: DOI: 10.1080/03610920801919692
URL: http://dx.doi.org/10.1080/03610920801919692

Communications in Statistics—Theory and Methods, 37: 2342–2357, 2008

ISSN: 0361-0926 print/1532-415X online
DOI: 10.1080/03610920801919692

Inference

Least Squares Volatility Change Point Estimation for Partially Observed Diffusion Processes

ALESSANDRO DE GREGORIO
AND STEFANO M. IACUS

Dipartimento di Scienze Economiche, Aziendali e Statistiche,
Milan, Italy

A one-dimensional diffusion process $X = \{X_t, 0 \leq t \leq T\}$, with drift $b(x)$ and diffusion coefficient $\sigma(\theta, x) = \sqrt{\theta}\sigma(x)$ known up to $\theta > 0$, is supposed to switch volatility regime at some point $t^* \in (0, T)$. On the basis of discrete time observations from X, the problem is the one of estimating the instant of change in the volatility structure t^* as well as the two values of θ, say θ_1 and θ_2, before and after the change point. It is assumed that the sampling occurs at regularly spaced times intervals of length Δ_n with $n\Delta_n = T$. To work out our statistical problem we use a least squares approach. Consistency, rates of convergence and distributional results of the estimators are presented under an high frequency scheme. We also study the case of a diffusion process with unknown drift and unknown volatility but constant.

Keywords Change point problem; Diffusion process; Discrete observations; Nonparametric estimator; Volatility regime switch.

Mathematics Subject Classification Primary 60K99; Secondary 62M99.

1. Introduction

Change-point problems have originally arisen in the context of quality control, but the problem of abrupt changes in general arises in many contexts like epidemiology, rhythm analysis in electrocardiograms, seismic signal processing, study of archeological sites, and financial markets.

Originally, the problem was considered for i.i.d samples (see Csörgő and Horváth, 1997; Hinkley, 1971; Inclan and Tiao, 1994) and it moved naturally into the time series context as economic time series often exhibit prominent evidence for

Received September 18, 2007; Accepted January 15, 2008
Address correspondence to Stefano M. Iacus, Dipartimento di Scienze Economiche, Aziendali e Statistiche, Via Conservatorio 7, Milan 20122, Italy; E-mail: stefano.iacus@unimi.it

structural change in the underlying model (see, for example, Chen et al., 2005; Kim et al., 2000; Lee et al., 2003; the papers cited therein).

In this article, we deal with a change-point problem for the volatility of a diffusion process observed at discrete times. The instant of the change in volatility regime is identified retrospectively by the method of the least squares along the lines proposed in Bai (1994). For continuous time observations of diffusion processes Lee et al. (2006) considered the change point estimation problem for the drift. In the present work, the drift coefficient of the stochastic differential equation is assumed known and if not it is estimated nonparametrically.

The article is organized as follows. Section 2 introduces the model of observation and the estimator of the change point instant and the estimators of the volatilities before and after the change. Section 3 analyzes the asymptotic properties of the estimators. In Sec. 4, we consider the case when the drift is unknown and the diffusion coefficient does not depend on the state of the process. All the proofs of the theorems are contained in the Appendix.

2. The Least Squares Estimator

We denote by $X = \{X_t, 0 \leq t \leq T\}$ the diffusion process, with state space $I = (l, r)$, $-\infty \leq l \leq r \leq +\infty$, solution of the stochastic differential equation:

$$\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sqrt{\theta}\sigma(X_t)\mathrm{d}W_t, \tag{2.1}$$

with $X_0 = x_0$ and $\{W_t, t \geq 0\}$ a standard Brownian motion. We suppose that the value of θ is θ_1 up to some unknown time $t^* \in (0, T)$ and θ_2 after, i.e., $\theta = \theta_1 \mathbf{1}_{\{t \leq t^*\}} + \theta_2 \mathbf{1}_{\{t > t^*\}}$. The parameters θ_1 and θ_2 belong to Θ, a compact set of $\mathbb{R}^+$. The coefficients $b : I \to \mathbb{R}$ and $\sigma : I \to (0, \infty)$ are supposed to be known with continuous derivatives. The continuity of the derivatives of $b(\cdot)$ and $\sigma(\cdot)$ assures that it exists a unique continuous process solution to (2.1), which is defined up an explosion time (see Arnold, 1974). Let $s(x) = \exp\{-\int_{x_0}^{x} 2b(u)/\sigma^2(u)\mathrm{d}u\}$ be the scale function (where x_0 is an arbitrary point inside I).

A1. $\lim_{x_1 \to l} \int_{x_1}^{x} s(u)\mathrm{d}u = -\infty$, $\lim_{x_2 \to r} \int_{x}^{x_2} s(u)\mathrm{d}u = +\infty$, where $l < x_1 < x < x_2 < r$.

Condition A1 guarantees that the exit time from I is infinite (see Karatzas and Shreve, 1991).

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

In order to obtain a simple least squares estimator, we follow the same approach proposed in Bai (1994). To this end, we make use of Euler approximation to the solution of (2.1), i.e.,

$$X_{i+1} = X_i + b(X_i)\Delta_n + \sqrt{\theta}\sigma(X_i)(W_{i+1} - W_i),$$

and introduce the quantities

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

which represent n independent standard normal variables.

We denote by $k_0 = [n\tau_0]$ and $k = [n\tau]$, $\tau, \tau_0 \in (0, 1)$, where $[x]$ is the integer part of the real value x. Given that $\Delta_n \to 0$, without loss of generality, we can assume that the process switches volatility regime exactly at time $t_i = t_{k_0} = k_0\Delta_n = t^*$. The least squares estimator of the change point is obtained as follows:

$$\begin{aligned}\hat{k}_0 &= \arg\min_k \left(\min_{\theta_1, \theta_2} \left\{ \sum_{i=1}^{k} (Z_i^2 - \theta_1)^2 + \sum_{i=k+1}^{n} (Z_i^2 - \theta_2)^2 \right\} \right) \\ &= \arg\min_k \left\{ \sum_{i=1}^{k} (Z_i^2 - \bar{\theta}_1)^2 + \sum_{i=k+1}^{n} (Z_i^2 - \bar{\theta}_2)^2 \right\}, \end{aligned} \tag{2.2}$$

