Statistica Sinica **26** (2016), 1305-1329
doi:http://dx.doi.org/10.5705/ss.202015.0215

GOODNESS-OF-FIT TEST FOR THE SVM BASED ON NOISY OBSERVATIONS

Liang-Ching Lin[1,3], Sangyeol Lee[2] and Meihui Guo[3]

[1]*National Cheng Kung University,* [2]*Seoul National University and* [3]*National Sun Yat-sen University*

Abstract: In financial high frequency data analysis, the efficient price of an asset is commonly assumed to follow a continuous-time stochastic volatility model, contaminated with a microstructure noise. In this study, we consider a goodness-of-fit test problem for the efficient price models based on discretely observed samples and employ a goodness-of-fit test based on the empirical characteristic function. We show that the proposed test is asymptotically a weighted sum of products of centered normal random variables. To evaluate the proposed test, we conducted a simulation study using a bootstrap method. A data analysis is provided for illustration.

Key words and phrases: Empirical characteristic function, goodness of fit, high frequency data, microstructure noise, stochastic volatility model.

1. Introduction

High-frequency financial time series provides a rich source of problems as to trading processes and market microstructure. In particular, owing to special characteristics that occur in this field, such as microstructure noise effects, the analysis of high-frequency data has brought a new challenge to economists and statisticians, see Tsay (2010, Chap. 5). It is conventionally assumed that, instead of observing the efficient log-price p at transaction time t_i, we observe p with noise:

$$\widetilde{p}_{t_i} = p_{t_i} + \eta_{t_i},$$

where $\{\eta_{t_i}\}$ are i.i.d. noises with mean zero and variance σ_η^2 and are independent of the process p. The noise term η represents a microstructure contamination owing to imperfections of trading processes. See, for instance, Aït-Sahalia, Mykland, and Zhang (2005) and Bandi and Russell (2006). This microstructure noise results from such information or non-information related factors as bid-ask spread, differences in trade sizes, informational asymmetries of traders, inventory control effects, and discreteness of price changes. It is well known that the microstructure noise dominates the signal in high-frequency data and creates problems in the model-free estimation of integrated volatility of high-frequency data. For example, the conventional realized volatility estimator diverges to infinity when the sampling frequency approaches zero: see Barndorff-Nielsen and Shephard (2002), Aït-Sahalia, Mykland, and Zhang (2005), Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2006), Bandi and Russell (2006), Fan and Wang (2007), Bandi and Russell (2008), Barndorff-Nielsen et al. (2008), Barndorff-Nielsen et al. (2009), and Reiss (2011).

In this study, we assume that the efficient log price process satisfies the stochastic volatility model (SVM)

$$\begin{cases} dp_t = \sigma_t dW_t, & p_0 = x_0, \\ d\sigma_t^2 = b(\sigma_t^2)dt + \sqrt{v(\sigma_t^2)}dB_t, & \sigma_0^2 = \zeta, \end{cases} \tag{1.1}$$

where $(B_t, W_t)_{t>0}$ is a two-dimensional standard Brownian motion, σ_t^2 is the instantaneous volatility at time t, and ζ is a positive random variable independent of (B_t, W_t). Empirical evidence suggests that the SVM approach provides a better modeling for high-frequency transaction data than the classical Black-Scholes constant volatility method. One may also consider SDE models with price jumps but, in such cases, the jump component can be smoothed by a wavelet method as in Fan and Wang (2007). Thus we focus on Model (1.1) with no price jumps.

Modeling of the SVM (1.1) emphasizes the specification of the diffusion coefficient v of the volatility process $\{\sigma_t^2\}$ that plays an important role in derivative pricing, portfolio allocation, and risk management. Since the diffusion coefficient v is uniquely determined by both the marginal distribution and autocorrelation function of σ_t^2, see Aït-Sahalia (1996a,b), Bibby, Skovgaard, and Sørensen (2005), and Chen, Gao, and Tang (2008), it can be well specified through a goodness-of-fit test for the stationary distribution of σ_t^2, see the hypothesis testing problem in (2.1). Motivated by this, Lin, Lee, and Guo (2013, 2014) studied a goodness-of-fit test for $\{\sigma_t^2\}$ of SVM (1.1) based on discretely sampled efficient log-price $\{p_t\}$, assuming no presence of microstructure noises. We aim to extend the method of Lin, Lee, and Guo (2013, 2014) to the observed price $\widetilde{p}$ of SVM (1.1) with microstructure noises η. Specifically, we use the goodness-of-fit test based on measuring differences between the empirical characteristic function (e.c.f.) and true parametric characteristic function (c.f.) divided by the characteristic function of the microstructure noise obtained from the hypothesized stochastic volatility model. This issue is more challenging than that of our previous study, since the volatility process is latent and the price process is contaminated with noise.

The organization of the paper is as follows. In Section 2, the goodness-of-fit test is introduced and its limiting null distribution is derived as a weighted sum of products of centered normal random variables. In Section 3, we study the moment estimators of the volatility model parameters and use two SVMs for illustration. In Section 4, we study noise parameter estimation and discuss the performances of model and noise parameter estimation. In Section 5, simulation and empirical studies are reported. Concluding remarks are provided in Section 6. Proofs are given in the Appendix.

2. Main Result

We need regularity conditions for b and v.

(A1) The functions $b(x)$ and $v(x)$ defined on $(0,\infty)$ satisfy $b(x)\in C^1$, $v(x)\in C^2$ for all $x>0$, and there exists $K>0$ such that, for all $x>0$, $|b(x)|\le K(1+|x|)$ and $v(x)\le K(1+x^2)$.

(A2) The scale and speed densities of the diffusion σ_t^2,

$$s(x)=\exp\left(-2\int_{x_0}^{x}\frac{b(u)}{v(u)}du\right)\quad\text{and}\quad m(x)=\frac{1}{v(x)s(x)},\ x>0,$$

satisfy

$$\int_{0^+}s(x)dx=+\infty,\ \int^{+\infty}s(x)dx=+\infty,\ \int_0^{+\infty}m(x)dx=M<+\infty,$$

where $\int_{0^+}$ denotes the integral over the interval $(0,c)$ for some $c>0$ and $\int^{+\infty}$ denotes the integral over the interval (c',∞) for some $c'>0$.

We impose conditions on the stationary density of σ_t^2,

$$f_{\sigma,\theta}(x)=\frac{1}{M}m(x)\mathbf{1}_{[x>0]},$$

where θ denotes the true parameter.

(A3) The initial random variable $\sigma_0^2=\zeta$ has the density function $f_{\sigma,\theta}$ and

$$\int_0^{\infty}|v|^{\nu}f_{\sigma,\theta}(v)dv<\infty\quad\text{for some }\nu\ge 2.$$

(A4) For all $q\ge 1$, there exist constants $C_q>0$ such that

$$E_\theta|\sigma_s-\sigma_t|^{2q}\le C_q|t-s|^q.$$

(A1) and (A2) ensure that the unique solution of σ_t^2 is positive and recurrent on $(0,\infty)$, whereas (A3) entails that it is strictly stationary, ergodic, and time-reversible. (A4) can be found in Prakasa Rao (1999) and Kessler (2000). With regard to the limit theorems of empirical processes and parameter estimation for Model (1.1), we refer to Genon-Catalot, Jeantheau, and Larédo (1998, 1999, 2000).

We assume $\widetilde{p}_t$ is observed at equispaced time points $(t_1, t_2, \cdots, t_n)$, where $t_i = ik_n$ with $k_n \to 0,\ nk_n \to \infty$, and $nk_n^2 \to 0$ as $n \to \infty$. In this case, we write the observed log return at time t_i as

$$\widetilde{r}_i = \widetilde{p}_{t_i} - \widetilde{p}_{t_{i-1}} = r_i + \varepsilon_i,$$

where $r_i = p_{t_i} - p_{t_{i-1}}$ denotes the nominal return, and $\varepsilon_i = \eta_{t_i} - \eta_{t_{i-1}}$. Since the η_{t_i}'s are i.i.d. random variables with variance σ_η^2, the noise process $\{\varepsilon_i\}$ is an MA(1) process with $Var(\epsilon_i) = \sigma_\varepsilon^2 = 2\sigma_\eta^2$. The distribution of ε_i can be obtained from the marginal distribution of η using a convolution method. We further assume that η_t has a stationary density $f_{\eta,\beta}$ wherein β denotes the true vector parameter.

