Nonparametric Combining Estimation of the Diffusion Coefficient of Diffusion Models *

YE XUGUO[1,2]　LING NENGXIANG[3]　YAO RENHAI[2]

([1]*Department of Mathematics, Southeast University, Nanjing, 210096*)

([2]*School of Mathematical Sciences, Kaili University, Kaili, 556011*)

([3]*School of Mathematics, Hefei University of Technology, Hefei, 230009*)

Abstract

In order to improve the accuracy of the diffusion coefficient estimation, we propose a new combining estimator to estimate the diffusion coefficient by dynamically integrating information from the time-domain and the state-domain. We find that the proposed estimator can effectively estimate the diffusion coefficient of diffusion models, as we show in this paper on simulated time series. Under certain conditions, the asymptotic normality is separately established for the proposed nonparametric estimators and the proposed theorem proves that the time-domain and state-domain estimators are asymptotically independent. Extensive simulations demonstrate the proposed estimator outperforms the other two estimators, and also outperforms the ones in the literature.

Keywords: Nonparametric estimation, diffusion coefficient, combining estimation, diffusion models.

AMS Subject Classification: 62G05, 62G20, 60J60.

§1. Introduction

Consider the problem of estimating the diffusion coefficient, $\sigma^2(\cdot)$, for a continuous-time diffusion process X_t satisfying the stochastic differential equation

$$\mathrm{d}X_t = \mu(X_t)\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \qquad t \in [0, T], \tag{1.1}$$

where W_t is a Wiener process on $[0, \infty]$. The function $\mu(\cdot)$ is a drift coefficient and $\sigma(\cdot)$ is referred to as a diffusion coefficient. As far as the model (1.1) is concerned, we know that it has wide applications in asset price, yields of bonds, financial markets; see the Vasicek

*The research was partly supported by the Guizhou Provincial Natural Science Foundation (LKK[2013]30, 2010076), the Project of Kaili University (Z1214, JG201313), Anhui Provincial Natural Science Foundation (1104 0606M03), the Project of Science and Research of Ministry of Education of China (309017).

Received September 10, 2011. Revised February 19, 2014.

doi: 10.3969/j.issn.1001-4268.2014.03.001

model (Vasicek, 1977), CIR model (Cox et al., 1985), CKLS model (Chan et al., 1992), the semiparametric model (Fan and Zhang, 2003), etc., for details.

The diffusion coefficient, most commonly known as spot volatility, plays a fundamental role in modern financial analysis. It is a measure of risk of a portfolio and is related to the Value-at-Risk, asset pricing, portfolio allocation, capital requirement and risk adjusted returns, among others. Various nonparametric estimators of the diffusion coefficient have been proposed in the finance literature, building on theoretical developments in the statistic literature. See, Florens-Zmirou (1993); Jiang and Knight (1997); Stanton (1997); Arfi (1998); Fan and Yao (1998); Jacod (2000); Hoffmann (1999); Fan and Zhang (2003); Bandi and Phillips (2003); Nicolau (2003); Aït-Sahalia and Mykland (2004); Jeffrey et al. (2004); Renò (2006); Arapis and Gao (2006); Bandi and Moloche (2008), etc..

For a given state variable X_t, most estimating methods are based on the assumption that the diffusion coefficient $\sigma^2(\cdot)$ depends on either price level or time level. But the economic conditions change from time to time. Thus, it is reasonable to expect that the diffusion coefficient prediction depends on both time and price level. For using more information to improve the accuracy of the diffusion coefficient estimation, how to combine information from both the time domain and state domain is a key problem. Some efforts have been made to propose the dependence of the estimator on both time and price level. Brenner, Harjes and Kroner (1996) proposed an interesting model on term interest rate that has some flavor of combining the time- and state-domain information in a parametric form. However, there is no formal work in the literature on efficiently integrating the time- and state-domain estimators. Recently, Fan et al. (2007) successfully proposed a integrated estimator to estimate the volatility function by weightedly integrating both time estimator and state estimator, and they pointed out that if the underlying process is continuous in time domain and stationary in state domain, such as model (1.1), both methods are applicable. Today, with advance of computer technology, the available data collection and storage are becoming more easy. The high-frequency financial data provide an incredible experiment more generally for analyzing financial markets. By using a two-step procedure in (2.1) — In the first step, estimate the time series $\widetilde{\sigma}_i$, which is implemented in the second step with equation (2.1), Renò (2008) constructed a nonparametric kernel estimator based on the combination of daily and high-frequency data.

On the one hand, in Renò's methodology, the advantage of estimating quadratic variation with intraday data is obviously the gain in precision. The proposed estimator is a fully nonparametric, in the sense that we impose very loose restrictions on the functional form of the drift term. They compared the proposed estimator with some existing methods in the literature, some results showed that the proposed estimator based on realized volatility is more precise. On the other hand, we should pay attention to the Fan's integrated method because of borrowing the strengths of both time- and state-domain estimators with aggregated information from the data. It is important to combine the estimators from the time-domain and state-domain separately, because this combination allows us to use more sampling information and to get better estimation. In fact, compared with the state-domain estimator, the integrated estimator will put more emphasis (weight) on recent data, in contrast with the time-domain estimator, it will use historical data to improve the efficiency. In short, two methods are two more effective approaches to nonparametrically estimate the diffusion coefficient. Whereas the former uses a dynamic weighting scheme to combine the two weekly dependent estimators in the diffusion estimation, the latter relies mainly on using a combination of daily and high-frequency data to construct a fully kernel estimator for estimating the diffusion coefficient.

Motivated by the empirical and theoretical success of combination estimation based on time-domain and state-domain in Fan et al. (2007) and based on combination of daily and high-frequency data in Renò (2008), we propose a new combining estimator and presents some new results. We use high-frequency data to estimate the diffusion coefficient in time domain and use daily data to estimate it in state domain. Then we define a new combining estimator by a dynamic integrated method of Fan et al. (2007). We find that the proposed estimator can effectively estimate the diffusion coefficient and some simulations demonstrate that the proposed estimator outperforms the ones in the literature, and also outperforms the other two estimators. Asymptotic normal behaviour as the time- and state-domain estimators is established under certain conditions, and we point out that the time-domain and state-domain estimators are asymptotically independent. This paper investigates how these two methods can be used together, and compares their combined use with the existing estimators. The main contribution of this paper is to propose a new combining estimator by combining the respective advantages of Fan's and Renò methods.

This paper is organized as follows. In Section 2, we firstly give some definitions for describing and illustrating our main results. Asymptotic normality is separately established for the proposed estimators. Section 3 present some simulations to evaluate the finite sample performance of the proposed estimator. We give some conclusions and the future work in Section 4. All the technical conditions and proofs are collected to Section 5.