where

$$\begin{aligned}\bar{\theta}_1 &= \arg\min_{\theta_1} \sum_{i=1}^{k} \left(Z_i^2 - \theta_1\right)^2 = \frac{1}{k}\sum_{i=1}^{k} Z_i^2 = \frac{S_k}{k} \\ \bar{\theta}_2 &= \arg\min_{\theta_2} \sum_{i=k+1}^{n} \left(Z_i^2 - \theta_2\right)^2 = \frac{1}{n-k}\sum_{i=k+1}^{n} Z_i^2 = \frac{S_{n-k}}{k}\end{aligned}$$

and $k = 1, \dots, n-1$. We introduce the following quantity:

$$U_k^2 = \sum_{i=1}^{k} \left(Z_i^2 - \bar{\theta}_1\right)^2 + \sum_{i=k+1}^{n} \left(Z_i^2 - \bar{\theta}_2\right)^2;$$

then we have that:

$$\hat{k}_0 = \arg\min_k U_k^2$$

To study the asymptotic properties of U_k^2 it is better to rewrite it in the following way:

$$U_k^2 = \sum_{i=1}^{n} \left(Z_i^2 - \overline{Z}_n\right)^2 - nV_k^2$$

where $\overline{Z}_n = \frac{1}{n}\sum_{i=1}^{n} Z_i^2 = \frac{1}{n}S_n$ and

$$V_k = \left(\frac{k(n-k)}{n^2}\right)^{\frac{1}{2}} (\bar{\theta}_2 - \bar{\theta}_1) = \frac{S_n D_k}{\sqrt{k(n-k)}}$$

with $D_k = k/n - S_k/S_n$.

This representation of U_k^2 is obtained by lengthy but straightforward algebra and it is rather useful because minimization of U_k^2 is equivalent to the maximization of V_k and hence D_k. So it is easier to consider the following estimator of k_0:

$$\hat{k}_0 = \arg\max_k |D_k| = \arg\max_k (k(n-k))^{\frac{1}{2}} |V_k|. \tag{2.3}$$

As a side remark, it can be noticed that for fixed k (and under suitable hypothesis), D_k can be seen as an approximate likelihood ratio statistics for testing the null hypothesis of no change in volatility (see, e.g., Inclan and Tiao, 1994). We do not discuss approximate likelihood approach in this article.

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

$$\hat{\theta}_1 = \frac{S_{\hat{k}_0}}{\hat{k}_0}, \quad \hat{\theta}_2 = \frac{S_{n-\hat{k}_0}}{n-\hat{k}_0}. \tag{2.4}$$

We will prove consistency of $\hat{k}_0$, $\hat{\theta}_1$, and $\hat{\theta}_2$ and also distributional results for these estimators.

Our first result concerns the asymptotic distribution of the statistic D_k under the condition that no change of volatility occurs during the interval $[0, T]$.

Theorem 2.1. *Assume that* H_0: $\theta_1 = \theta_2 = 1$, *then we have that*:

$$\sqrt{\frac{n}{2}} |D_k| \xrightarrow{d} |W^0(\tau)|, \tag{2.5}$$

where $\{W^0(\tau), 0 \leq \tau \leq 1\}$ *is a Brownian bridge.*

Corollary 2.1. *From Theorem* 2.1 *we derive immediately that for* $\delta \in (0, 1/2)$:

$$\sqrt{\frac{n}{2}} \sup_{\delta n \leq k \leq (1-\delta)n} |D_k| \xrightarrow{d} \sup_{\delta \leq \tau \leq (1-\delta)} |W^0(\tau)|, \tag{2.6}$$

$$\sqrt{\frac{n}{2}} \sup_{\delta n \leq k \leq (1-\delta)n} |V_k| \xrightarrow{d} \sup_{\delta \leq \tau \leq (1-\delta)} (\tau(1-\tau))^{-1/2} |W^0(\tau)|. \tag{2.7}$$

The last asymptotic results are useful to test if a change point occurred in $[0, T]$. In particular, it is possible to obtain the asymptotic critical values for the distribution (2.7) by means of the same arguments used in Csörgő and Horváth (1997, p. 25).

3. Asymptotic Properties of the Estimator

We study the main asymptotic properties of the least squares estimator $\hat{k}_0$. We start analyzing the consistency and the rate of convergence of the change-point estimator (2.3). It is convenient to note that the rate of convergence is particularly important not only to describe how fast the estimator converges to the true value, but also to get the limiting distribution. The next Theorem represents our first result on the consistency.

Theorem 3.1. *The estimator $\hat{\tau}_0 = \frac{\hat{k}_0}{n}$ satisfies*

$$|\hat{\tau}_0 - \tau_0| = n^{-1/2}(\theta_2 - \theta_1)^{-1} O_p(\sqrt{\log n}). \tag{3.1}$$

Theorem 3.1 implies consistency of our estimator, in fact we have that $n^{\beta}(\hat{\tau}_0 - \tau_0) \to 0$ in probability for any $\beta \in (0, 1/2)$. We are able to improve the rate of convergence of $\hat{\tau}_0$.

Theorem 3.2. *We have the following result*:

$$\hat{\tau}_0 - \tau_0 = O_p\left(\frac{1}{n(\theta_2 - \theta_1)^2}\right). \tag{3.2}$$

It is also possible to derive the asymptotic distribution of $\hat{\tau}_0$ under our limiting framework for small variations of the rate of change of the direction. The case $\vartheta_n = \theta_2 - \theta_1$ equal to a constant is less interesting because when ϑ_n is large the estimate of k_0 is quite precise.

A2. Assume that:

$$\vartheta_n \to 0, \quad \frac{\sqrt{n}\vartheta_n}{\sqrt{\log n}} \to \infty.$$

Under A2 the consistency of $\hat{\tau}_0$ follows immediately either from Theorem 3.1 or Theorem 3.2. In order to obtain the next result, it is useful to observe that:

$$\hat{k}_0 = \arg\max_k V_k^2 = \arg\max_k n\left(V_k^2 - V_{k_0}^2\right) \tag{3.3}$$

and to define a two-sided Brownian motion $\mathcal{W}(u)$ in the following manner:

$$\mathcal{W}(u) = \begin{cases} W_1(-u), & u < 0 \\ W_2(u), & u \geq 0 \end{cases}, \tag{3.4}$$

where W_1, W_2 are two independent Brownian motions. Now we present the following convergence in distribution result.