Let $\{f_{\sigma,\theta} : \theta \in \Theta \subset R^d\}$ and $\{f_{\eta,\beta} : \beta \in B \subset R^{d_1}\}$ be families of density functions and suppose that one wishes to test the hypotheses

$$\mathcal{H}_0 : \sigma_t^2 \sim f_{\sigma,\theta} \text{ and } \ \eta_t \sim f_{\eta,\beta} \ \text{ for some } \theta \in \Theta,\ \beta \in B \text{ vs. } \mathcal{H}_1 : \text{ not } \mathcal{H}_0. \quad (2.1)$$

Set $\xi_i = \sigma_{t_{i-1}}(W_{t_i} - W_{t_{i-1}})/k_n^{1/2}$ and

$$\hat{\xi}_i = \frac{(p_{t_i} - p_{t_{i-1}})}{k_n^{1/2}} = \int_{t_{i-1}}^{t_i} \sigma_s d\frac{W_s}{k_n^{1/2}} = \xi_i + \Delta_{ni} \quad (2.2)$$

with $\Delta_{ni} = \int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})dW_s/k_n^{1/2}$. It can be seen that under $\mathcal{H}_0$, due to (A4), $E_\theta|\Delta_{ni}|^{2q} = O(k_n^q)$ for any $q = 1, 2, \ldots$(cf., Lee (2010)).

For estimation of θ and β, we need the following conditions.

(A5) Let $\psi : \Theta \times R \to R^d$ be a vector-valued function and take $U_n(\theta) = \sum_{j=1}^n \psi(\theta, \hat{\xi}_j)$ with $\psi(\theta, x) = (\psi_1(\theta, x), \ldots, \psi_d(\theta, x))'$. Let $\hat{\theta}$ be the solution of $U_n(\theta) = 0$ based on the full sample $\{\hat{\xi}_j\}_{j=1}^n$ with decreasing sampling intervals. Then, under $\mathcal{H}_0$, for $i = 1, \ldots, d$,

$$\sup_t \Big\{ \Big| \frac{1}{n} \sum_{j=1}^n \Big(\frac{\partial \psi_i}{\partial \theta_i}(t, \hat{\xi}_j) - \frac{\partial \psi_i}{\partial \theta_i}(\theta, \hat{\xi}_j) \Big) \Big| : |t - \theta_i| \le a_n \Big\} \xrightarrow[n\to\infty]{P} 0$$

whenever $a_n \to 0$, and $\hat{\theta} \xrightarrow{P} \theta$ as $n \to \infty$.

For the vector of real-valued functions g, it holds that

$$\hat{\theta} = \theta + h(\theta, \tilde{\mathbf{r}}) + o_p\left({n_k}^{-1/2}\right), \quad (2.3)$$

$$h(\theta,\tilde{\mathbf{r}}) = \frac{1}{n}\sum_{j=1}^{n} g(\theta,\hat{\xi}_j) + o_p\left({n_k}^{-1/2}\right), \tag{2.4}$$

where $\tilde{\mathbf{r}} = (\tilde{r}_1,\ldots,\tilde{r}_n)'$ and $n_k = nk_n$.

The function $g(\theta,x) = (g_1(\theta,x),\ldots,g_d(\theta,x))'$ satisfies $E_\theta g_r(\theta,\xi_j) = 0$, $E_\theta|g_r(\theta,\xi_j)|^{2\nu} < \infty$ for some $\nu > (2-\kappa)/(1-\kappa)$, $0<\kappa<1$, and

$$|g_r(\theta,x_1) - g_r(\theta,x_2)| \le w_r(x_1,x_2,\theta)|x_1-x_2|$$

for some real-valued continuous functions $w_r \ge 0$ satisfying

$$\sup_{j,k\in\mathbb{N}} E_\theta \sup_{a\in[-A,A]}\left[\left|w_r(\xi_j+a,\xi_k,\theta)\right|^{2\nu} + \left|w_r(\xi_1+a,\tilde{\xi}_1,\theta)\right|^{2\nu}\right] < \infty$$

for any independent copy $\tilde{\xi}_1$ of ξ_1 and for some $A>0$, we write $\mathcal{W} = (w_1,\ldots,w_d)$.

(A6) $\hat{\beta} = \beta + O_p(n^{-1/2})$.

Remark 1. We illustrate that the method of moment estimator satisfies (2.3) and (2.4). For example, consider the Heston model defined in (3.2), and let $S_L(\mathbf{a}^*)$ and $\hat{m}_4$ denote the estimators of integrated volatility and quarticity, respectively, as described in Section 3. Set

$$U(\theta) = (U_1(\theta),U_2(\theta))^T = \left(\frac{\alpha}{\lambda}, 3(\frac{\alpha}{\lambda^2}+\frac{\alpha^2}{\lambda^2})\right)^T, U(\hat{\theta}) = \left(\frac{\hat{\alpha}}{\hat{\lambda}}, 3(\frac{\hat{\alpha}}{\hat{\lambda}^2}+\frac{\hat{\alpha}^2}{\hat{\lambda}^2})\right)^T,$$

$$\psi(\theta,x) = (\psi_1(\theta,x)\psi_2(\theta,x))^T = \left(x^2-\frac{\alpha}{\lambda}, x^4 - 3(\frac{\alpha}{\lambda^2}+\frac{\alpha^2}{\lambda^2})\right)^T, \tag{2.5}$$

$$y(\theta,\tilde{r}_1,\ldots,\tilde{r}_n) = \left(S_L(\mathbf{a}^*) - \frac{\alpha}{\lambda}, \quad 3\hat{m}_4 - 3(\frac{\alpha}{\lambda^2}+\frac{\alpha^2}{\lambda^2})\right)^T.$$

Then, by (2.5), we have

$$U(\hat{\theta}) = U(\theta) + y(\theta,\tilde{\mathbf{r}}) + o_p(n_k^{-1/2}). \tag{2.6}$$

We can see that

$$\frac{1}{n}\sum_{j=1}^{n} E\psi(\theta,\hat{\xi}_j) = \begin{pmatrix} \frac{1}{nk_n}E\int_0^{nk_n}\sigma_s^2 ds - \frac{\alpha}{\lambda} \\ \frac{3}{nk_n}E\int_0^{nk_n}\sigma_s^4 ds - 3(\frac{\alpha}{\lambda^2}+\frac{\alpha^2}{\lambda^2}) \end{pmatrix},$$

and thus

$$\frac{1}{n}\sum_{j=1}^{n}\psi(\theta,\hat{\xi}_j) - \frac{1}{n}\sum_{j=1}^{n}E\psi(\theta,\hat{\xi}_j) = O_p(\sqrt{k_n}), \tag{2.7}$$

$$y(\theta, \tilde{\mathbf{r}}) - \frac{1}{n}\sum_{j=1}^{n} E\psi(\theta, \hat{\xi}_j) = O_p(n^{-1/4}). \tag{2.8}$$

For the orders of (2.7) and (2.8), see, for example, Andersen et al. (2001) and Lin and Guo (2015).