§2. Methodology and Main Results

2.1 The Estimator are Based on the Estimation of Quadratic Variation between Observations

Suppose that we observe a discrete sample $\{\widehat{X}_{t_i}\}_{i=1}^{n}$ at n equally spaced time points from the diffusion process (1.1). $T = n\Delta$ is the time span of the sample, where Δ is the step size between observations and is fixed. We build our theory on Nadaraya-Watson estimators of the kind:

$$\widehat{\sigma}^2(x)=\sum_{i=1}^n K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big)\widetilde{\sigma}_i^2\Big/\sum_{i=1}^n K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big), \tag{2.1}$$

where $\widetilde{\sigma}_i$ is a consistent estimate of the volatility at time t_i and h_n is a bandwidth. The most popular fully nonparametric estimator of the diffusion coefficient is that proposed by Florens-Zmirou (1993). Her estimator is obtained by setting $\widetilde{\sigma}_i^2=(n/T)(\widehat{X}_{t_i}-\widehat{X}_{t_{i-1}})^2$ in (2.1), and is described by

$$\widehat{\sigma}^2_{\rm FZ}(x)=\sum_{i=1}^n K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big)\frac{n}{T}(\widehat{X}_{t_i}-\widehat{X}_{t_{i-1}})^2\Big/\sum_{i=1}^n K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big). \tag{2.2}$$

An estimator similar to that of Florens-Zmirou estimator has been proposed by Bandi and Phillips (2003). In the framework of (2.1), the estimator in Bandi and Phillips (2003) is obtained by setting

$$\widetilde{\sigma}_i^2=\frac{n}{T}\frac{1}{m_i}\sum_{j=0}^{m_i}(\widehat{X}_{t_{i,j+1}}-\widehat{X}_{t_{i,j}})^2, \tag{2.3}$$

where m_i is the number of times that $|\widehat{X}_{t_k}-\widehat{X}_{t_i}|\le\varepsilon_s$, ε_s is a parameter to be selected, $t_{i,j}$ is a subset of indexes such that

$$t_{i,0}=\inf\{k\ge 0:|\widehat{X}_{t_k}-\widehat{X}_{t_i}|\le\varepsilon_s\}\quad\text{and}\quad t_{i,j+1}=\inf\{t\ge t_{i,j}+\Delta:|\widehat{X}_{t_k}-\widehat{X}_{t_i}|\le\varepsilon_s\}.$$

A different estimator of the diffusion coefficient with the same asymptotic properties can be devised as follows: in the estimator (2.2), we replace $(n/T)(\widehat{X}_{t_i}-\widehat{X}_{t_{i-1}})^2$ with the integrated volatility in the interval $[t_{i-1},t_i]$, i.e. we set $\widetilde{\sigma}_i^2=(n/T)\displaystyle\int_{t_{i-1}}^{t_i}\sigma_s^2{\rm d}s$ in (2.1). This estimator is proposed by Renò (2008) and is defined by

$$\widehat{\sigma}^2_{\rm Renò}(x)=\sum_{i=1}^n K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big)\frac{n}{T}\int_{t_{i-1}}^{t_i}\sigma_s^2{\rm d}s\Big/\sum_{i=1}^n K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big). \tag{2.4}$$

Such an estimator would be unfeasible, but we can substitute the integrated variance with a consistent estimate of it, namely, realized volatility[1] (Barndorff-Nielsen and Shephard, 2002a, 2002b; Andersen et al., 2001a, 2001b, 2003). Based on the estimation of quadratic variation between observations by means of realized volatility, Renò (2008) proposed a new nonparametric estimator of the diffusion coefficient, and showed that it is consistent and asymptotically normally distributed. Define the local time as

$$L_t(x)=\lim_{\varepsilon\to 0}\frac{1}{2\varepsilon}\int_0^t I_{[x-\varepsilon,x+\varepsilon]}(X_s){\rm d}s.$$

[1]Renò (2008) use the realized volatility to estimate integrated volatility in (2.4), but we also can substitute it with the realized bi-power variation (See, Barndorff-Nielsen and Shephard, 2003, 2004a, 2004b, 2005, 2006; Andersen et al., 2004, 2007). In the simulations, we check out their performances. Specifically, Section 3 can be seen.

The local time of a diffusion is estimated by the following approximation

$$L_t^n(x)=\frac{T}{nh_n}\sum_{i=1}^{[nt/T]}K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big),$$

where $[x]$ is the integer part of x.

Lemma 2.1 (cf. Florens-Zmirou, 1993) If $n\to\infty$, we have $nh_n^4\to 0$, then $L_t^n(x)\to L_t(x)$ in the $\mathcal{L}^2$ sense. The convergence is almost sure if $\log n/(nh_n^2)\to 0$.

Let

$$V_n(x)=\frac{T}{nh_n}\sum_{i=1}^{n}K\Big(\frac{\widehat{X}_{t_i}-x}{h_n}\Big)\Big(\frac{\widehat{X}_{t_i}-\widehat{X}_{t_{i-1}}}{\sqrt{T/n}}\Big)^2,$$

we have

Lemma 2.2 (cf. Florens-Zmirou, 1993) If $nh_n^4\to 0$ as $n\to\infty$, then $V_n(x)$ converges to $\sigma^2(x)L_T(x)$ in the $\mathcal{L}^2$ sense.

Therefore, Renò (2008) proved that the proposed estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ is consistent and asymptotically normally distributed.

Lemma 2.3 (cf. Renò, 2008) If $nh_n^4\to 0$ as $n\to\infty$, then $\widehat{\sigma}^2_{\text{Renò}}(x)$ is consistent estimator of $\sigma^2(x)$ in the $\mathcal{L}^2$ sense.

Lemma 2.4 (cf. Renò, 2008) If $nh_n^3\to 0$, as $n\to\infty$, then

$$\sqrt{nh_n}\Big(\frac{\widehat{\sigma}^2_{\text{Renò}}(x)}{\sigma^2(x)}-1\Big)\xrightarrow{d}\frac{1}{\sqrt{L_T(x)}}N(0,1),$$

where the convergence $\xrightarrow{d}$ is in distribution.

2.2 The Estimator are Based on the Combining Estimation

Firstly, the strategy for combination is to introduce a dynamic weighting scheme W_t ($0\leq W_t\leq 1$), to combine the two weekly dependent estimators. Then we define the combining estimator based on time- and state-domain smoothing as,

$$\widehat{\sigma}^2_{I,(s,t)}=W_t\widehat{\sigma}^2_{t,\text{time}}+(1-W_t)\widehat{\sigma}^2_{t,\text{state}}. \tag{2.5}$$

Since $\widehat{\sigma}^2_{t,\text{time}}$ and $\widehat{\sigma}^2_{t,\text{state}}$ are asymptotically independent.[2] By minimizing the variance of the combining estimator, we can get the dynamic optimal weight,

$$W_t=\frac{\mathsf{Var}\,(\widehat{\sigma}^2_{t,\text{state}})}{\mathsf{Var}\,(\widehat{\sigma}^2_{t,\text{state}})+\mathsf{Var}\,(\widehat{\sigma}^2_{t,\text{time}})}. \tag{2.6}$$

[2]In Section 3, we prove that $\widehat{\sigma}^2_{t,\text{time}}$ and $\widehat{\sigma}^2_{t,\text{state}}$ are asymptotically independent, and thus they are close to independent in finite sample.

Estimated value of W_t is given by estimating unknown variances in (2.6). Therefore, we get the feasible integrated estimator

$$\widehat{\sigma}^2_{I,(s,t)} = \widehat{W}_t\widehat{\sigma}^2_{t,\text{time}} + (1-\widehat{W}_t)\widehat{\sigma}^2_{t,\text{state}}. \tag{2.7}$$

On the one hand, the time-domain method has been extensively studied in the literature (Robinson, 1997; Härdle et al., 2003; Fan et al., 2003; Mercurio and Spokoiny, 2004; Fan et al., 2007). These methods depend on the assumption that the coefficients are smooth so that they can be locally approximated by a constant. That's to say, the basic idea is to localizing in time, resulting in a time-domain smoothing. If high-frequency financial data can be collected in local time (a day), the availability of high-frequency intraday data allows us to accurately estimate the spot volatility (See, Foster and Nelson, 1996; Aït-Sahalia et al., 2005; Fan and Wang, 2008).