Theorem 3.3. *Under Assumption* A2 *we have that*:

$$\frac{n\vartheta_n^2(\hat{\tau}_0 - \tau_0)}{2\tilde{\theta}^2} \xrightarrow{d} \arg\max_v \left\{\mathcal{W}(v) - \frac{|v|}{2}\right\}, \tag{3.5}$$

where $W(v)$ is a two-sided Brownian motion and $\tilde{\theta}$ is a consistent estimator for θ_1 or θ_2.

Let θ_0 be the limiting value of both θ_1 and θ_2. Using the consistency result, we are able to obtain the asymptotic distributions for the estimators $\hat{\theta}_1$, $\hat{\theta}_2$, defined in (2.4).

Theorem 3.4. *Under Assumption* A2 *we have that*:

$$\sqrt{n}\begin{pmatrix}\hat{\theta}_1-\theta_1\\ \hat{\theta}_2-\theta_2\end{pmatrix}\xrightarrow{d} N(0,\Sigma), \tag{3.6}$$

where

$$\Sigma=\begin{pmatrix}2\tau_0^{-1}\theta_0^2 & 0\\ 0 & 2(1-\tau_0)^{-1}\theta_0^2\end{pmatrix}. \tag{3.7}$$

Remark 3.1. It is easy to verify that the Theorems presented in this section are also true if we consider an horizon time tending to infinite, i.e., $\Delta_n\to 0$, $n\Delta_n=T\to\infty$ as $n\to\infty$.

Remark on Ergodic Case. If the Euler approximation is not admissible, it is worth to consider the ergodic case. Let $m(u)=(\sigma^2(u)s(u))^{-1}$ be the speed measure of the diffusion process X_t. We introduce the following assumptions:

A3. $\int_l^r m(x)dx<+\infty$.

A4. $X_0=x_0$ has distribution P^0.

Assumptions A1, A3, and A4 imply that the process X_t is also ergodic and strictly stationary with invariant distribution $P^0(\mathrm{d}x)=\pi(x)\mathrm{d}x$, $\pi(x)=m(x)/\int_l^r m(u)\mathrm{d}u$. Under the additional condition:

A5. $\lim_{x\to 0}\sigma(x)\pi(x)=0$ or $\lim_{x\to\infty}\sigma(x)\pi(x)=0$ and

$$\lim_{x\to 0}\left|\frac{\sigma(x)}{2b(x)-\sigma(x)\sigma'(x)}\right|<\infty \quad\text{or}\quad \lim_{x\to\infty}\left|\frac{\sigma(x)}{2b(x)-\sigma(x)\sigma'(x)}\right|<\infty$$

the observed data $X_i, i=0,1,\ldots,n$, is a strictly stationary β-mixing sequence satisfying $k^\delta\beta_k\to 0$ as $k\to\infty$ for some fixed $\delta>1$ (see, e.g., Aït-Sahalia, 1996).

Under this setup and $n\Delta_n=T\to\infty$, similar results to the ones presented in the above can be obtained using the same techniques of Chen et al. (2005).

4. Estimation of the Change Point with Unknown Drift

We want to analyze the change point problem for a diffusion process X_t, when the drift coefficient $b(\cdot)$ is unknown, while the diffusion coefficient is supposed unknown but independent from the state of the process X_t. In other words, the process X_t is the solution of the following reduced stochastic differential equation:

$$\mathrm{d}X_t=b(X_t)\mathrm{d}t+\sqrt{\theta}\,\mathrm{d}W_t, \tag{4.1}$$

and the observation scheme and the asymptotics are as in Sec. 2. Let $K\geq 0$ be a kernel function, i.e., K is symmetric and continuously differentiable, with $\int_{\mathbb{R}}uK(u)\mathrm{d}u=0$, $\int_{\mathbb{R}}K^2(u)\mathrm{d}u<\infty$ and such that $\int_{\mathbb{R}}K(u)\mathrm{d}u=1$. We start introducing the following quantities:

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

where

$$\hat{b}(x) = \frac{\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)\frac{X_{i+1}-X_i}{\Delta_n}}{\sum_{i=1}^{n} K\left(\frac{X_i - x}{h_n}\right)} \tag{4.2}$$

is a nonparametric estimator of the drift constructed using the full sample and h_n is the bandwidth defined as in Silverman (1986). The estimator (4.2) is a particular case of the nonparametric estimator studied in Bandi and Phillips (2003). Unfortunately, for fixed T the drift coefficient cannot be estimated consistently. So in this section we assume that $n\Delta_n = T \to \infty$, $\Delta_n \to 0$ as $n \to \infty$. Moreover, some relationship between the bandwidth of the kernel and the mesh Δ_n should be required. Let X be a solution of $\mathrm{d}X_t = b(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t$ and define as:

$$\overline{L}_X(t, x) = \frac{1}{\sigma^2(x)} \lim_{\epsilon\to 0} \frac{1}{2\epsilon}\int_0^t \mathbf{1}_{\{x,x+\epsilon\}}(X_s)\sigma^2(X_s)\mathrm{d}s$$

the chronological local time of X. For the model (4.1), the chronological local time is simply:

$$\overline{L}_X(t, x) = \lim_{\epsilon\to 0} \frac{1}{2\epsilon}\int_0^t \mathbf{1}_{\{x,x+\epsilon\}}(X_s)\mathrm{d}s.$$

Bandi and Phillips (2003) showed that under the following additional assumption a consistent estimator of $b(\cdot)$ is given by (4.2).

A6. Assume that:

$$\frac{\overline{L}_X(T, x)}{h_n}\sqrt{\Delta_n \log\left(\frac{1}{\Delta_n}\right)} = o_{a.s.}(1)$$

and $h_n\overline{L}_X(T, x) \overset{a.s.}{\to} \infty$, with $n\Delta_n = T \to \infty$ and $\Delta_n \to 0$ as $n \to \infty$. In the stationary case, A6 may be replaced by $\frac{T}{h_n}\sqrt{\Delta_n \log(\frac{1}{\Delta_n})} = o_{a.s.}(1)$ and standard results on the estimation of first moments for discrete-time series hold (see Härdle, 1990; Pagan and Ullah, 1999).