By the Mean Value Theorem, we have

$$U(\hat{\theta}) - U(\theta) = A(\theta^*)(\hat{\theta} - \theta), \tag{2.9}$$

where $A(\theta^*) = [[a_{ik}]]$, $a_{ik} = \frac{\partial U_i}{\partial \theta_k}(\theta^*)$, and θ^* lies on the line segment determined by $\hat{\theta}$ and θ. If A is invertible, by plugging (2.9) to (2.6), we have $\hat{\theta} = \theta + h(\theta, \tilde{\mathbf{r}}) + o_p(n_k^{-1/2})$, where $h(\theta, \tilde{\mathbf{r}}) = A^{-1}y(\theta, \tilde{\mathbf{r}})$ and (2.3) holds. Then, multiplying (2.7) and (2.8) with A^{-1}, we get

$$\frac{1}{n}\sum_{j=1}^{n} g(\theta, \hat{\xi}_j) - \frac{1}{n}\sum_{j=1}^{n} Eg(\theta, \hat{\xi}_j) = O_p(\sqrt{k_n}), \tag{2.10}$$

$$h(\theta, \tilde{\mathbf{r}}) - \frac{1}{n}\sum_{j=1}^{n} Eg(\theta, \hat{\xi}_j) = O_p(n^{-1/4}), \tag{2.11}$$

where $g(\theta, \hat{\xi}_j) = A^{-1}\psi(\theta, \hat{\xi}_j)$. Finally, combining (2.10) and (2.11), we have

$$h(\theta, \tilde{\mathbf{r}}) = \frac{1}{n}\sum_{j=1}^{n} g(\theta, \hat{\xi}_j) + O_p(\sqrt{k_n}) + O_p\left(n^{-1/4}\right) = \frac{1}{n}\sum_{j=1}^{n} g(\theta, \hat{\xi}_j) + o_p(n_k^{-1/2}),$$

where we have used the fact that $O_p\left(n^{-1/4}\right) = o_p\left({n_k}^{-1/2}\right)$ and $O_p\left(\sqrt{k_n}\right) = o_p\left({n_k}^{-1/2}\right)$, since $k_n \to 0$, $nk_n \to \infty$, and $nk_n^2 \to 0$ as $n \to \infty$. Hence, (2.4) holds.

Let

$$\hat{\phi}_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{it\tilde{r}_j/k_n^{1/2}}$$

be the empirical characteristic function (e.c.f.) based on the observed log returns, $\hat{\phi}_\xi(t) = E_\theta(e^{it\xi_1})\big|_{\theta=\hat{\theta}}$ be the characteristic function (c.f.) of ξ_1 with θ replaced by its estimator $\hat{\theta}$, and $\hat{\phi}_\eta(t) = E_\beta\left(e^{it\eta_1}\right)\big|_{\beta=\hat{\beta}}$ be the c.f. of η_1 with β replaced by its estimator $\hat{\beta}$. We need the following conditions.

(A7) $\phi_\xi(t) = E_\theta(e^{it\xi_1})$ is continuously differentiable with respect to θ and $\nabla\phi_\xi(t) = (\partial\phi_\xi(t)/\partial\theta_1, \ldots, \partial\phi_\xi(t)/\partial\theta_d)'$ satisfies

$$\int \left\| \frac{\partial\phi_\xi(t)}{\partial\theta_i} \right\|^{2\nu} dG(t) < \infty.$$

(A8) (i) The characteristic function of η satisfies

$$|\phi_\eta(t)| = e^{-\alpha_0(\beta)|t|^{\alpha_1}+R(\beta,t)} \quad \text{as } t \to \infty,$$

where $\|\nabla_\beta \alpha_0(\beta)\|_1 = O(1)$, $0 < \alpha_1 < 1$, and $R(\beta,t) = o(|t|^{\alpha_1})$ with $\|\nabla_\beta R(\beta,t)\|_1 = o(|t|^{\alpha_1})$.

(ii) The characteristic function of η satisfies

$$|\phi_\eta(t)| = \alpha_0(\beta)|t|^{-\alpha_1(\beta)} + R(\beta,t) \quad \text{as } t \to \infty,$$

where $\|\nabla_\beta \alpha_0(\beta)\|_1 = O(1)$, $0 < \alpha_1(\beta)$, $\|\nabla_\beta \alpha_1(\beta)\|_1 = O(1)$, and $R(\beta,t) = o(|t|^{-\alpha_1(\beta)})$ with $\|\nabla_\beta R(\beta,t)\|_1/R(\beta,t) = O(|t|^{-\alpha_3})$ for some $\alpha_3 > 0$.

Remark 2. Condition (A8)(i) is related to the supersmoothness case of Fan (1991) that includes the t and generalized error distributions. Condition (A8)(ii) is related to the ordinary smoothness case of Fan (1991) that includes the exponential and gamma distributions.

Consider the characteristic function based test statistic:

$$\hat{T}_n = n_k \int \left| \frac{\hat{\phi}_n(t) - \hat{\phi}_\xi(t)\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} \right|^2 dG(t). \tag{2.12}$$

Since the asymptotic distribution of $\hat{T}_n$ is hard to derive directly, we introduce

$$\hat{T}_n^* = n_k \int \Big| \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j} - \hat{\phi}_\xi(t) \Big|^2 dG(t),$$

the characteristic function-based test statistic for the noiseless case; its limiting null distribution can be seen in Lin, Lee, and Guo (2013). Similarly to Section 3.1 there, we can get

$$\Big| \hat{T}_n^* - n_k \int \Big| \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j} - E_\theta e^{it\hat{\xi}_1} - (\nabla\phi_\xi(t))' \frac{1}{n}\sum_{j=1}^n \dot{g}(\theta,\hat{\xi}_j) \Big|^2 dG(t) \Big| = o_p(1),$$

where $\dot{g}(\theta,\hat{\xi}_j) = g(\theta,\hat{\xi}_j) - E_\theta g(\theta,\hat{\xi}_1)$ for $j = 1,2,\ldots,n$, and the expectation is under the stationary law of $\hat{\xi}_1$. Subsequently, $\hat{T}_n^*$ should have the same limiting null distribution as

$$n_k \int \Big| \hat{\phi}_n(t) - E_\theta e^{it\hat{\xi}_1} - (\nabla\phi_\xi(t))' \frac{1}{n}\sum_{j=1}^n \dot{g}(\theta,\hat{\xi}_j) \Big|^2 dG(t),$$

which is also named as $\hat{T}_n^*$ without any confusion. A result ensures that $\hat{T}_n^*$ can be approximated by a degree-2 degenerate V-statistic (cf., Lemma 3.1 of Lin, Lee, and Guo (2013)).

Lemma 1. *If* (A5) *holds, $\hat{T}_n^*$ is a degree-2 degenerate V-statistic,*

$$\hat{T}_n^* = \frac{n_k}{n^2}\sum_{j=1}^{n}\sum_{k=1}^{n} k(\hat{\xi}_j, \hat{\xi}_k; \theta), \tag{2.13}$$

where $E_\theta k(x, \hat{\xi}_1) = E_\theta k(\hat{\xi}_1, x) = 0$ for any $x \in \mathbb{R}$ and

$$k(x, y; \theta) = \mathrm{Re}\,\{I_1(x, y) + I_2(y) + I_3(x, y)\} \tag{2.14}$$

with

$$I_1(x, y) = \int \Big[e^{-itx}\left(e^{ity} - E_\theta e^{it\hat{\xi}_1}\right) - e^{-itx}(\nabla\phi_\xi(t))'\dot{g}(\theta, y)\Big]\, dG(t),$$

$$I_2(y) = \int \Big[E_\theta e^{-it\hat{\xi}_1}\left(-e^{ity} + E_\theta e^{it\hat{\xi}_1}\right) + E_\theta e^{-it\hat{\xi}_1}(\nabla\phi_\xi(t))'\dot{g}(\theta, y)\Big]\, dG(t),$$

$$I_3(x, y) = \int \Big[(\nabla\phi_\xi(-t))'\dot{g}(\theta, x)\left(-e^{ity} + E_\theta e^{it\hat{\xi}_1}\right) + (\nabla\phi_\xi(-t))'g(\theta, x)(\nabla\phi_\xi(t))'\dot{g}(\theta, y)\Big]\, dG(t).$$