The quadratic variation of X_t has expression $[X,X]_t = \int_0^t \sigma_s^2 \mathrm{d}s$. Suppose that we observe X_t at N discrete time points $t_i = iT/N$, $i = 1, 2, \ldots, N$. Our goal is to estimate the spot volatility $\sigma_t^2 = \mathrm{d}[X,X]_t/\mathrm{d}t$. Suppose $M(\cdot)$ is a kernel with support on $[-1,1]$. We define the time-domain kernel type estimator as

$$\widehat{\sigma}^2_t(x) = \frac{1}{b}\sum_{t_i=t-b}^{t+b} M\Big(\frac{t_i-t}{b}\Big)(X_{t_i} - X_{t_{i-1}})^2,$$

where b is bandwidth. If $M(\cdot) = 1$, then the estimator results in a rolling average

$$\widehat{\sigma}^2_t(x) = \frac{1}{b}\sum_{t_i=t-b}^{t+b}(X_{t_i} - X_{t_{i-1}})^2 = \frac{1}{b}[X,X]_b.$$

If $M(\cdot)$ is one side kernel with support on $[-1,0]$, it yields an estimator that uses the immediate past data,

$$\widehat{\sigma}^2_{\mathrm{MA},t}(x) = \frac{1}{b}\sum_{t_i=t-b}^{t} M\Big(\frac{t_i-t}{b}\Big)(X_{t_i} - X_{t_{i-1}})^2. \tag{2.8}$$

Theorem 2.1 Under conditions (A.1)-(A.4), we have

$$\sqrt{Nb}(\widehat{\sigma}^2_{\mathrm{MA},t}(x) - \sigma_t^2) \xrightarrow{d} N\Big(0, 2\sigma_t^4 \cdot \int M^2(x)\mathrm{d}x\Big).$$

On the other hand, we exclude the N recent data points used in the time-domain estimation. Suppose that the historical data at time t are $\widehat{X}_{t_i}$, $i = 0, 1, \ldots, n$ from process (1.1) with a sampling interval Δ. Let $f(x) = \mu(x)\Delta^{1/2}$. By Euler approximation scheme, we get $\widehat{Y}_i \approx \mu(\widehat{X}_{t_i})\Delta^{1/2} + \sigma(\widehat{X}_{t_i})\varepsilon_i$, where $\widehat{Y}_i = \Delta^{-1/2}(\widehat{X}_{t_i} - \widehat{X}_{t_{i-1}})$ and $\varepsilon_i \sim_{\text{i.i.d.}} N(0,1)$ for $i = 1, 2, \ldots, n$. Denote the squared residuals by $\widehat{R}_i = \{\widehat{Y}_i - \widehat{f}(\widehat{X}_{t_i})\}^2$. Then the local linear estimator of f and σ_t^2 are $\widehat{f} = \widehat{a}$ and $\widehat{\sigma}^2_{\mathrm{LLE};s} = \widehat{\alpha}$ given by

$$(\widehat{a}, \widehat{b}) = \arg\min_{a} \sum_{i=1}^{n}\{\widehat{Y}_i - a - b(\widehat{X}_{t_i} - x)\}^2 U_h(\widehat{X}_{t_i} - x)$$

and

$$(\widehat{\alpha},\widehat{\beta})=\arg\min_{\alpha,\beta}\sum_{i=1}^{n}\{\widehat{R}_i-\alpha-\beta(\widehat{X}_{t_i}-x)\}^2W_h(\widehat{X}_{t_i}-x),$$

where $U_h(\cdot)$ and $W_h(\cdot)$ are kernel functions, h is bandwidth. Then the local linear estimator can be expressed as

$$\widehat{\sigma}^2_{\mathrm{LLE},s}=\sum_{i=1}^{n}\omega_{2,i}(x)\widehat{R}_i, \tag{2.9}$$

where $\omega_{2,i}(x)=W((\widehat{X}_{t_i}-x)/h)(S_2(x)-(\widehat{X}_{t_i}-x)S_1(x))/(S_0(x)S_2(x)-S_1^2(x))$ and $S_j(x)=\sum\limits_{i=1}^{n}(\widehat{X}_{t_i}-x)^jW((\widehat{X}_{t_i}-x)/h)$.

Remark 1 Stanton (1997) and Fan et al. (2003) showed that $\widehat{Y}_i^2$ instead of $\widehat{R}_i$ in (2.9) also can be used for the estimation of σ_t^2. Moreover, the variance of the estimator can be approximated as $\mathsf{Var}\,(\widehat{\sigma}^2_{\mathrm{LLE},s}|\widehat{X}_{t_i})\approx 2\sigma_t^4\sum\limits_{i=1}^{n}\omega_{2,i}^2$.

Theorem 2.2 Let $v_j=\displaystyle\int u^jK^2(u)\mathrm{d}u$ for $j=0,1,2$. Suppose that the second derivatives $\mu(\cdot)$ and $\sigma^2(\cdot)$ exist in a neighborhood of x, then

$$\sqrt{nh}(\widehat{\sigma}^2_{\mathrm{LLE},s}-\sigma_t^2-\theta_n)\xrightarrow{d}N(0,2\sigma_t^4p^{-1}(x)e_1^T(H^{-1}S^{-1})^TS^*H^{-1}S^{-1}e_1),$$

where $\theta_n=(1/2)h^2\ddot{\sigma}^2(x)\sigma_w^2+o(h^2)$ and $\sigma_w^2=\displaystyle\int u^2w(u)\mathrm{d}u$.

The asymptotic independence between the time- and state-domain estimators is the basic condition of the combining method. The following theorem shows that both the time- and state-domain estimators are asymptotically independent and the combining estimator $\widehat{\sigma}^2_{I,(s,t)}$ exists the asymptotic normality.

Theorem 2.3 Suppose that the second derivatives $\mu(\cdot)$ and $\sigma^2(\cdot)$ exist in a neighborhood of x. Under conditions (A.1)-(A.4) and (B.1)-(B.4). Then we present result at the current time t_N.

The asymptotic independence of both time-domain estimator and state-domain estimator

$$(\sqrt{Nb}(\widehat{\sigma}^2_{\mathrm{MA},t}(x)-\sigma^2_{t_N}),\sqrt{nh_2}(\widehat{\sigma}^2_{\mathrm{LLE},s}-\sigma^2_{t_N}-\theta_n))^T\xrightarrow{d}N\left(\begin{pmatrix}0\\0\end{pmatrix},\begin{pmatrix}V_1&0\\0&V_2\end{pmatrix}\right),$$

where $V_1=2\sigma^4_{t_N}\cdot\displaystyle\int_{-1}^{1}M^2(x)\mathrm{d}x$ and $V_2=2\sigma^4_{t_N}p^{-1}(x)e_1^T(H^{-1}S^{-1})^TS^*H^{-1}S^{-1}e_1$.

The asymptotic normality of $\widehat{\sigma}^2_{I,(s,t_N)}$: if the limit $d=\lim\limits_{n\to\infty}N/nh_2$ exists, then

$$\sqrt{nh_2/S^2}[\widehat{\sigma}^2_{I,(s,t_N)}-\sigma^2_{t_N}]\xrightarrow{d}N(0,1),$$

where $S^2=W_t^2V_1/d+(1-W_t)^2V_2$.

Based on the asymptotic independence between the time- and state-domain estimators, an dynamic weight of the integrated estimator is given by (2.6). Because that $\widehat{\sigma}^2_{\mathrm{MA},t_N}(x)$ and $\widehat{\sigma}^2_{\mathrm{LLE},s}$ are consistent, we have

$$\widehat{W}_t = \frac{\widehat{\mathsf{Var}}\,(\widehat{\sigma}^2_{\mathrm{LLE},s})}{\widehat{\mathsf{Var}}\,(\widehat{\sigma}^2_{\mathrm{LLE},s}) + \widehat{\mathsf{Var}}\,(\widehat{\sigma}^2_{\mathrm{MA},t_N}(x))}$$

Remark 2 Note that the theoretically optimal weight minimizing the variance in Theorem 2.3 is $W_{t,\mathrm{opt}} = V_2/(V_1/d + V_2)$.