The least squares estimator now takes the following form:

$$\tilde{k}_0 = \arg\min_k \left\{ \sum_{i=1}^{k} \left(\widehat{Z}_i^2 - \bar{\theta}_1^*\right)^2 + \sum_{i=k+1}^{n} \left(\widehat{Z}_i^2 - \bar{\theta}_2^*\right)^2 \right\}, \tag{4.3}$$

where

$$\bar{\theta}_1^* = \frac{\widehat{S}_k}{k}, \quad \bar{\theta}_2^* = \frac{\widehat{S}_{n-k}}{n-k},$$

with $\widehat{S}_k = \sum_{i=1}^{k} \widehat{Z}_i^2$, $\widehat{S}_{n-k} = \sum_{i=k+1}^{n} \widehat{Z}_i^2$. From (4.3), by the same steps considered in Sec. 2, we derive:

$$\widehat{V}_k = \left(\frac{k(n-k)}{n^2}\right)^{\frac{1}{2}} (\bar{\theta}_2^* - \bar{\theta}_1^*) \tag{4.4}$$

We are able to show that the asymptotic properties of the estimator $\tilde{k}_0$ defined in (4.3), are equal to the ones of $\hat{k}_0$.

Theorem 4.1. *Under the additional condition* A6, *the same results presented in Theorems* 3.1–3.4 *hold for the estimator* $\tilde{k}_0$.

Remark 4.1. In order to apply the theoretic results, an important issue in the nonparametric context is the selection of bandwidth. As shown by Bandi and Phillips (2003), for the practical implementation is often sufficient to choose the bandwidth proportional to $n^{-\nu}$, $\nu \in (0, 1/2)$. In other words, we can fix $h_n = cn^{-\nu}$, where the proportionality constant c might be obtained by means of the methods for bandwidth selection in density estimation (see, for example, Härdle, 1990; Pagan and Ullah, 1999).

Another way to work out this problem is a cross-validation method as considered in Chen et al. (2005). Hence, the optimal bandwidth $\hat{h}_n$ is obtained in the following manner:

$$\hat{h}_n = \arg\min_{h_n} \sum_{i=0}^{n} \left(\frac{X_{i+1} - X_i}{\Delta_n} - \hat{b}(X_i) \right)^2 w(X_i),$$

where $w(\cdot)$ is a weight function.

Appendix

As in Chen et al. (2005), some of the proofs are based on the ones in Bai (1994). We adapt Bai's theorems making the appropriate (but crucial) adjustments when needed, skipping all algebraic calculations which can be found in the original article of the author.

Proof of Theorem 2.1. By setting $\xi_i = Z_i^2 - 1$, under H_0 we note that $E(\xi_i) = 0$ and $Var(\xi_i) = 2$. Let us introduce the quantity

$$Y_n(\tau) = \frac{1}{\sqrt{2n}} \mathcal{S}_{[n\tau]} + (n\tau - [n\tau]) \frac{1}{\sqrt{2n}} \xi_{[n\tau]+1},$$

where $\mathcal{S}_n = \sum_{i=1}^{n} \xi_i$. It's not hard to see that

$$\left| \frac{1}{\sqrt{2n}} \sum_{i=1}^{[n\tau]} \xi_i - \frac{\sqrt{\tau}}{\sqrt{2[n\tau]}} \sum_{i=1}^{[n\tau]} \xi_i \right| \xrightarrow{p} 0, \tag{5.1}$$

and $Var\left(\frac{\sqrt{\tau}}{\sqrt{2[n\tau]}} \sum_{i=1}^{[n\tau]} \xi_i\right) = \tau$. Since the Lindeberg condition is true:

$$\sum_{i=1}^{[n\tau]} \frac{E\{\mathbf{1}_{|\xi_i| \geq \varepsilon\sqrt{2n}\}} \xi_i^2\}}{2n} \to 0, \tag{5.2}$$

we can conclude that:

$$\frac{1}{\sqrt{2n}} \mathcal{S}_{[n\tau]} \xrightarrow{d} N(0, \tau). \tag{5.3}$$

By Donsker's theorem we have $Y_n \xrightarrow{d} W(\tau)$ that implies $Y_n(\tau) - \tau Y_n(1) \xrightarrow{d} W^0(\tau)$, where $W(\tau)$ and $W^0(\tau)$ are, respectively, a standard Brownian motion and a Brownian bridge. Let $[n\tau] = k$, $k = 1, 2, \dots, n$, then:

$$\begin{aligned} Y_n(\tau) - \tau Y_n(1) &= \frac{1}{\sqrt{2n}} \mathscr{S}_{[n\tau]} - \frac{\tau}{\sqrt{2n}} \mathscr{S}_n + (n\tau - [n\tau]) \frac{1}{\sqrt{2n}} \xi_{[n\tau]+1} \\ &= \frac{1}{\sqrt{2n}} \left[\mathscr{S}_k - \frac{k}{n} \mathscr{S}_n \right] + (n\tau - [n\tau]) \frac{1}{\sqrt{2n}} \xi_{[n\tau]+1}. \end{aligned} \tag{5.4}$$

Now, by observing that:

$$\mathscr{S}_k - \frac{k}{n} \mathscr{S}_n = \left[\sum_{i=1}^{k} (Z_i^2 - 1) - \frac{k}{n} \sum_{i=1}^{n} (Z_i^2 - 1) \right] = -D_k \sum_{i=1}^{n} Z_i^2,$$

from (5.4) we have that:

$$\sqrt{\frac{n}{2}} |D_k| \frac{\sum_{i=1}^n Z_i^2}{n} = \left| X_n(\tau) - \tau X_n(1) - \frac{(n\tau - [n\tau])}{\sqrt{2n}} \xi_{[n\tau]+1} \right|. \tag{5.5}$$

As $n \to \infty$, $\Delta_n \to 0$ we get that $\frac{\sum_{i=1}^n Z_i^2}{n} \to 1$ and $\frac{(n\tau - [n\tau])}{\sqrt{2n}} \xi_{[n\tau]+1} \xrightarrow{p} 0$. Hence the thesis of Theorem 3.1 follows. □