We decompose (2.13) using wavelet functions. By (A5) and (2.14),

$$\int\int k(x, y)^2 d\tilde{F}(x; \theta) d\tilde{F}(y; \theta) < \infty, \tag{2.15}$$

where $\tilde{F}(x; \theta)$ denotes the stationary distribution of $\hat{\xi}_1$. Let Φ be a Lipschitz-continuous scale function and Ψ be a Lipschitz-continuous wavelet mother function with a compact support such that $\int_{-\infty}^{\infty}\Phi(x)dx = 1$ and $\int_{-\infty}^{\infty}\Psi(x)dx = 0$. Define the sequence of wavelet functions

$$\Phi_{j,l}(x) = 2^{j/2}\Phi(2^j x - l),\ \Psi_{j,l}(x) = 2^{j/2}\Psi(2^j x - l),\ j \in \mathbb{N} \cup \{0\},\ l \in \mathbb{Z};$$

this is an orthonormal basis of L_2-space satisfying

$$\int \Phi_{j,l}(x)\Phi_{j',l'}(x)dx = \begin{cases} 1, & \text{if } j = j' \ \&\ l = l', \\ 0, & \text{otherwise.} \end{cases}$$

Owing to (2.15), the kernel function k has a decomposition (cf., Daubechies (2002)) in the L_2-sense,

$$k(x, y) = \sum_{j=0}^{\infty}\sum_{k_1, k_2=-\infty}^{\infty} \lambda_{j;k_1,k_2}\varphi_{j,k_1}(x)\varphi_{j,k_2}(y),$$

where

$$\varphi_{j,k_1} = \begin{cases} \Phi_{j,k_1},\ j = 0, \\ \Psi_{j,k_1},\ j \in \mathbb{N}, \end{cases}$$

$$\lambda_{j;k_1,k_2} = \int\int \dot{h}^{(c)}(x, y)\varphi_{j,k_1}(x)\varphi_{j,k_2}(y)dxdy. \tag{2.16}$$

We have two conditions on the scale density of (σ_t^2) in (A1), (B1) $\int_{0^+} s(v)\left[\int_0^v f_{\sigma,\theta}(u)du\right]^2 dv < \infty;$

(B2) $\int^{\infty} s(v)\left[\int_v^{\infty} u^{\nu/2} f_{\sigma,\theta}(u)du\right]^2 dv < \infty$

to obtain the following (cf., Theorem 3.1 of Lin, Lee, and Guo (2013)).

Theorem 1. *Let* $\{p_t\}$ *and* $\{\sigma_t^2\}$ *be as at* (1.1). *Suppose that* (A1)$\sim$(A5), (A7), (B1), *and* (B2) *hold and the distribution function* G *satisfies*

$$\lim_{t\to\infty} \frac{dG(t)}{d\tilde{F}(t;\theta)} = 0.$$

If the α*-mixing coefficients of* $\{\sigma_t\}$ *satisfy* $\alpha_\sigma(m) = O(e^{-am})$ *for some* $a > 0$, *then, under* $\mathcal{H}_0$ *as* $n \to \infty$,

$$\hat{T}_n^* \xrightarrow{d} Z \equiv \sum_{j=0}^{\infty} \sum_{k_1,k_2=-\infty}^{\infty} \lambda_{j;k_1,k_2} Z_{j,k_1} Z_{j,k_2}, \tag{2.17}$$

where $Z_{j,k}$, $j = 0, 1, 2, \ldots$, $k = 0, \pm 1, \pm 2, \ldots$, *are correlated centered normally distributed random variables and* $\lambda_{j;k_1,k_2}$ *are the wavelet coefficients of the kernel function* $\dot{h}^{(c)}$ *in* (2.14) (*cf.,* (2.16)).

Theorem 2. *Under the assumptions of Theorem* 1, (A6), *and* (A8)(i), *under* $\mathcal{H}_0$,

(i) $\hat{T}_n - \hat{T}_n^* = o_p(1)$;

(ii) $\hat{T}_n$ *has the limiting distribution as at* (2.17).

Theorem 3. *Under the assumptions of Theorem* 1, (A6), *and* (A8)(ii), *under* $\mathcal{H}_0$,

(i) $\hat{T}_n - \hat{T}_n^* = o_p(1)$;

(ii) $\hat{T}_n$ *has the limiting distribution as at* (2.17).

3. Volatility Parameter Estimation

In this section, we consider two SVM examples as illustrations. The characteristic functions and parameter estimations based on the method of moment estimates are provided. The integrated second and fourth moments of the efficient returns $\{r_i\}$ are

$$m_2 = E\left(\int_0^{nk_n} \sigma_s^2 ds\right), \quad m_4 = E\left(\int_0^{nk_n} \sigma_s^4 ds\right), \tag{3.1}$$

and their corresponding estimators are denoted by $\hat{m}_2$ and $\hat{m}_4$.

Example 1 (Heston model)**.** The process $\{p_t\}$ of the Heston (1993) model satisfies

$$\begin{cases} dp_t = \sigma_t dW_t, \\ d\sigma_t^2 = -\rho(\sigma_t^2 - \mu)dt + \omega\sqrt{\sigma_t^2} dB_t, \end{cases} \tag{3.2}$$

where $\{W_t : t \geq 0\}$ and $\{B_t : \ t \geq 0\}$ are independent Wiener processes. The volatility σ_t^2 has a stationary $Gamma(\alpha, \lambda)$ distribution with $\alpha = 2\rho\mu/\omega^2, \lambda = 2\rho/\omega^2$: see, for example, Bibby, Skovgaard, and Sørensen (2005). In view of Proposition 4.1 of Lin, Lee, and Guo (2013), it can be seen that the characteristic function of $\{\xi_i\}$ is

$$\phi_H = \left(\frac{2\lambda}{2\lambda + t^2}\right)^{\alpha}.$$

The two moment equations for the parameters α and λ are given by

$$m_2 = E\left(\int_0^{nk_n} \sigma_s^2 ds\right) = nk_n \frac{\alpha}{\lambda},$$
$$m_4 = E\left(\int_0^{nk_n} \sigma_s^4 ds\right) = nk_n \left(\frac{\alpha}{\lambda^2} + \frac{\alpha^2}{\lambda^2}\right),$$

and one has

$$\hat{\alpha} = \frac{\hat{m}_2^2}{nk_n\hat{m}_4 - \hat{m}_2^2}, \qquad \hat{\lambda} = \frac{nk_n\hat{m}_2}{nk_n\hat{m}_4 - \hat{m}_2^2}.$$

Example 2 (Stein and Stein model)**.** The process $\{p_t\}$ of the Stein and Stein (1991) model satisfies

$$\begin{cases} dp_t = \sigma_t dW_t, \\ d\sigma_t = -\rho(\sigma_t - \mu)dt + \omega dB_t. \end{cases} \tag{3.3}$$

The volatility σ_t has a stationary $N(\mu, \tau^2)$ distribution with $\tau^2 = \omega^2/(2\rho)$: see, for example, Bibby, Skovgaard, and Sørensen (2005). Owing to Proposition 4.1 of Lin, Lee, and Guo (2013), it can be seen that the characteristic function of $\{\xi_i\}$ is

$$\phi_S = \sqrt{\frac{1}{1 + t^2\tau^2}} \exp\left\{-\frac{\mu^2 t^2}{2(1 + t^2\tau^2)}\right\}.$$

The moment equations are

$$\begin{cases} m_2 = nk_n\left(\mu^2 + \tau^2\right), \\ m_4 = nk_n\left(\mu^4 + 6\mu^2\tau^2 + 3\tau^4\right), \end{cases}$$

so the moment estimators of μ and τ are

$$\hat{\mu} = \left[\frac{3\hat{m}_2^2 - nk_n\hat{m}_4}{2(nk_n)^2}\right]^{1/4}, \qquad \hat{\tau}^2 = \frac{2\hat{m}_2 - \sqrt{6\hat{m}_2^2 - 2nk_n\hat{m}_4}}{2nk_n}.$$