Thus, this results in the dynamically feasible combining estimator

$$\widehat{\sigma}^2_{I,(s,t_N)} = \widehat{W}_{t_N} * \widehat{\sigma}^2_{\mathrm{MA},t_N} + (1 - \widehat{W}_{t_N}) * \widehat{\sigma}^2_{\mathrm{LLE},s}. \tag{2.10}$$

§3. Numerical Analysis

To facilitate the presentation, we use the simple abbreviations in Table 1 to denote six diffusion coefficient estimators.

Table 1 Six diffusion coefficient estimators

$\widehat{\sigma}^2_{\mathrm{FZ}}(x)$: An popular nonparametric estimator in (2.2)
$\widehat{\sigma}^2_{\mathrm{BP}}(x)$: An improved nonparametric estimator in (2.3)
$\widehat{\sigma}^2_{\text{Renò}}(x)$: The new fully nonparametric estimator in (2.4)
$\widehat{\sigma}^2_{\mathrm{MA},t}(x)$: The time-domain estimation method in (2.8)
$\widehat{\sigma}^2_{\mathrm{LLE},s}(x)$: The state-domain estimation method in (2.9)
$\widehat{\sigma}^2_{I,(s,t_N)}(x)$: The new integrated estimator in (2.10)

The following five measures are employed to assess the performance of different procedures for estimating the diffusion coefficient.

Measure 1: Mean Error (ME):

$$\mathrm{ME} = \frac{1}{m}\sum_{i=1}^{m}(\widehat{\sigma}^2_i - \sigma^2_t).$$

Measure 2: Root Mean Square Error (RMSE):

$$\mathrm{RMSE} = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(\widehat{\sigma}^2_i - \sigma^2_t)^2}.$$

Measure 3: Ideal Mean Absolute Deviation Error (IMADE):

$$\mathrm{IMADE} = \frac{1}{m}\sum_{i=1}^{m}|\widehat{\sigma}^2_i - \sigma^2_t|.$$

Measure 4: Ideal Square Root Absolute Deviation Error (IRADE):

$$\mathrm{IRADE}=\frac{1}{m}\sum_{i=1}^{m}|\widehat{\sigma}_i-\sigma_t|.$$

Measure 5: Relative Ideal Mean Absolute Deviation Error (RIMADE):

$$\mathrm{RIMADE}=\frac{1}{m}\sum_{i=1}^{m}\frac{|\widehat{\sigma}_i^2-\sigma_t^2|}{\sigma_t^2}.$$

Generally speaking, the smaller the calculated value, the better the estimated approach.

Example 1 We use the Cox-Ingersoll-Ross model (CIR) as our data-generating process

$$\mathrm{d}X_t=k(\alpha-X_t)\mathrm{d}t+\sigma X_t^{1/2}\mathrm{d}W_t,\qquad t\geq t_0,\tag{3.1}$$

where the spot rate, X_t, moves around its long-run equilibrium level α at speed k. When the condition $2k\alpha\geq\sigma^2$ holds, this process is shown to be positive and stationary. We simulate the sample paths of the process by using the Euler scheme. In our implementation, the values of the model parameters are cited from Fan and Zhang (2003), that is, $k=0.21459$, $\alpha=0.08571$, $\sigma=0.07830$. Throughout the paper we simulate daily observations, that is $\Delta=1/252$, where the number of days per year is assumed to be 252.

The estimator (2.4) is implemented as follows. The distance between two adjacent observations is 1/252 (daily data). For each observation at time t_i, we divide this interval into m steps, and we estimate the integrated volatility by using realized volatility or realized bi-power variation (intraday data) as follows:

$$\begin{aligned}\widetilde{\sigma}_i^2 &= \frac{n}{T}\sum_{j=0}^{m-1}\big(X_{i+(j+1)/m}-X_{i+j/m}\big)^2\\ &\quad\text{or}\ \frac{n}{T}\mu_1^{-2}\frac{n}{n-1}\sum_{j=0}^{m-2}\big|X_{i+(j+1)/m}-X_{i+j/m}\big|\big|X_{i+(j+2)/m}-X_{i+(j+1)/m}\big|,\end{aligned}\tag{3.2}$$

where $\mu_a=\mathsf{E}(|Z|^a)$ and $Z\sim N(0,1)$ $(a>0)$. For the estimator (2.4), we compute (3.2) with $m=500$.

The estimator (2.10) is implemented as follows. On each simulated sample path, we estimate σ_t^2 over $T=1$ day (i.e., $T=N\Delta_t=1/252$, since the parameter values are all annualized) by using $[Y,Y]_b^{(\mathrm{all})}$. We use daily data to estimate it in state domain. That is to say, we set the first N observations as the time-domain data and the last n observations as the state-domain data, and the start value $X_{t_0}=0.1$. We focus on an interior state point $x=0.1$. A simple rule of thumb bandwidth formula is used (Scott, 1992): $h_n=h_s\widehat{\sigma}n^{-1/5}$, where h_s is a real constant (here $h_s=1.06$), and $\widehat{\sigma}$ is the sample standard deviation, and the gaussian density function is a common kernel $K(s)=(1/\sqrt{2\pi})\mathrm{e}^{-s^2/2}$. The experimental results are based on 1,000 replications.

Table 2 Simulated results for the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$

$\widehat{\sigma}^2_{\text{Renò}}(x)$	Measure				
$\widetilde{\sigma}_i^2$	ME	RMSE	IMADE	IRADE	RIMADE
Based on RV	-1.3419×10^{-6}	1.2149×10^{-5}	8.6622×10^{-6}	1.7506×10^{-4}	1.4129×10^{-2}
Based on BPV	-2.5028×10^{-6}	1.2567×10^{-5}	8.9990×10^{-6}	1.8204×10^{-4}	1.4678×10^{-2}

The results show that the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ based on realized volatility outperforms based on realized bi-power variation. Moreover, Monte Carlo simulations show that the precision of the high-frequency estimator increases with m, as we expected. In reality, there are two problems. One is the disadvantage of this estimator is that it requires much more data. Every $\widetilde{\sigma}_i^2$ prediction is based on the intraday data. Other one is taking too much computational time. When using the more days, the greater the time consumption.

Table 3 Simulated results for the six estimators

Method	Measure				
	ME	RMSE	IMADE	IRADE	RIMADE
$\widehat{\sigma}^2_{\text{FZ}}(x)$	-3.1199×10^{-5}	1.0914×10^{-4}	8.9035×10^{-5}	1.8476×10^{-3}	1.4522×10^{-1}
$\widehat{\sigma}^2_{\text{LLE},s}(x)$	-1.3718×10^{-5}	7.5369×10^{-5}	5.5776×10^{-5}	1.1441×10^{-3}	9.0975×10^{-2}
$\widehat{\sigma}^2_{\text{BP}}(x)$	-3.4813×10^{-5}	6.6186×10^{-5}	5.0805×10^{-5}	1.0561×10^{-3}	8.2868×10^{-2}
$\widehat{\sigma}^2_{\text{Renò}}(x)$	-2.0127×10^{-6}	1.2745×10^{-5}	9.0184×10^{-6}	1.8239×10^{-4}	1.4710×10^{-2}
$\widehat{\sigma}^2_{\text{MA},t}(x)$	3.9242×10^{-7}	7.9969×10^{-6}	6.3762×10^{-6}	1.2870×10^{-4}	1.0400×10^{-2}
$\widehat{\sigma}^2_{\text{Int},(s,t)}(x)$	9.0786×10^{-8}	7.9376×10^{-6}	6.3006×10^{-6}	1.2721×10^{-4}	1.0277×10^{-2}

To examine the efficiency of the integrated estimator, we simulate the price as discussed above. The results, summarized in Table 3, shows that the performance of our proposed combining estimator uniformly dominates the other estimators because of its lowest ME, RMSE, IMADE, IRADE and RIMADE. The results show that the proposed combining estimator performs closely to the time-domain method. This is mainly because high-frequency data is used so that the time-domain is far more informative than the state-domain, supporting our statement at the end of the first paragraph in Section 1. Due to the estimated method weightedly emphasis on daily realized volatility, the results demonstrate that Renò's estimator do not seem to produce more efficient estimation results than the others. Our proposed integrated method features high speed of convergence and precision and is more reasonable compared with others, but less than the use of data in Renò's estimator.