Proof of Theorem 3.1. By the same arguments of Bai (1994), Sec. 3 and by using the formulas (10)–(14) therein, we have that:

$$|\hat{\tau}_0 - \tau_0| \le C_{\tau_0} (\theta_2 - \theta_1)^{-1} \sup_k |V_k - EV_k|, \tag{5.6}$$

where C_{τ_0} is a constant depending only on τ_0. Furthermore,

$$V_k - EV_k = \frac{1}{\sqrt{n}} \sqrt{\frac{k}{n}} \frac{1}{\sqrt{n-k}} \sum_{i=k+1}^{n} \left(Z_i^2 - \theta_2 \right) + \frac{1}{\sqrt{n}} \sqrt{1 - \frac{k}{n}} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} \left(Z_i^2 - \theta_1 \right);$$

then we can write

$$|V_k - EV_k| \le \frac{1}{\sqrt{n}} \left\{ \frac{1}{\sqrt{n-k}} \left| \sum_{i=k+1}^{n} \left(Z_i^2 - \theta_2 \right) \right| + \frac{1}{\sqrt{k}} \left| \sum_{i=1}^{k} \left(Z_i^2 - \theta_1 \right) \right| \right\}. \tag{5.7}$$

By applying Hajék–Renyi inequality for martingales we have that:

$$Pr\left\{ \max_{1 \le k \le n} \left| \frac{\sum_{i=1}^k (Z_i^2 - \theta_1)}{c_k} \right| > \alpha \right\} \le \frac{1}{\alpha^2} \sum_{k=1}^{n} \frac{E(Z_k^2 - \theta_1)^2}{c_k^2} = \frac{2\theta_1^2}{\alpha^2} \sum_{k=1}^{n} \frac{1}{c_k^2}. \tag{5.8}$$

Choosing $c_k = \sqrt{k}$ and observing that $\sum_{k=1}^n k^{-1} \le C \log n$, for some $C > 0$ (see, e.g., Bai, 1994), we have that:

$$\max_{1 \le k \le n} \frac{1}{\sqrt{k}} \sum_{i=1}^{k} Z_i = O_p\left(\sqrt{\log n}\right). \tag{5.9}$$

The same result holds for $\frac{1}{\sqrt{n-k}}\left|\sum_{i=k+1}^{n}(Z_i^2-\theta_2)\right|$; then from the relationships (5.7) and (5.9) we obtain the result (3.1). □

Proof of Theorem 3.2. We use the same framework of the proof of the Proposition 3 in Bai (1994), therefore we omit the details.

We choose a $\delta > 0$ such that $\tau_0 \in (\delta, 1-\delta)$. Since $\hat{k}_0/n$ is consistent for τ_0, for every $\varepsilon > 0$, $Pr\{\hat{k}_0/n \notin (\delta, 1-\delta)\} < \varepsilon$ when n is large. In order to prove (3.2), it is sufficient to show that $Pr\{|\hat{\tau}_0 - \tau_0| > M(n\vartheta_n^2)^{-1}\}$ is small when n and M are large, where $\vartheta_n = \theta_2 - \theta_1$. We are interested to study the behavior of V_k for $n\delta \le k \le n(1-\delta)$, $0 < \delta < 1$. We define for any $M > 0$ the set $D_{n,M} = \{k : n\delta \le k \le n(1-\delta), |k-k_0| > M\vartheta_n^{-2}\}$. Then we have that:

$$Pr\left\{|\hat{\tau}_0 - \tau_0| > M\left(n\vartheta_n^2\right)^{-1}\right\} \le \varepsilon + Pr\left\{\sup_{k\in D_{n,M}} |V_k| \ge |V_{k_0}|\right\},$$

for every $\varepsilon > 0$. Thus, we study the behavior of $Pr\{\sup_{k\in D_{n,M}} |V_k| \ge |V_{k_0}|\}$. It is possible to prove that:

$$\begin{aligned} Pr\left\{\sup_{k\in D_{n,M}} |V_k| \ge |V_{k_0}|\right\} &\le Pr\left\{\sup_{k\in D_{n,M}} V_k - V_{k_0} \ge 0\right\} + Pr\left\{\sup_{k\in D_{n,M}} V_k + V_{k_0} \le 0\right\} \\ &= P + Q. \end{aligned} \tag{5.10}$$

Furthermore,

$$\begin{aligned} Q &\le 2Pr\left\{\sup_{k\le n(1-\delta)} \frac{1}{n-k}\left|\sum_{i=k+1}^{n}\left(Z_i^2-\theta_2\right)\right| \ge \frac{1}{4}EV_{k_0}\right\} \\ &\quad + 2Pr\left\{\sup_{k\ge n\delta} \frac{1}{k}\left|\sum_{i=1}^{k}\left(Z_i^2-\theta_1\right)\right| \ge \frac{1}{4}EV_{k_0}\right\}. \end{aligned} \tag{5.11}$$

By observing that $\sum_{i=m}^{\infty} i^{-2} = O(m^{-1})$, the Hajék–Renyi inequality yields:

$$Pr\left\{\max_{k\ge m}\left|\frac{1}{k}\sum_{i=1}^{k}\left(Z_i^2-\theta_1\right)\right| > \alpha\right\} \le \frac{C_1}{\alpha^2 m}, \tag{5.12}$$

for some constant $C_1 < \infty$. The inequality (5.12) implies that (5.11) tends to zero as n tends to infinity.

Let $d(k) = \sqrt{((k/n)(1-k/n))}, k = 1, 2, \ldots, n$, for the first term in the right-hand of (5.10) we have that:

$$\begin{aligned} P &\le Pr\left\{\sup_{k\in D_{n,M}} \frac{n}{|k_0-k|}|G(k)| > \frac{\vartheta_n C_{\tau_0}}{2}\right\} + Pr\left\{\sup_{k\in D_{n,M}} \frac{n}{|k_0-k|}|H(k)| > \frac{\vartheta_n C_{\tau_0}}{2}\right\}, \\ &= P_1 + P_2, \end{aligned} \tag{5.13}$$

where

$$G(k) = d(k_0)\frac{1}{k_0}\sum_{i=1}^{k_0}\left(Z_i^2-\theta_1\right) - d(k)\frac{1}{k}\sum_{i=1}^{k}\left(Z_i^2-\theta_1\right) \tag{5.14}$$

$$H(k) = d(k)\frac{1}{n-k}\sum_{i=k+1}^{n}\left(Z_i^2 - \theta_2\right) - d(k_0)\frac{1}{n-k_0}\sum_{i=k_0+1}^{n}\left(Z_i^2 - \theta_2\right). \qquad (5.15)$$