Here the $\{r_i\}$ are unobservable, and the observed returns $\{\tilde{r}_i\}$ are contaminated by microstructure noise. Thus, to estimate $\hat{m}_2$ and $\hat{m}_4$, a method of filtering out the noise process is required. For estimating m_4, we adopt the estimator of Jacod et al. (2009, Remark 4) to obtain

$$\hat{m}_4 = \frac{1}{3c^2\psi_2^2}\sum_{i=0}^{n-l_n+1}(\bar{r}_i^n)^4 - \frac{k_n\psi_1}{c^4\psi_2^2}\sum_{i=0}^{n-2l_n+1}(\bar{r}_i^n)^2\sum_{j=i+l_n}^{i+2l_n-1}\tilde{r}_j^2 + \frac{k_n\psi_1^2}{4c^4\psi_2^2}\sum_{i=1}^{n-2}\tilde{r}_i^2\tilde{r}_{i+2}^2,$$

where $\psi_1 = 1,\ \psi_2 = 1/12,\ l_n = \lfloor ck_n^{-1/2}\rfloor,\ c = 3,$

$$\bar{r}_i^n = \sum_{j=1}^{l_n-1} g\left(\frac{j}{l_n}\right)\tilde{r}_{i+j},$$

and $g(x) = \min\{x, 1-x\}$.

There are various methods for estimating $\int_0^1 \sigma_s^2 ds$, the integrated volatility, in the fixed-span in-fill setting ($k_n \to 0$ and $nk_n \to$ constant). For example, the two-scaled estimator of Zhang, Mykland, and Aït-Sahalia (2005); the multi-scaled estimator of Zhang (2006); the kernel estimator of Barndorff-Nielsen et al. (2009); the pre-averaging estimator of Jacod et al. (2009); the optimal restricted quadratic estimator of Lin and Guo (2015). Here we employ the quadratic estimator of Lin and Guo (2015) because of its finite sample efficiency.

The quadratic estimator of Lin and Guo (2015) is

$$\begin{aligned} S_L(\boldsymbol{a}) &= a_0\sum_{j=1}^{n+1}\widetilde{r}_j^2 + a_1\sum_{j=1}^{n}\widetilde{r}_j\widetilde{r}_{j+1} + \ldots + a_\ell\sum_{j=1}^{n+1-\ell}\widetilde{r}_j\widetilde{r}_{j+\ell} \\ &= a_0L_0 + a_1L_1 + \ldots + a_\ell L_\ell, \end{aligned}$$

where $L_h = \sum\widetilde{r}_j\widetilde{r}_{j+h}$ denotes the lag h sample autocovariance. We set $a_0 = 1$ and $a_1 = 2$ to ensure the unbiasedness of S_L: see Lemma 1 of Lin and Guo (2015). The optimal weights, $\boldsymbol{a}^* = (a_2^*, \ldots, a_\ell^*)$, are chosen to minimize the finite sample variance and to satisfy the system of equations (for details, see (16) of Lin and Guo (2015)):

$$\begin{cases} \mu_2a_2 + \rho_3a_3 + \gamma_4a_4 = -\gamma_2 - 2\rho_2, \\ \rho_3a_2 + \mu_3a_3 + \rho_4a_4 + \gamma_5a_5 = -2\gamma_3, \\ \gamma_{h+2}a_h + \rho_{h+2}a_{h+1} + \mu_{h+2}a_{h+2} + \rho_{h+3}a_{h+3} + \gamma_{h+4}a_{h+4} = 0 \end{cases} \tag{3.4}$$

for $2 \leqslant h \leqslant \ell - 2$, where $\mu_h = E\left[Var_{\mathcal{G}}(L_h)\right]$, $\rho_h = E\left[Cov_{\mathcal{G}}(L_{h-1},\ L_h)\right]$, $\gamma_h = E\left[Cov_{\mathcal{G}}(L_{h-2},\ L_h)\right]$, and $\mathcal{G}$ is the σ-field generated by $\{\sigma_t,\ t \geq 0\}$.

Lemma 2. *Assume the log price $\tilde{p}_t$'s are observed at equispaced time points $\{t_1, \ldots, t_n\}$ where $t_i = ik_n$ where $k_n \to 0, nk_n \to \infty$, and $nk_n^2 \to 0$ as $n \to \infty$. Let $K = E(\eta_t^4)/(E\eta_t^2)^2$. Then,*

$$\mu_0 = 2k_n E\left(\int_0^{nk_n} \sigma_s^4 ds\right) + 4\sigma_\varepsilon^2 E\left(\int_0^{nk_n} \sigma_s^2 ds\right) + (nK-1)\sigma_\varepsilon^4 + o(nk_n^3),$$

$$\mu_1 = k_n E\left(\int_0^{nk_n} \sigma_s^4 ds\right) + 2\sigma_\varepsilon^2 E\left(\int_0^{nk_n} \sigma_s^2 ds\right) + 2A_1 + \frac{(K+4)n-6}{4}\sigma_\varepsilon^4 + o(nk_n^3),$$

$$\rho_1 = -2\sigma_\varepsilon^2 E\left(\int_0^{nk_n} \sigma_s^2 ds\right) - 2A_1 - \frac{(K+1)n-2}{2}\sigma_\varepsilon^4,$$

$$\mu_h = k_n E\left(\int_0^{nk_n} \sigma_s^4 ds\right) + 2\sigma_\varepsilon^2 E\left(\int_0^{nk_n} \sigma_s^2 ds\right) + A_h + B_h + \frac{3n-3h}{2}\sigma_\varepsilon^4 + o(nk_n^3),$$

$$\rho_h = -\sigma_\varepsilon^2 E\left(\int_0^{nk_n} \sigma_s^2 ds\right) - \frac{1}{2}A_h - \frac{1}{2}B_h - \frac{2n-2h+1}{2}\sigma_\varepsilon^4, \ 2 \leqslant h \leqslant \ell,$$

where $A_h = \sigma_\varepsilon^2 E\left(\int_{(n-h)k_n}^{nk_n} \sigma_s^2 ds\right)$ and $B_h = \sigma_\varepsilon^2 E\left(\int_0^{hk_n} \sigma_s^2 ds\right)$, $1 \leqslant h \leqslant \ell$.

The proof of Lemma 2 is given in the Appendix. The results for the γ_h's are those in Lemma 2 of Lin and Guo (2015).

By dividing both sides of (3.4) by $n\sigma_\varepsilon^4$ and ignoring the $O(\ell k_n/n)$ terms, we obtain the system of equations

$$\begin{cases} \dot{\mu}_2 a_2 + \dot{\rho}_3 a_3 + \dot{\gamma}_4 a_4 + \dot{\gamma}_2 + 2\dot{\rho}_2 = 0, \\ \dot{\rho}_3 a_2 + \dot{\mu}_3 a_3 + \dot{\rho}_4 a_4 + \dot{\gamma}_5 a_5 + 2\dot{\gamma}_3 = 0, \\ \dot{\gamma}_{h+2} a_h + \dot{\rho}_{h+2} a_{h+1} + \dot{\mu}_{h+2} a_{h+2} + \dot{\rho}_{h+3} a_{h+3} + \dot{\gamma}_{h+4} a_{h+4} = 0, \end{cases} \tag{3.5}$$

where $2 \leqslant h \leqslant \ell - 2$ and

$$\dot{\mu}_h = S_{nr}^2 + 2S_{nr} + \frac{3n-3h}{2n}, \quad \dot{\rho}_h = -S_{nr} - \frac{2n-2h+1}{2n},$$

$$\dot{\gamma}_2 = \frac{n-1}{2n}, \quad \dot{\gamma}_h = \frac{n-h+1}{4n}, \quad S_{nr} = \frac{E\int_0^{nk_n} \sigma_s^2 ds}{n\sigma_\varepsilon^2},$$

for $2 \leq h \leq \ell$. We then use (3.5) to solve the optimal weights $\mathbf{a}^*$. The numerator of S_{nr} is slightly modified from that of Lin and Guo (2015) for our setting, and the system of equations (3.5) depends only on S_{nr}. Thus, we can use the recursive algorithm proposed by Lin and Guo (2015), Section 3, that employs the Newton-Raphson and Gauss-Seidel methods to solve $\mathbf{a}^*$. We derive the asymptotic distribution of $S_L(\mathbf{a}^*)$ via an approach similar to Lin and Guo (2015, Thm. 1).