Remark 3 In fact, we also consider the Vasicek model (Vasicek, 1977) as our data-generating process: $\mathrm{d}X_t = k(\alpha - X_t)\mathrm{d}t + \sigma \mathrm{d}W_t$. For this model the diffusion coefficient is constant and equal to σ^2. Set $\alpha = 8.3\%$, $k = 0.5$ and $\sigma = 3\%$. These parameters values are given by Renò et al. (2006). Unreported simulation results also show that the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ based on realized volatility outperforms based on realized bi-power variation. There is no substantial difference on the performance of all estimators in the Vasicek model and the CIR model.

Example 2 We consider the geometric Brownian model (GBM)

$$\mathrm{d}X_t = (\mu + \sigma^2/2)X_t\mathrm{d}t + \sigma X_t\mathrm{d}B_t. \tag{3.3}$$

This model (3.3) is a non-stationary process to which we check if our method continues to apply. Note that the celebrated Black-Scholes option price formula is derived based on the Osborne's assumptions that the stock price follows the GBM model. We simulate 1,000 times with $\Delta = 1/252$, the corresponding approximate process with parameters $\mu = 0.087$ and $\sigma = 0.178$, starting at $X_{t_0} = 1.0$ (Fan and Zhang, 2003). We choose the Euler approximation scheme to simulate the process (3.3). For each scheme, 1,000 sample paths of length 1,000 are generated. We focus on an interior state point $x = 1.0$ and the bandwidth parameter $h_n = h_s\widehat{\sigma}n^{-1/5}$, where h_s is a real constant, and $\widehat{\sigma}$ is the sample standard deviation.

Table 4 Simulated results for the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ and the normal kernel

Bandwidth	$\widehat{\sigma}^2_{\text{Renò}}(x)$	Measure				
h_s	$\widetilde{\sigma}^2_i$	ME	RMSE	IMADE	IRADE	RIMADE
1.06	RV	1.2721×10^{-3}	2.6284×10^{-3}	1.6728×10^{-3}	4.5713×10^{-3}	5.2796×10^{-2}
	BPV	1.2105×10^{-3}	2.5972×10^{-3}	1.6490×10^{-3}	4.5087×10^{-3}	5.2047×10^{-2}
3	RV	5.2023×10^{-3}	8.5507×10^{-3}	6.0348×10^{-3}	1.5708×10^{-2}	1.9047×10^{-1}
	BPV	5.1299×10^{-3}	8.4943×10^{-3}	5.9882×10^{-3}	1.5596×10^{-2}	1.8900×10^{-1}
5	RV	1.0119×10^{-2}	1.6274×10^{-2}	1.1525×10^{-2}	2.8485×10^{-2}	3.6373×10^{-1}
	BPV	1.0037×10^{-2}	1.6202×10^{-2}	1.1465×10^{-2}	2.8354×10^{-2}	3.6186×10^{-1}

For a given kernel function, the choice of an effective bandwidth parameter is very important to the performance of all nonparametric kernel estimator. For different values of h_s, the results are summarized in Table 4, which shows that the performance of the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ based on realized bi-power variation outperforms based on realized volatility because of its lower ME, RMSE, IMADE, IRADE and RIMADE. These results show that the performance of the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ in Table 2 is opposite because of the non-stationarity of the process. Furthermore, we find that $h_s = 1.06$ is highly suitable for estimating the diffusion coefficient.

It is widely recognized that the choice of the kernel function is much less important than the choice of an appropriate smoothing parameter. But in Table 5, we find that the performance of $\widehat{\sigma}^2_{\text{Renò}}(x)$ with Epanechnikov kernel outperforms it with normal kernel.

Table 5 Simulated results for the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ and the Epanechnikov kernel $K(u)=0.75(1-u^2)I(|u|\le 1)$

Bandwidth	$\widehat{\sigma}^2_{\text{Renò}}(x)$	Measure				
h_s	$\widehat{\sigma}^2_i$	ME	RMSE	IMADE	IRADE	RIMADE
1.06	RV	4.7519×10^{-4}	1.2136×10^{-3}	7.1224×10^{-4}	1.9734×10^{-3}	2.2479×10^{-2}
	BPV	4.1594×10^{-4}	1.1993×10^{-3}	7.0966×10^{-4}	1.9677×10^{-3}	2.2398×10^{-2}
3	RV	2.0657×10^{-3}	3.9695×10^{-3}	2.5272×10^{-3}	6.8119×10^{-3}	7.9763×10^{-2}
	BPV	1.9974×10^{-3}	3.9280×10^{-3}	2.4908×10^{-3}	6.7171×10^{-3}	7.8615×10^{-2}
5	RV	3.9680×10^{-3}	6.8042×10^{-3}	4.5870×10^{-3}	1.2074×10^{-2}	1.4477×10^{-1}
	BPV	3.8969×10^{-3}	6.7541×10^{-3}	4.5436×10^{-3}	1.1966×10^{-2}	1.4340×10^{-1}

In order to obtain the performance of each estimator in the literature, all estimators are used to estimate the diffusion coefficient of the GBM model. The results are summarized in Tables 6 and 7.

Table 6 Simulated results for the six estimators and the normal kernel

Method	Measure				
	ME	RMSE	IMADE	IRADE	RIMADE
$\widehat{\sigma}^2_{\text{BP}}(x)$	-1.2782×10^{-2}	1.2954×10^{-2}	1.2796×10^{-2}	4.0752×10^{-2}	4.0388×10^{-1}
$\widehat{\sigma}^2_{\text{FZ}}(x)$	1.6266×10^{-3}	5.5479×10^{-3}	3.5327×10^{-3}	9.5701×10^{-3}	1.1150×10^{-1}
$\widehat{\sigma}^2_{\text{LLE},s}(x)$	1.3039×10^{-3}	6.4728×10^{-3}	3.7006×10^{-3}	9.9748×10^{-3}	1.1680×10^{-1}
$\widehat{\sigma}^2_{\text{Renò}}(x)$	1.2582×10^{-3}	2.5878×10^{-3}	1.6449×10^{-3}	4.4956×10^{-3}	5.1916×10^{-2}
$\widehat{\sigma}^2_{\text{MA},t}(x)$	4.1041×10^{-5}	5.0724×10^{-4}	4.0375×10^{-4}	1.1333×10^{-3}	1.2743×10^{-2}
$\widehat{\sigma}^2_{\text{Int},(s,t)}(x)$	4.6800×10^{-5}	5.0459×10^{-4}	4.0231×10^{-4}	1.1292×10^{-3}	1.2697×10^{-2}

Table 7 Simulated results for the six estimators and the Epanechnikov kernel $K(u)=0.75(1-u^2)I(|u|\le 1)$