We prove that P_1 tends to zero when n and M are large. Thus, we consider only $k \le k_0$ or more precisely those values of k such that $n\delta \le k \le n\tau_0 - M\vartheta_n^{-2}$. For $k \ge n\delta$, we have:

$$|G(k)| \le \frac{k_0 - k}{n\delta k_0}\left|\sum_{i=1}^{k_0}\left(Z_i^2 - \theta_1\right)\right| + B\frac{k_0 - k}{n}\frac{1}{n\delta}\left|\sum_{i=1}^{k}\left(Z_i^2 - \theta_1\right)\right| + \frac{1}{n\delta}\left|\sum_{i=k+1}^{k_0}\left(Z_i^2 - \theta_1\right)\right|, \qquad (5.16)$$

where $B \ge 0$ satisfies $|d(k_0) - d(k)| \le B|k_0 - k|/n$. By means of (5.8), (5.12), and (5.16), we obtain:

$$\begin{aligned} P_1 &\le Pr\left\{\frac{1}{n\tau_0}\left|\sum_{i=1}^{[n\tau_0]}\left(Z_i^2 - \theta_1\right)\right| > \frac{\delta\vartheta_n C_{\tau_0}}{6}\right\} \\ &\quad + Pr\left\{\sup_{1\le k\le n}\frac{1}{n}\left|\sum_{i=1}^{k}\left(Z_i^2 - \theta_1\right)\right| > \frac{\delta\vartheta_n C_{\tau_0}}{6B}\right\} \\ &\quad + Pr\left\{\sup_{k\le n\tau_0 - M\vartheta_n^{-2}}\frac{1}{n\tau_0 - k}\left|\sum_{i=k+1}^{[n\tau_0]}\left(Z_i^2 - \theta_1\right)\right| > \frac{\delta\vartheta_n C_{\tau_0}}{6}\right\} \\ &\le \frac{36\theta_1^2}{(\delta C_{\tau_0})^2\tau_0 n\vartheta_n^2} + \frac{36\theta_1^2 B^2}{(\delta C_{\tau_0})^2 n\vartheta_n^2} + \frac{36\theta_1^2}{\delta C_{\tau_0}^2 M}. \end{aligned}$$

When n and M are large, the last three terms are negligible. Analogously, we derive the proof of P_2. □

Proof of Theorem 3.3. The proof follows the same steps in Bai (1994, Theorem 1), hence we only sketch the parts of the proof that differ. We consider only $v \le 0$ because of symmetry. Let $K_n(v) = \{k : k = [k_0 + v\vartheta_n^{-2}], -M \le v \le 0, M > 0\}$ and

$$\Lambda_n(v) = n\left(V_k^2 - V_{k_0}^2\right) \qquad (5.17)$$

with $k \in K_n(v)$. We note that:

$$n\left(V_k^2 - V_{k_0}^2\right) = 2nEV_{k_0}(V_k - V_{k_0}) + 2n(V_{k_0} - EV_{k_0})(V_k - V_{k_0}) + n(V_k - V_{k_0})^2 \qquad (5.18)$$

The last two terms in (5.18) are negligible on $K_n(v)$. Since $\sqrt{n}(V_{k_0} - EV_{k_0})$ is bounded by (5.7), we have to show that $\sqrt{n}|V_k - V_{k_0}|$ is bounded. In particular, we can write:

$$\sqrt{n}|V_k - V_{k_0}| \le \sqrt{n}|G(k) + H(k)| + \sqrt{n}|EV_k - EV_{k_0}|,$$

where $G(k)$ and $H(k)$ are defined, respectively, in (5.14) and (5.15). The upper bound (5.16) is $o_p(1)$, because the first term is such that:

$$\sqrt{n}\frac{k_0-k}{n\delta k_0}\left|\sum_{i=1}^{k_0}(Z_i^2-\theta_1^2)\right| \leq \frac{M}{\delta\tau_0 n\vartheta_n^2}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{k_0}(Z_i^2-\theta_1)\right|$$
$$= \frac{O_p(1)}{n\vartheta_n^2} = o_p(1), \tag{5.19}$$

similarly for the second term and for the third term we apply the invariance principle (5.3). Now we explicit the limiting distribution for

$$2nEV_{k_0}(V_k - V_{k_0}) = 2\sqrt{\tau_0(1-\tau_0)}n\vartheta_n(V_{[k_0+v\vartheta_n^{-2}]} - V_{k_0}). \tag{5.20}$$

For simplicity we shall assume that $k_0 + v\vartheta_n^{-2}$ and $v\vartheta_n^{-2}$ are integers. We observe that:

$$n\vartheta_n(V_k - V_{k_0}) = n\vartheta_n(G(k)+H(k)) - n\vartheta_n(EV_{k_0} - EV_k), \tag{5.21}$$

where $G(k)$, $H(k)$ are defined in the expressions (5.14), (5.15). We can rewrite $G(k)$ as follows:

$$G(k) = d(k_0)\frac{k-k_0}{kk_0}\sum_{i=1}^{k_0}\left(Z_i^2-\theta_1\right) + \frac{d(k_0)-d(k)}{k}\sum_{i=1}^{k}\left(Z_i^2-\theta_1\right)$$
$$+ d(k_0)\frac{1}{k}\sum_{i=k+1}^{k_0}\left(Z_i^2-\theta_1\right). \tag{5.22}$$

By the same arguments used to prove (5.19), we can show that the first two terms in (5.22) multiplied by $n\vartheta_n$ are negligible on $K_n(M)$. Furthermore, $d(k_0) = \sqrt{\tau_0(1-\tau_0)}$ and $n/k \to 1/\tau_0$ for $k \in K_n(M)$, then we get that:

$$n\vartheta_n G\left(k_0+v\vartheta_n^{-2}\right) = n\vartheta_n d(k_0)\frac{1}{k}\sum_{i=k+1}^{k_0}(Z_i^2-\theta_1) + o_p(1)$$
$$= d(k_0)\frac{n}{k}\left\{\vartheta_n\sum_{i=1}^{|v|\vartheta_n^{-2}}\left(Z_{i+k}^2-\theta_1\right)\right\} + o_p(1)$$
$$\xrightarrow{d} \frac{\sqrt{(1-\tau_0)\tau_0}}{\tau_0}\sqrt{2}\theta_1 W_1(-v) \tag{5.23}$$

where in the last step we have used the invariance principle (5.3). Analogously, we can show that:

$$n\vartheta_n H(k_0+v\vartheta_n^{-2}) \xrightarrow{d} \frac{\sqrt{(1-\tau_0)\tau_0}}{1-\tau_0}\sqrt{2}\theta_1 W_1(-v). \tag{5.24}$$