4. Noise Parameter Estimation

In this section, we discuss the estimators of the noise distribution parameters by using the method of moment estimator. We suggest using the lag-1 sample autocovariance and the fourth moment of the observed returns to obtain the moment estimators:

$$\begin{aligned} E\left(\frac{1}{n}\sum_{j=1}^{n}\tilde{r}_j\tilde{r}_{j+1}\right) &= -E(\eta_t^2), \\ E\left(\frac{1}{n}\sum_{j=1}^{n}\tilde{r}_j^4\right) &= 3k_n^2 m_4 + 12k_n m_2 E(\eta_t^2) + 2E(\eta_t^4) + 6(E(\eta_t^2))^2, \end{aligned} \tag{4.1}$$

where m_2 and m_4 are the ones defined in (3.1).

For example, if η_t has a scaled t distribution, the moment estimators of s and ν can be solved from

$$\begin{aligned} \frac{1}{n}\sum_{j=1}^{n}\tilde{r}_j\tilde{r}_{j+1} &= \frac{-s^2\nu}{\nu-2}, \\ \frac{1}{n}\sum_{j=1}^{n}\tilde{r}_j^4 &= 3k_n^2 m_4 + 12k_n m_2\frac{s^2\nu}{\nu-2} + 12\frac{s^4\nu^2(\nu-3)}{(\nu-4)(\nu-2)^2}. \end{aligned} \tag{4.2}$$

As another example, if $\eta_t \overset{d}{=} \mathcal{E}(\beta) - \beta^{-1}$, where $\mathcal{E}(\beta)$ is the exponential distribution with expected value β^{-1}, $\varepsilon_t = \eta_t - \eta_{t-1}$ follows a double exponential distribution with parameter β. Then the moment estimator of β can be solved from

$$\frac{1}{n}\sum_{j=1}^{n}\tilde{r}_j\tilde{r}_{j+1} = -\frac{2}{\beta^2}.$$

We investigated the accuracy of the moment estimators for the volatility and noise distributions. Recall that

$$S_{nr} = \frac{E\left(\int_0^{nk_n}\sigma_s^2 ds\right)}{2n\sigma_\eta^2},$$

the signal-to-noise ratio. For the Heston model in (3.2), the parameters were set at $\rho = 10$, $\mu = 0.00048$, and $\omega = \sqrt{\rho\mu}/2$. For the Stein and Stein model in (3.3), the parameters were set at $\rho = 5$, $\mu = 0.02$, and $\omega = \sqrt{0.0008}$. In these settings, the numerator of S_{nr} was 4.8×10^{-4}. The sample size was $n = 2 \times 10^5$. To investigate the performance of the moment estimators, we conducted a simulation study for three values of S_{nr}. The parameters of noise distribution were $\nu = 5$, $s = 0.00035$ ($S_{nr} = 0.12$), $\nu = 8$, $s = 0.0002$ ($S_{nr} = 0.45$), and $\nu = 10$, $s = 0.0001$ ($S_{nr} = 1.9$) for the scaled t distribution, and were

Table 1. Relative errors of parameter estimation.

	Heston Model				Stein and Stein Model		
	$S_{nr}=0.12$	$S_{nr}=0.45$	$S_{nr}=1.9$		$S_{nr}=0.12$	$S_{nr}=0.45$	$S_{nr}=1.9$
$\hat{\alpha}_o$	0.1466	0.1537	0.1524	$\hat{\mu}_o$	0.0421	0.0429	0.0429
$\hat{\lambda}_o$	0.1571	0.1637	0.1641	$\hat{\tau}_o$	0.0933	0.0910	0.0904
$\hat{\alpha}$	0.2247	0.1584	0.1569	$\hat{\mu}$	0.0543	0.0455	0.0451
$\hat{\lambda}$	0.2265	0.1656	0.1659	$\hat{\tau}$	0.1451	0.0940	0.0910
$\hat{\nu}$	0.3586	0.2409	0.4141	$\hat{\nu}$	0.2813	0.2763	0.4418
$\hat{s}$	0.1534	0.1249	0.3958	$\hat{s}$	0.1389	0.1186	0.4113
$\hat{\beta}$	0.0304	0.1074	0.2457	$\hat{\beta}$	0.0347	0.1098	0.3849

$\beta = 2,200$ ($S_{nr} = 0.12$), β =4,300 ($S_{nr} = 0.45$), and $\beta = 9,000$ ($S_{nr} = 1.9$) for the exponential distribution.

The relative errors of $(\hat{\alpha}, \hat{\lambda})$ and $(\hat{\mu}, \hat{\tau})$ are listed in Table 1. The results demonstrate that the relative errors of the estimators of noise parameters ($\hat{\nu}$, $\hat{s}$ and $\hat{\beta}$) increase as S_{nr} increases, whereas those of the model parameter estimators $(\hat{\alpha}, \hat{\lambda})$ for the Heston model and $(\hat{\mu}, \hat{\tau})$ for the Stein and Stein model decrease as S_{nr} increases. To obtain bench mark relative errors, we considered the moment estimators for noise free SVMs, $\epsilon_i = 0$ for all i. In the same parameter settings as before, we denote the moment estimators by $(\hat{\alpha}_o, \hat{\lambda}_o)$ and $(\hat{\mu}_o, \hat{\tau}_o)$ for the SVM with no microstructure noise. To showcase the microstructure noise effect on the parameter estimation, the relative errors of $(\hat{\alpha}_o, \hat{\lambda}_o)$ and $(\hat{\mu}_o, \hat{\tau}_o)$ are also listed in Table 1, serving as benchmark values. These results indicate that $(\hat{\alpha}, \hat{\lambda})$ and $(\hat{\alpha}_o, \hat{\lambda}_o)$ perform comparably when $S_{nr} > 0.45$. An analogous phenomenon can be found in $(\hat{\mu}, \hat{\tau})$ and $(\hat{\mu}_o, \hat{\tau}_o)$.

5. Simulation and Empirical Results

5.1. Simulation results

In this simulation study, we evaluated the performance of our test

$$\hat{T}_n = nk_n \int \Big| \frac{\hat{\phi}_n(t) - \phi_{\hat{\theta}_n}(t)}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} \Big|^2 dG(t) \tag{5.1}$$

with

$$\hat{\phi}_n(t) = \frac{1}{n}\sum_{j=1}^{n} e^{it\tilde{r}_j/k_n^{1/2}},$$

$$\phi_{\hat{\theta}_n}(t) = E_\theta(e^{it\xi_1}) E_\beta\left(e^{it\eta_1/k_n^{1/2}}\right) E_\beta\left(e^{-it\eta_1/k_n^{1/2}}\right)\Big|_{\{\theta=\hat{\theta}_n, \beta=\hat{\beta}\}},$$

where $\hat{\theta}_n$ and $\hat{\beta}$ are the moment estimators of the parameters of the volatility and the microstructure noise, respectively.