Method	Measure				
	ME	RMSE	IMADE	IRADE	RIMADE
$\widehat{\sigma}^2_{\text{BP}}(x)$	-1.2880×10^{-2}	1.3338×10^{-2}	1.3014×10^{-2}	4.1765×10^{-2}	4.1074×10^{-1}
$\widehat{\sigma}^2_{\text{FZ}}(x)$	1.5154×10^{-3}	6.7970×10^{-3}	4.5191×10^{-3}	1.2307×10^{-2}	1.4263×10^{-1}
$\widehat{\sigma}^2_{\text{LLE},s}(x)$	1.4196×10^{-3}	7.8987×10^{-3}	4.9018×10^{-3}	1.3300×10^{-2}	1.5471×10^{-1}
$\widehat{\sigma}^2_{\text{Renò}}(x)$	3.5770×10^{-4}	1.0410×10^{-3}	6.6232×10^{-4}	1.8422×10^{-3}	2.0904×10^{-2}
$\widehat{\sigma}^2_{\text{MA},t}(x)$	2.4492×10^{-5}	5.3035×10^{-4}	4.2390×10^{-4}	1.1901×10^{-3}	1.3379×10^{-2}
$\widehat{\sigma}^2_{\text{Int},(s,t)}(x)$	2.7013×10^{-5}	5.2925×10^{-4}	4.2264×10^{-4}	1.1866×10^{-3}	1.3340×10^{-2}

The results show that the performance of the proposed integrated estimator is more prominent than other estimators. It is easy to show that the the proposed integrated estimator is effective. Moreover, the estimator $\widehat{\sigma}^2_{\text{MA},t}(x)$ also performs better than the other estimators. But we find that the estimator $\widehat{\sigma}^2_{\text{Renò}}(x)$ result in a poor estimation may because of the non-stationarity of the process. The conclusion similar to Example 1 can be drawn from this example. This shows that our integrated method continues to perform better than the others for this non-stationary case. When we use the Epanechnikov kernel to replace the the normal kernel in the simulation study, we find that the results not to change, as seen in Table 7.

§4. Conclusions

This paper introduces nonparametric estimation of the diffusion coefficient of diffusion models. A new dynamical combining estimator to aggregate the information from the time-domain and state-domain is proposed and studied. The performance of the proposed estimator is assessed on simulations of several popular diffusion models. Some simulations illustrated that the proposed combining estimator is effectively aggregating the information from both the time and the state domains, and has advantages over some previous methods. Our study has also revealed that proper use of information from both the time-domain and state-domain makes volatility forecasting more accurately. Our method exploits the continuity in the time-domain and stationarity in the state-domain.

The results briefly presented in this paper are encouraging, but they are also preliminary in many respects. First, a serious limitation to this approach is that the assumption of a continuous sample path for asset prices may be too restrictive. We can extend the estimators to allow for jumps in the equation driving the observable variable, as well as to estimate leverage when the observable variable and the latent volatility factor are correlated. Second, the choice of the bandwidth parameter can be refined, for example using automated techniques instead of the simple rule of thumb adopted here. Third, one important limitation is that they have been studied for single-factor models only. In practice, it is well known that single-factor models are too naive, both for stock prices and spot rate modeling. Finally, using intraday data to directly implement the estimator proposed by Florens-Zmirou (1993) could be misleading, since intraday data display pronounced seasonalities and microstructure effects that could seriously distort the estimation.

Appendix

We firstly introduce the following technical conditions which are necessary in the proofs. Throughout the proofs, we denotes by C a generic constant.

(A.1) The drift term μ_t and diffusion term σ_t in (1.1) satisfy

$$\sup\{|\mu_t-\mu_s|,|t-s|\le c\}=O_p(c^{1/2}|\log c|^{1/2})$$

and

$$\sup\{|\sigma_t-\sigma_s|,|t-s|\le c\}=O_p(c^{1/2}|\log c|^{1/2}),$$

for any $t,s\in[0,T]$ and c is a positive constant.

(A.2) $\sup\Big\{\Big|\int_{t_{i-1}}^{t_i}(\sigma_s-\sigma_{t_{i-1}})\mathrm{d}W_s\Big|^2, i=1,\ldots,n\Big\}=O_p(N^{-2+c})$, where c is an arbitrarily small positive number.

(A.3) $b\backsim N^{-1/2}/\log N$, $M(\cdot)$ is twice differentiable with support $[-1,1]$ and

$$\int_{-1}^{1}M(x)\mathrm{d}x=1.$$

(A.4) There exists a positive constant C, such that

$$\mathsf{E}|\mu(X_s)|^{2(p+\delta)}\le C\quad\text{and}\quad\mathsf{E}|\sigma(X_s)|^{2(p+\delta)}\le C,\qquad\text{for any } s\in[t-\eta,t],$$

where η is some positive constant, p is an integer not less than 2 and $\delta>0$.

(B.1) The discrete observations $\{\widehat{X}_{t_i}\}_{i=0}^{n}$ satisfy the stationarity condition of Banon (1978). Furthermore, a stationary process X_t is said to satisfy the condition $G_2(s,\alpha)$ of Rosenblatt (1970).

(B.2) The conditional density $p_l(y|x)$ of $\widehat{X}_{t_{i+l}}$ for given $\widehat{X}_{t_i}$ is continuous in the arguments (x,y) and is bounded by a constant (independent of l).

(B.3) The kernel functions $U_h(\cdot)$ and $W_h(\cdot)$ are bounded, symmetric probability density function with compact support $[-1,1]$. Furthermore, they are continuously differentiable function.

(B.4) As $n\to\infty$, $nh\to\infty$ and $nh^5\to0$ and $nh\Delta\to0$.

Proof of Theorem 2.1 The proof is completed by using the same lines in Fan and Wang (2008). □

Proof of Theorem 2.2 Without loss of generality, we assume that $f(x)=0$, hence $\widehat{R}_i=\widehat{Y}_i^2$. Let

$$Y=(\widehat{Y}_1^2,\widehat{Y}_2^2,\ldots,\widehat{Y}_n^2),\qquad W=\mathrm{diag}\Big\{W\Big(\frac{\widehat{X}_{t_1}-x}{h}\Big),W\Big(\frac{\widehat{X}_{t_2}-x}{h}\Big),\ldots,W\Big(\frac{\widehat{X}_{t_n}-x}{h}\Big)\Big\}$$

and

$$X=\begin{pmatrix}1 & \widehat{X}_{t_1}-x\\ \vdots & \vdots\\ 1 & \widehat{X}_{t_n}-x\end{pmatrix}.$$

Denote by $m_i=\mathsf{E}[\widehat{Y}_i^2|X_{t_i}]$, $m=(m_1,m_2,\ldots,m_n)^T$ and $e_1=(1,0)^T$. Then it can be written that

$$\widehat{\sigma}^2_{\mathrm{LLE},s}=e_1^T(X^TWX)^{-1}X^TWY,$$

$$\begin{aligned}\widehat{\sigma}^2_{\text{LLE},s} - \sigma_t^2 &= e_1^T(X^TWX)^{-1}X^TW\{m - X\beta_N\} + e_1^T(X^TWX)^{-1}X^TW\{Y - m\} \\ &= e_1^TB + e_1^Tb,\end{aligned}$$

where $\beta_N = (m(x), m'(x))^T$ with $m(x) = \mathsf{E}[\widehat{Y}_1^2|\widehat{X}_{t_i} = x]$. By Fan and Yao (1998), the bias vector B converges in probability to a vector B with $B = O(h^2) = o(1/\sqrt{nh})$. In the following, we will show that the centralized vector b is asymptotically normal.

Put $u = n^{-1}H^{-1}X^TW(Y - m)$, where $H = \text{diag}\{1, h\}$, then by Fan and Yao (2003) the vector b can be written as

$$b = p^{-1}(x)H^{-1}S^{-1}u(1 + o_p(1)), \tag{5.1}$$

where $S = (\mu_{i+j-2})$, $i, j = 1, 2$ with $\mu_j = \displaystyle\int u^j k(u)\text{d}u$.

For any constant vector c, we define

$$Q_n = c^Tu = \frac{1}{2}\sum_{i=1}^{n}\{\widehat{Y}_i^2 - m_i\}C_h(\widehat{X}_{t_i} - x),$$

where $C(u) = c_1W(u) + c_2uW(u)$ with $C_u(h) = C(u/h)/h$.