Since

$$n\vartheta_n(EV_{k_0} - EV_k) \to \frac{|v|}{1\sqrt{\tau_0(1-\tau_0)}} \tag{5.25}$$

we obtain that:

$$\Lambda_n(v) \xrightarrow{d} 2\left\{\sqrt{2}\theta_1 W_1(-v) - \frac{|v|}{2}\right\}. \tag{5.26}$$

In the same way, for $v > 0$, we can prove that:

$$\Lambda_n(v) \xrightarrow{d} 2\left\{\sqrt{2}\theta_1 W_2(v) - \frac{|v|}{2}\right\}. \tag{5.27}$$

By applying the continuous mapping theorem and Theorem 3.2,

$$\frac{n\vartheta_n^2(\hat{\tau}_0 - \tau_0)}{2\tilde{\theta}^2} \xrightarrow{d} \frac{1}{2\theta_1^2} \arg\max_v \Lambda_n(v). \tag{5.28}$$

Since $aW(v) \stackrel{d}{=} W(a^2 v)$, $a \in \mathbb{R}$, a change in variable transforms $\arg\max_v \Lambda_n(v)$ into $2\theta_1^2 \arg\max_v \left\{W(v) - \frac{|v|}{2}\right\}$, which concludes the proof. □

Proof of Theorem 3.6. We start noticing that:

$$\begin{aligned}
&\sqrt{n}(\hat{\theta}_1(\hat{k}_0) - \hat{\theta}_1(k_0)) \\
&= \sqrt{n}\left(\frac{1}{\hat{k}_0}\sum_{i=1}^{\hat{k}_0} Z_i^2 - \frac{1}{k_0}\sum_{i=1}^{k_0} Z_i^2\right) \\
&= \mathbf{1}_{\{\hat{k}_0 \le k_0\}}\left(\sqrt{n}\frac{k_0 - \hat{k}_0}{k_0\hat{k}_0}\sum_{i=1}^{k_0}(Z_i^2 - \theta_1) - \sqrt{n}\frac{1}{\hat{k}_0}\sum_{i=\hat{k}_0}^{k_0}(Z_i^2 - \theta_1)\right) \\
&\quad + \mathbf{1}_{\{\hat{k}_0 > k_0\}}\left(\sqrt{n}\frac{k_0 - \hat{k}_0}{k_0\hat{k}_0}\sum_{i=1}^{k_0}(Z_i^2 - \theta_1) + \sqrt{n}\frac{1}{\hat{k}_0}\sum_{i=k_0}^{\hat{k}_0}(Z_i^2 - \theta_2) + \sqrt{n}\vartheta_n\frac{\hat{k}_0 - k_0}{\hat{k}_0}\right).
\end{aligned} \tag{5.29}$$

Since $k_0 = [\tau_0 n]$, $\hat{k}_0 = k_0 + O_p(\vartheta_n^{-2})$, and $n\vartheta_n^2 \to \infty$, we have that (5.29) is $(\sqrt{n}\vartheta_n)^{-1}O_p(1)$, which converges to zero in probability. Then $\hat{\theta}_1(\hat{k}_0)$ and $\hat{\theta}_1(k_0)$ have the same limiting distribution. Obviously, the same result holds for $\hat{\theta}_2$. Now it is easy to show that the limiting distribution of $\sqrt{n}(\hat{\theta}_1(k_0), \hat{\theta}_2(k_0))$ is equal to (3.6). □

Proof of Theorem 4.1. To prove our thesis it is sufficient to show that:

$$\sqrt{n}(\widehat{V}_k - \widehat{V}_{k_0} - (V_k - V_{k_0})) \xrightarrow{p} 0, \tag{5.30}$$

where V_k and V_{k_0} in this proof are the statistics obtained by setting $\sigma(X_t) = 1$ into Eq. (2.1). We rewrite the left-hand side of (5.30) as follows:

$$\sqrt{n}\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}\left[\frac{1}{n-k}(\widehat{S}_{n-k} - S_{n-k}) - \frac{1}{k}(\widehat{S}_k - S_k)\right]$$
$$- \sqrt{n}\sqrt{\frac{k_0}{n}\left(1-\frac{k_0}{n}\right)}\left[\frac{1}{n-k_0}(\widehat{S}_{n-k_0} - S_{n-k_0}) - \frac{1}{k_0}(\widehat{S}_{k_0} - S_{k_0})\right]. \qquad (5.31)$$

The first term of (5.31) is equal to:

$$\sqrt{n}\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}\left[\frac{1}{n-k}(\widehat{S}_{n-k} - S_{n-k}) - \frac{1}{k}(\widehat{S}_k - S_k)\right]$$
$$= \sqrt{n}\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}\left[\frac{1}{n-k}\sum_{i=k+1}^{n}(\widehat{Z}_i^2 - Z_i^2) - \frac{1}{k}\sum_{i=1}^{k}(\widehat{Z}_i^2 - Z_i^2)\right]$$
$$= \sqrt{\frac{n}{n-k}}\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}\frac{1}{\sqrt{n-k}}\sum_{i=k+1}^{n}(\widehat{Z}_i^2 - Z_i^2)$$
$$- \sqrt{\frac{n}{k}}\sqrt{\frac{k}{n}\left(1-\frac{k}{n}\right)}\frac{1}{\sqrt{k}}\sum_{i=1}^{k}(\widehat{Z}_i^2 - Z_i^2). \qquad (5.32)$$

We observe that:

$$\frac{1}{\sqrt{k}}\sum_{i=1}^{k}(\widehat{Z}_i^2 - Z_i^2) = \frac{2}{\sqrt{k}}\sum_{i=1}^{k} Z_i(\widehat{Z}_i - Z_i) + \frac{1}{\sqrt{k}}\sum_{i=1}^{k}(\widehat{Z}_i - Z_i)^2$$
$$= 2\mathcal{Z}_1 + \mathcal{Z}_2. \qquad (5.33)$$