We considered the null hypotheses $\mathcal{H}_{0,ij} : \sigma_t \sim f_i$ and $\eta_t \sim f_\eta^{(j)}$ for $i = 1, 2$ and $j = 1, 2$, where $f_1 \sim \sqrt{Gamma}$ (the Heston SVM, see (3.2)), $f_2 \sim Normal$ (the Stein and Stein SVM, see (3.3)), $f_\eta^{(1)} \sim st_\nu$ and $f_\eta^{(2)} \sim \mathcal{E}(\beta) - \beta^{-1}$. In the next section, we discuss the choice of microstructure noise distribution based on empirical data. Under $\mathcal{H}_{0,i1}$, we have

$$\hat{\phi}_\eta(\frac{t}{\sqrt{k_n}})\hat{\phi}_\eta(-\frac{t}{\sqrt{k_n}}) = \frac{(B_{\nu/2}(\sqrt{\nu}|st/k_n^{1/2}|))^2(\sqrt{\nu}|st/k_n^{1/2}|)^\nu}{(\Gamma(\nu/2))^2 2^{\nu-2}}\Big|_{\{\nu=\hat{\nu},s=\hat{s}\}},$$

where $B_{\nu/2}$ denotes a modified Bessel function of the third kind with index $\nu/2$. Under $\mathcal{H}_{0,i2}$, we have

$$\hat{\phi}_\eta(\frac{t}{\sqrt{k_n}})\hat{\phi}_\eta(-\frac{t}{\sqrt{k_n}}) = \frac{\beta^2}{\beta^2 + k_n^{-1}t^2}\Big|_{\{\beta=\hat{\beta}\}}.$$

As in Lin, Lee, and Guo (2013), we adopted a strong order-one approximation of the Ornstein-Uhlenbeck process to attain better approximation. See, for example, Schurz (2000, p.242) and Fan (2005):

$$\begin{cases} p_{t+1} = p_t + \sigma_t\sqrt{k_n}Z_t, \\ \sigma_t^2 = \sigma_{t-1}^2 + \rho(\kappa - \sigma_{t-1}^2)k_n + v(\sigma_{t-1})\sqrt{k_n}W_t + \frac{1}{2}v(\sigma_{t-1})v'(\sigma_{t-1})k_n(W_t^2 - 1). \end{cases}$$

Our simulation scheme was similar to that of Lin, Lee, and Guo (2013), and the key steps were as follows.

1. Simulate a sample $\{p_i\}_{1\le i\le n}$ from a hypothesized SVM and $\{\eta_i\}_{1\le i\le n}$ with the corresponding true parameters θ and $\beta = (s, \nu)$.
2. Obtain the log prices $\tilde{p}_i = p_i + \eta_i$, the log returns $\tilde{r}_i = \tilde{p}_i - \tilde{p}_{i-1}$, and the normalized returns $\hat{\xi}_i$ defined in (2.2).
3. The parameter estimators of noise distribution and the model parameters, denoted by $\hat{\beta}_n$ and $\hat{\theta}_n$, respectively, are obtained from (4.2) and (3.3) for $\mathcal{H}_{0,1}$ and (4.2) and (3.4) for $\mathcal{H}_{0,2}$. Finally, $\hat{T}_n$ is obtained from (5.1).
4. Generate B bootstrap samples of size n by replacing $\hat{\theta}_n$ and $\hat{\beta}_n$ to the model and noise parameters, respectively. Similarly to Steps 2 and 3, obtain the bootstrap moment estimators to construct the bootstrap test statistics $\hat{T}_n^{*b}$ from (5.1), $b = 1, \ldots, B$. As in Section 5 of Lin, Lee, and Guo (2013), we simply set $\rho = 10$.
5. Use the B bootstrap test statistics $\hat{T}_n^{*b}$ to estimate the sample $(1-\alpha)$th quantile. Repeat Steps 1 to 3 1,000 times to obtain the sizes and powers.

The parameter settings were the same as described in Section 4, which correspond to three S_{nr} values (0.12, 0.45 and 1.9). The sizes and powers of the test

Table 2. The sizes and powers (in percentage) of $\hat{T}_n$ for $\mathcal{H}_{0,i1}$, $i = 1, 2$ versus five $\mathcal{H}_1$.

	$\mathcal{H}_{0,11} : \sqrt{Gamma}$			$\mathcal{H}_{0,21}$: Normal		
$\mathcal{H}_1$	$S_{nr} = 0.12$	$S_{nr} = 0.45$	$S_{nr} = 1.9$	$S_{nr} = 0.12$	$S_{nr} = 0.45$	$S_{nr} = 1.9$
$\sqrt{Gamma}$	5.2	5.3	4.6	78.1	89.9	83.4
Normal	48.1	35.5	76.2	3.6	4.3	5.3
Uniform	99.9	99.9	100.0	94.8	87.3	96.8
F	82.5	70.8	90.2	90.6	99.7	99.4
IG	33.2	93	93	70.3	98.7	99.2

Table 3. The sizes and powers (in percentage) of $\hat{T}_n$ for $\mathcal{H}_{0,i2}$, $i = 1, 2$ versus five $\mathcal{H}_1$.

	$\mathcal{H}_{0,12} : \sqrt{Gamma}$			$\mathcal{H}_{0,22}$: Normal		
$\mathcal{H}_1$	$S_{nr} = 0.12$	$S_{nr} = 0.45$	$S_{nr} = 1.9$	$S_{nr} = 0.12$	$S_{nr} = 0.45$	$S_{nr} = 1.9$
$\sqrt{Gamma}$	3.6	4.7	5.8	61.7	61.5	50.9
Normal	82.3	93.8	90.8	6.5	5.1	5.2
Uniform	95.9	88.9	70.8	76.5	61.5	60.5
F	70.9	99.8	65.7	55.3	61.6	61.7
IG	68.7	82.7	94.1	67.1	75.8	87.8

statistic $\hat{T}_n$ for $\mathcal{H}_{0,ij}$, $i = 1, 2$, $j = 1, 2$, versus $\mathcal{H}_1$'s corresponding to the volatility distributions $\sqrt{Gamma}$, Normal, Uniform, F, and Inverse Gamma (IG), are presented in Table 2 ($\mathcal{H}_{0,i1}$, $i = 1, 2$,) and Table 3 ($\mathcal{H}_{0,i2}$, $i = 1, 2$). These results support the validity of our test.

5.2. Data analysis

We considered the ultra high frequency tick-by-tick data of 13 stocks listed on the New York Stock Exchange (NYSE): ABT, AMD, BAC, C, GE, JNJ, JPM, KO, MCD, MER, NOK, PEP, XOM. The normal trading hours of NYSE is 6.5 hours from 9:30 to 16:00. Here, we use the previous tick interpolation scheme (see, for example, Dacorogna et al. (2001)) to obtain the equi-spaced log prices $\tilde{p}_{t_i}$'s for each stock. To preprocess the suspicious jumps, we applied the wavelet method of Fan and Wang (2007). The following analyses are based on the log returns after the jumps are smoothed.

We first discuss appropriate microstructure noise distributions through a high frequency data analysis. For this, we consider three stocks with different transaction frequencies: ABT (low frequency), GE (middle frequency), and JPM (high frequency). The nominal returns are $r_i = \int_{t_{i-1}}^{t_i} \sigma_s dW_s = \sigma_{t_{i-1}}(W_{t_i} - W_{t_{i-1}}) + k_n^{1/2}\Delta_{ni}$, where $\Delta_{ni} = \int_{t_{i-1}}^{t_i}(\sigma_s - \sigma_{t_{i-1}})dW_s/k_n^{1/2}$, and thus,

$$Var(r_i) = k_n E\Big(\frac{\sigma_{t_{i-1}}(W_{t_i} - W_{t_{i-1}})}{k_n^{1/2}} + \Delta_{ni}\Big)^2 = O(k_n),$$

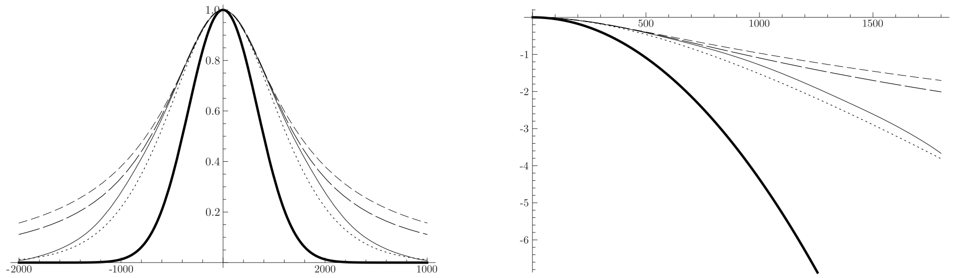

Figure 1. The left panel includes the empirical characteristic function (solid line) of the JPM and the fitted characteristic functions of ε_t for η_t following normal (thick line), generalized error (- - -), scaled log uniform (— — —) and scaled t $(\cdots)$ distributions. The right panel includes the corresponding log characteristic functions.