Applying the "big-block" and "small-block" arguments in Fan and Yao (2003, Theorem 6.3), we obtain

$$\theta^{-1}(x)\sqrt{nh}Q_n \to N(0, 1), \tag{5.2}$$

where $\theta^2(x) = 2p(x)\sigma_t^4\displaystyle\int C^2(u)\text{d}u$. Therefore, we have

$$\sqrt{nh}c_Tu \to N\Big(0, 2p(x)\sigma_t^4\int C^2(u)\text{d}u\Big).$$

Because that Q_n is a linear transform of u,

$$\sqrt{nh}u \to N(0, 2\sigma_t^4p(x)S^*/nh),$$

where $S^* = (v_{i+j-2})$, $i, j = 1, 2$ with $v_j = \displaystyle\int u^jK^2(u)\text{d}u$.

We can reduce to $u \to N(0, 2\sigma_t^4p(x)S^*/nh)$. Because that $b = p^{-1}(x)H^{-1}S^{-1}u(1 + o_p(1))$, we have

$$b \to N(0, 2\sigma_t^4p(x)^{-1}(H^{-1}S^{-1})^TS^*H^{-1}S^{-1}/nh).$$

Thus, we have $\widehat{\sigma}^2_{\text{LLE},s} - \sigma_t^2 - \theta_n \to N(0, 2\sigma_t^4p(x)^{-1}e_1^T(H^{-1}S^{-1})^TS^*e_1H^{-1}S^{-1}/nh)$. □

Proof of Theorem 2.3 Based on the above results, we will decompose Q_n into two parts Q_n' and Q_n'', which satisfy that

(i) $nh\mathsf{E}[\theta^{-1}(x)Q_n']^2 \le (h/n)(h^{-1}a_n(1 + o(1)) + no(h^{-1})) \to 0$;

(ii) Q_n'' is identically distributed as Q_N and is asymptotically independent of $\widehat{\sigma}^2_{\text{MA},t}$.

Define

$$Q_N' = \frac{1}{n}\sum_{i=1}^{a_n}\{\widehat{Y}_i^2 - \mathsf{E}[\widehat{Y}_i^2|\widehat{X}_{t_i}]\}C_h(\widehat{X}_{t_i} - x) \tag{5.3}$$

and $Q''_n = Q_n - Q'_n$, where a_n is a positive integer with $a_n = o(n)$ and $a_n\Delta \to \infty$. Let $v_{n,l} = v_{l+1}\sqrt{h}$ and $v_i = \{\widehat{Y}_i - m_i\}C_h(\widehat{X}_{t_i} - x)$ $(i = 1, 2, \ldots, n)$, then by Fan and Yao (2003),

$$\mathsf{Var}\,[\theta^{-1}(x)v_{n,0}] = (1 + o(1)) \quad \text{and} \quad \sum_{l=1}^{n-1} |\mathsf{Cov}\,(v_{n,0}, v_{n,l+1})| = o(1),$$

which yields the result in (i). This combined with (5.1), (i) and (5.2) leads to

$$\theta^{-1}(x)\sqrt{nh}Q''_n \to N(0, 1).$$

According to the stationarity conditions of Banon (1978) and a stationary process X_t is said to satisfy the condition $G_2(s, \alpha)$ of Rosenblatt (1970) and the Proposition 2.6 of Fan and Yao (2003) imply that the $\rho(l)$ of $\{\widehat{X}_{t_i}\}$ decays exponentially and the strong-mixing coefficient $\alpha(l) \le \rho(l)$, it follows that

$$|\mathsf{E}\exp\{i\zeta(Q''_n + \widehat{\sigma}^2_{\mathrm{MA},t})\} - \mathsf{E}\exp\{i\zeta Q''_n\}\mathsf{E}\exp\{i\zeta\widehat{\sigma}^2_{\mathrm{MA},t}\}| \le 16\alpha(s_n) \to 0$$

for any $\zeta \in R$. By theorem of Volkonskii and Rozanov (1959), we get the asymptotic independence of $\widehat{\sigma}^2_{\mathrm{MA},t}$ and Q''_n.

By (i), $\sqrt{nh}Q'_n$ is asymptotically negligible. This together with Theorem 2.3 leads to

$$d_1\theta^{-1}(x)\sqrt{nh}Q_n + d_2V_2^{-1/2}\sqrt{Nb}[\widehat{\sigma}^2_{\mathrm{MA},t} - \sigma_t^2] \to N(0, d_1^2 + d_2^2)$$

for any $d_1, d_2 \in R$, where $V_2 = 2\sigma_t^4 \cdot \displaystyle\int_{-1}^{1} M^2(x)\mathrm{d}x$. Since Q_n is a linear transform of u,

$$V^{-1/2}\begin{pmatrix} \sqrt{nh}u \\ \sqrt{Nb}[\widehat{\sigma}^2_{\mathrm{MA},t} - \sigma_t^2] \end{pmatrix} \longrightarrow N(0, I_2),$$

which $V = \mathrm{blockdiag}\{V_1, V_2\}$ with $V_1 = 2\sigma_t^4 p(x)S^*$ where $S^* = (v_{i+j-2})$, $i, j = 1, 2$ with $v_j = \displaystyle\int u^j K^2(u)\mathrm{d}u$. This combined with equation (5.1) gives the joint asymptotic normality of b and $\widehat{\sigma}^2_{\mathrm{MA},t}$. Note that $B = o_p(1/\sqrt{nh})$, it follows that

$$\Sigma_2^{-1/2}\begin{pmatrix} \sqrt{nh}[\widehat{\sigma}^2_{\mathrm{LLE},s} - \sigma_t^2 - \theta_n] \\ \sqrt{Nb}[\widehat{\sigma}^2_{\mathrm{MA},t} - \sigma_t^2] \end{pmatrix} \longrightarrow N(0, I_2),$$

where $\Sigma_2^{-1/2} = \mathrm{diag}\{2\sigma_t^4 p(x)^{-1}e_1^T(H^{-1}S^{-1})^T S^* H^{-1}S^{-1}e_1, V_2\}$. Note that $\widehat{\sigma}^2_{\mathrm{MA},t}$ and $\widehat{\sigma}^2_{\mathrm{LLE},s}$ are asymptotically independent, it follows that the asymptotical normality of $\widehat{\sigma}^2_{I,(s,t)}$ holds. $\square$

References

[1] Vasicek, O., An equilibrium characterization of the term structure, *Journal of Financial Economics*, **5(2)**(1977), 177–188.

[2] Cox, J.C., Ingersoll, J.E. and Ross, S.A., A theory of the term structure of interest rates, *Econometrica*, **53(2)**(1985), 385–407.

[3] Chan, K.C., Karolyi, G.A., Longstaff, F.A. and Sanders, A.B., An empirical comparison of alternative models of the short-term interest rate, *The Journal of Finance*, **47(3)**(1992), 1209–1227.

[4] Florens-Zmirou, D., On estimating the diffusion coefficient from discrete observations, *Journal of Applied Probability*, **30(4)**(1993), 790–804.

[5] Jiang, G.J. and Knight, J.L., A nonparametric approach to the estimation of diffusion processes, with an application to a short-term interest rate model, *Econometric Theory*, **13(5)**(1997), 615–645.

[6] Stanton, R., A nonparametric model of term structure dynamics and the market price of interest rate risk, *The Journal of Finance*, **52(5)**(1997), 1973–2002.

[7] Arfi, M., Non-parametric variance estimation from ergodic samples, *Scandinavian Journal of Statistics*, **25(1)**(1998), 225–234.

[8] Fan, J.Q. and Yao, Q.W., Efficient estimation of conditional variance functions in stochastic regression, *Biometrika*, **85(3)**(1998), 645–660.