The next step is to prove that (5.33) tends to zero in probability. In fact, by simple calculations we can write:

$$\mathcal{Z}_1 = \frac{\sqrt{\theta_1}\sqrt{\Delta_n}}{\sqrt{k}}\sum_{i=1}^{k}(b(X_i) - \hat{b}(X_i))\frac{W_{i+1} - W_i}{\sqrt{\Delta_n}}$$
$$= \frac{\sqrt{\theta_1}\sqrt{\Delta_n}}{\sqrt{k}}\sum_{i=1}^{k}\frac{(W_{i+1} - W_i)}{\sqrt{\Delta_n}}\frac{\sum_{j=1}^{n} K(\frac{X_j - X_i}{h_n})(b(X_i) - \frac{X_{j+1} - X_j}{\Delta_n})}{\sum_{j=1}^{n} K(\frac{X_j - X_i}{h_n})}.$$

Then we have that:

$$E(\mathcal{Z}_1)^2 = \frac{\theta_1\Delta_n}{k}E\left(\sum_{i=1}^{k}\frac{(W_{i+1} - W_i)}{\sqrt{\Delta_n}}\frac{\sum_{j=1}^{n} K(\frac{X_j - X_i}{h_n})(b(X_i) - \frac{X_{j+1} - X_j}{\Delta_n})}{\sum_{j=1}^{n} K(\frac{X_j - X_i}{h_n})}\right)^2.$$

We note the following fact:

$$
\begin{aligned}
&E(X_{j+1} - X_j - b(X_i)\Delta_n)^2 \\
&= E\left(\sqrt{\theta_1}(W_{j+1} - W_j) + \int_{t_j}^{t_{j+1}} [b(X_s) - b(X_i)]ds\right)^2 \\
&\leq 2\left\{\theta_1 E(W_{j+1} - W_j)^2 + E\left(\int_{t_j}^{t_{j+1}} [b(X_s) - b(X_i)]ds\right)^2\right\} \\
&= 2\left\{\theta_1 \Delta_n + E\left(\int_{t_j}^{t_{j+1}} [b(X_s) - b(X_i)]ds\right)^2\right\}. \qquad (5.34)
\end{aligned}
$$

Furthermore, the drift coefficient $b(\cdot)$ has continuous derivatives, therefore it is locally Lipschitz (see, e.g., Aït-Sahalia, 1996) and this bound follows:

$$
E\left(\int_{t_j}^{t_{j+1}} [b(X_s) - b(X_i)]ds\right)^2 = O\big(\Delta_n^2 \kappa_n^2\big) \qquad (5.35)
$$

where

$$
\kappa_n = \max_{j \leq n} \sup_{j\Delta_n \leq s \leq (j+1)\Delta_n} |X_s - X_j|.
$$

By Levy's modulus of continuity for diffusions (see Karatzas and Shreve, 1991),

$$
Pr\left\{\limsup_{\Delta_n \to 0} \frac{\kappa_n}{(\Delta_n \log(1/\Delta_n))^{1/2}} = C_0\right\} = 1
$$

with C_0 a suitable constant, we derive that:

$$
\kappa_n = O((\Delta_n \log(1/\Delta_n))^{1/2}). \qquad (5.36)
$$

Then by means of relationships (5.34), (5.35), and (5.36) we can claim that:

$$
\frac{X_{j+1} - X_j - b(X_i)\Delta_n}{\Delta_n} = O_p(1). \qquad (5.37)
$$

The expression (5.37) allows us to write

$$
E(\mathscr{L}_1)^2 \leq C\frac{\theta_1 \Delta_n}{k} E\left(\sum_{i=1}^{k} \frac{(W_{i+1} - W_i)}{\sqrt{\Delta_n}}\right)^2 = C\theta_1\Delta_n \to 0
$$

for some real constant C. The same arguments permit us to obtain that:

$$
\mathscr{L}_2 = \frac{\Delta_n}{\sqrt{k}} \sum_{i=1}^{k} (b(X_i) - \hat{b}(X_i))^2 \xrightarrow{p} 0.
$$

Similarly, we can develop the second term of (5.31), so the proof is complete. □

Acknowledgments

We thank an anonymous referee for his remarks. The research of the authors has been supported by MIUR grant F.I.R.B. RBNE03E3KF_004.

References

Aït-Sahalia, Y. (1996). Testing continuous-time models of the spot interest rate. *Rev. Fin. Stud.* 2:385–426.

Arnold, L. (1974). *Stochastic Differential Equations: Theory and Applications*. New York: Wiley.

Bai, J. (1994). Least squares estimation of a shift in linear processes. *J. Times Ser. Anal.* 15:453–472.

Bandi, F. M., Phillips, P. C. B. (2003). Fully nonparametric estimation of scalar diffusion models. *Econometrica* 71:241–283.

Chen, G., Choi, Y. K., Zhou, Y. (2005). Nonparametric estimation of structural change points in volatility models for time series. *J. Econometrics* 126:79–144.

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

Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge: Cambridge University Press.

Hinkley, D. V. (1971). Inference about the change-point from cumulative sum tests. *Biometrika* 58:509–523.

Inclan, C., Tiao, G. C. (1994). Use of cumulative sums of squares for retrospective detection of change of variance. *J. Amer. Statist. Assoc.* 89:913–923.

Karatzas, I., Shreve, S. E. (1991). *Brownian Motion and Stochastic Calculus*. New York: Springer-Verlag.

Kim, S., Cho, S., Lee, S. (2000). On the cusum test for parameter changes in GARCH(1, 1) models. *Commun. Statist. Theor. Meth.* 29:445–462.

Lee, S., Ha, J., Na, O., Na, S. (2003). The Cusum test for parameter change in time series models. *Scand. J. Statist.* 30:781–796.

Lee, S., Nishiyama, Y., Yoshida, N. (2006). Test for parameter change in diffusion processes by cusum statistics based on one-step estimators. *Ann. Inst. Statist. Mat.* 58:211–222.

Pagan, A., Ullah, A. (1999). *Nonparametric Statistics*. Cambridge: Cambridge University Press.

Silverman, B. W. (1986). *Density Estimation*. London: Chapman and Hall.