[9] Jacod, J., Non-parametric kernel estimation of the coefficient of a diffusion, *Scandinavian Journal of Statistics*, **27(1)**(2000), 83–96.

[10] Hoffmann, M., L^p estimation of the diffusion coefficient, *Bernoulli*, **5(3)**(1999), 447–481.

[11] Fan, J.Q. and Zhang, C.M., A reexamination of diffusion estimators with applications to financial model validation, *Journal of the American Statistical Association*, **98(461)**(2003), 118–134.

[12] Bandi, F.M. and Phillips, P.C.B., Fully nonparametric estimation of scalar diffusion models, *Econometrica*, **71(1)**(2003), 241–283.

[13] Nicolau, J., Bias reduction in nonparametric diffusion coefficient estimation, *Econometric Theory*, **19(5)**(2003), 754–777.

[14] Aït-Sahalia, Y. and Mykland, P.A., Estimators of diffusions with randomly spaced discrete observations: a general theory, *The Annals of Statistics*, **32(5)**(2004), 2186–2222.

[15] Jeffrey, A., Kristensen, D., Linton, O., Nguyen, T. and Phillips, P.C.B., Nonparametric estimation of a multifactor Heath-Jarrow-Morton model: an integrated approach, *Journal of Financial Econometrics*, **2(2)**(2004), 251–289.

[16] Renò, R., Nonparametric estimation of stochastic volatility models, *Economics Letters*, **90(3)**(2006), 390–395.

[17] Arapis, M. and Gao, J., Empirical comparison in short-term interest rate models using nonparametric methods, *Journal of Financial Econometrics*, **4(2)**(2006), 310–345.

[18] Bandi, F. and Moloche, G., On the functional estimation of multivariate diffusion processes, Work paper, University of Chicago and MIT, 2008.

[19] Brenner, R.J., Harjes, R.H. and Kroner, K.F., Another look at models of the short-term interest rate, *Journal of Financial and Quantitative Analysis*, **31(1)**(1996), 85–107.

[20] Fan, J.Q., Fan, Y.Y. and Jiang, J.C., Dynamic integration of time- and state-domain methods for volatility estimation, *Journal of the American Statistical Association*, **102(478)**(2007), 618–631.

[21] Renò, R., Nonparametric estimation of the diffusion coefficient of stochastic volatility models, *Econometric Theory*, **24(5)**(2008), 1174–1206.

[22] Barndorff-Nielsen, O.E. and Shephard, N., Econometric analysis of realized volatility and its use in estimating stochastic volatility models, *Journal of the Royal Statistical Society: Series B*, **64(2)**(2002a), 253–280.

[23] Barndorff-Nielsen, O.E. and Shephard, N., Estimating quadratic variation using realized variance, *Journal of Applied Econometrics*, **17(5)**(2002b), 457–478.

[24] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., The distribution of realized exchange rate volatility, *Journal of the American Statistical Association*, **96(453)**(2001a), 42–55.

[25] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Ebens, H., The distribution of realized stock return volatility, *Journal of Financial Economics*, **61(1)**(2001b), 43–76.

[26] Andersen, T.G., Bollerslev, T., Diebold, F.X. and Labys, P., Modeling and forecasting realized volatility, *Econometrica*, **71(2)**(2003), 579–625.

[27] Barndorff-Nielsen, O.E. and Shephard, N., Realized power variation and stochastic volatility models, *Bernoulli*, **9(2)**(2003), 243–265.

[28] Barndorff-Nielsen, O.E. and Shephard, N., Power and bipower variation with stochastic volatility and jumps, *Journal of Financial Econometrics*, **2(1)**(2004a), 1–37.

[29] Barndorff-Nielsen, O.E. and Shephard, N., Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics, *Econometrica*, **72(3)**(2004b), 885–925.

[30] Barndorff-Nielsen, O.E. and Stelzer, R., Absolute moments of generalized hyperbolic distributions and approximate scaling of normal inverse Gaussian Lévy processes, *Scandinavian Journal of Statistics*, **32(4)**(2005), 617–637.

[31] Barndorff-Nielsen, O.E. and Shephard, N., Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics*, **4(1)**(2006), 1–30.

[32] Andersen, T.G., Bollerslev, T. and Meddahi, N., Analytical evaluation of volatility forecasts, *International Economic Review*, **45(4)**(2004), 1079–1110.

[33] Andersen, T.G., Bollerslev, T. and Diebold, F.X., Roughing it up: including jump components in the measurement, modeling and forecasting of return volatility, *The Review of Economics and Statistics*, **89(4)**(2007), 701–720.

[34] Robinson, P.M., Large-Sample inference for nonparametric regression with dependent errors, *The Annals of Statistics*, **25(5)**(1997), 2054–2083.

[35] Härdle, W., Herwartz, H. and Spokoiny, V., Time inhomogeneous multiple volatility modeling, *Journal of Financial Econometrics*, **1(1)**(2003), 55–95.

[36] Fan, J.Q., Jiang, J.C., Zhang, C.M. and Zhou, Z.W., Time-dependent diffusion models for term structure dynamics, *Statistica Sinica*, **13(4)**(2003), 965–992.

[37] Mercurio, D. and Spokoiny, V., Statistical inference for time-inhomogeneous volatility models, *The Annals of Statistics*, **32(2)**(2004), 577–602.

[38] Foster, D.P. and Nelson, D.B., Continuous record asymptotics for rolling sample variance estimators, *Econometrica*, **64(1)**(1996), 139–174.

[39] Aït-Sahalia, Y., Mykland, P.A. and Zhang, L., How often to sample a continuous-time process in the presence of market microstructure noise, *The Review of Financial Studies*, **18(2)**(2005), 351–416.

[40] Fan, J.Q. and Wang, Y.Z., Spot volatility estimation for high-frequency data, *Statistics and Its Interface*, **1(2)**(2008), 279–288.

[41] Renò, R., Roma, A. and Schaefer, S., A comparison of alternative non-parametric estimators of the short rate diffusion coefficient, *Economic Notes*, **35(3)**(2006), 227–252.

[42] Banon, G., Nonparametric identification for diffusion processes, *SIAM Journal on Control and Optimization*, **16(3)**(1978), 380–395.

[43] Rosenblatt, M., Density estimates and Markov sequences, in *Nonparametric Techniques in Statistical Inference* (Editor: Puri, M.L.), Cambridge, U.K.: Cambridge University Press, 1970, 199–213.

[44] Volkonskii, V.A. and Rozanov, Yu.A., Some limit theorems for random functions. I., *Theory of Probability and Its Applications*, **4(2)**(1959), 178–197.

[45] Scott, D.W., *Multivariate Density Estimation: Theory, Practice, and Visualization*, New York: John Wiley & sons, 1992.

[46] Fan, J.Q. and Yao, Q.W., *Nonlinear Time Series: Nonparametric and Parametric Methods*, New York: Springer-Verlag, 2003.

扩散模型中扩散系数的非参数组合估计

叶绪国[1,2]　凌能祥[3]　姚仁海[2]

([1]东南大学数学系, 南京, 210096; [2]凯里学院数学科学学院, 凯里, 556011)

([3]合肥工业大学数学学院, 合肥, 23009)

为了提高扩散系数估计的准确度, 我们利用动态组合时间域与状态域信息提出一个新的组合估计量. 我们发现所提组合估计量能有效估计扩散模型的扩散系数, 正如在本文中模拟所示. 在一定的条件下, 建立了估计量的渐进正态性, 并证明了时间域估计量与状态域估计量是渐进独立的. 大量的模拟展示了所提组合估计量优于单域估计量, 也优于本文所提估计量.

关 键 词: 非参数估计, 扩散系数, 组合估计, 扩散模型.

学 科 分 类 号: O212.1.