which implies $r_i = O_P(k_n^{1/2})$. Hence, when n is large (k_n is small) and x is not so small, we have

$$P\left(\tilde{r}_i \leq x\right) = P\left(r_i + \varepsilon_i \leq x\right) \approx P\left(\varepsilon_i \leq x\right). \tag{5.2}$$

This suggests that at the highest observed frequency, the empirical distribution of the observed returns might resemble the microstructure noise distribution. Since the microstructure noise $\varepsilon_t = \eta_t - \eta_{t-1}$ and η_t's are i.i.d., the distribution of ε_t can be obtained as the convolution of the density function of η_t. As for the candidate distributions of $\{\eta_t\}$, below we considered a scaled t distribution, a normal distribution, a generalized error distribution, and an exponential distribution.

Consider the JPM case. In Figure 1 (left panel), we plot the empirical characteristic function (solid line) of the JPM and the fitted characteristic functions of $\varepsilon_t = \eta_t - \eta_{t-1}$ for η_t following a normal (thick line), a generalized error (- - -), an exponential (— — —), and a scaled t $(\cdots)$ distribution, respectively. The corresponding log empirical/fitted characteristic functions versus $\log(t)$ are plotted in the right panel of Figure 1. The parameters of these fitted characteristic functions were estimated by the method of moments. The scaled t distribution visually makes the best fit for the microstructure noise distribution of the JPM. Likewise, the scaled t distribution provides the best fit in the cases of ABT and GE. From this, we selected the scaled t distribution as our candidate distribution for η_t.

To investigate the microstructure noise effect on model testing, both 2-minute and 5-minute returns were considered. We utilized the high frequency transaction data of the 21 trading days in the period 2002/01/02$\sim$2002/01/31. As with the setting of Lin, Lee, and Guo (2013), we regarded one hour as a time unit and overnight returns were ignored. For the 5-minute returns, we set the sampling time length at 60min$\times k_n \approx$ 5min. Thus, for each stock through 21 normal trading days, the sample size was $n = 1{,}638$. For 2-minute returns, the sample size was $n = 4{,}095$ for each stock. We utilized high frequency returns to test the null hypotheses $\mathcal{H}_{0,i1} : \sigma_t \sim f_i$ and $\eta_t \sim$ scaled $t, i = 1, 2$, where $f_1 \sim \sqrt{Gamma}$ (Heston model) and $f_2 \sim Normal$ (Stein and Stein model) for each stock.

We performed the proposed test $\hat{T}_n$ (cf. (5.1)) at the nominal level 5%. For a comparison, we also considered the test $\tilde{T}_n$ in Lin, Lee, and Guo (2013), where

$$\tilde{T}_n = n_k \int \Big| \frac{1}{n} \sum_{j=1}^{n} e^{it\tilde{r}_j/k_n} - \tilde{\phi}_\xi(t) \Big|^2 dG(t),$$

$\tilde{r}_j,\ j = 1, \ldots, n$, are high frequency returns, and $\tilde{\phi}_\xi(t) = E_\theta(e^{it\xi_1})\big|_{\theta=\tilde{\theta}}$ with the parameter $\tilde{\theta}$ estimated based on the noise free model, see Section 4 of Lin, Lee, and Guo (2013).

We chose both the 2-minute and 5-minute returns due to the reasons addressed below. Since the variance of the efficient returns is proportional to the sampling frequency (see Section 3 for detail), in the 5-minute return case, the test statistic $\tilde{T}_n$ should have a tendency to have a smaller bias owing to the microstructure noise. However, at the same time, the total sample size decreases and this results in an increase of the variance of $\tilde{T}_n$. Conversely, the total sample size increases in the 2-minute return case and the variance of $\tilde{T}_n$ gets lower than that in the 5-minute return case. In the meantime, the effect of the microstructure noise becomes more prominent and this increases the bias owing to the microstructure noise. The results are summarized as follows.

(i) $\hat{T}_n$ accepts $\mathcal{H}_{0,11}$ and rejects $\mathcal{H}_{0,21}$ for all 13 stocks in both the 2-minute or 5-minute return cases.

(ii) $\tilde{T}_n$ yields the same result for the three stocks ABT, BAC, and PEP, but rejects both $\mathcal{H}_{0,11}$ and $\mathcal{H}_{0,21}$ in the cases of AMD, C, GE, JPM, and MCD. A main problem in using $\tilde{T}_n$ is that the obtained result varies with the sampling frequency since the microstructure noise term is not taken into consideration. For example, in the cases of JNJ, MER, and XOM, $\tilde{T}_n$ accepts $\mathcal{H}_{0,11}$ and rejects $\mathcal{H}_{0,21}$ in the 5-minute return case, while it rejects both hypotheses in the 2-minute return case. For KO and NOK, though, $\tilde{T}_n$ accepts both the $\mathcal{H}_{0,11}$ and $\mathcal{H}_{0,21}$ when 5-minute returns are used, it rejects $\mathcal{H}_{0,21}$ when 2-minute returns are used.

The summary in (ii) indicates that $\hat{T}_n$ yields more consistent results, reflects the situation a lot better, and yields more accurate results than $\tilde{T}_n$.

To explore the power of $\hat{T}_n$ in testing the microstructure noise distribution, we considered the null hypothesis: $\mathcal{H}_{0,12} : \sigma_t \sim \sqrt{Gamma}$ (Heston model) and

$\eta_t \sim \mathcal{E}(\beta) - \beta^{-1}$. Comparing $\mathcal{H}_{0,12}$ to $\mathcal{H}_{0,11}$, we kept the same distribution assumption on σ_t, yet changed the one on η_t. The test $\hat{T}_n$ rejected $\mathcal{H}_{0,12}$ for all 13 stocks for the 2-minute return case. This indicates that the proposed $\hat{T}_n$ has power in testing the the distribution assumption on η_t, and that the scaled t distribution is preferable to $\mathcal{E}(\beta) - \beta^{-1}$ for the microstructure noise distribution.

6. Concluding Remarks

In this study, a goodness-of-fit test is proposed for continuous time stochastic volatility models contaminated with microstructure noises. A focus is made on the stationary marginal distribution of the volatility process. The proposed test is designed to measure the deviations between the empirical and hypothesized true characteristic functions divided by the characteristic function of the microstructure noise. It is shown that under the null, the proposed test asymptotically follows a weighted sum of products of centered normal random variables. A simulation study was conducted to evaluate the proposed test. Our data analysis shows that our test outperforms the test of Lin, Lee, and Guo (2013) in terms of accuracy and practicability. Overall, our findings support the validity of the proposed test in the presence of microstructure noise.

Acknowledgement

The first and third authors acknowledge that this work was supported in part by the National Science Council, Taiwan, under grant NSC 100-2118-M-110-003 and NSC 103-2118-M-110-003. The first author acknowledges that this work was supported by Ministry of Science and Technology, Taiwan, under grant MOST 104-2118-M-006-001. The second author acknowledges that this work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIP) (No. 2015R1A2A2A010003894). This work is partially supported by the National Center for Theoretical Sciences (South), Tainan, Taiwan.

Appendix. Proofs

The proof of Theorem 1. Since (2.12) is reexpressed as

$$\hat{T}_n = nk_n \int \left| \frac{\hat{\phi}_n(t) - \hat{\phi}_\xi(t)\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} \right|^2 dG(t)$$

$$= nk_n \int \left| \frac{\frac{1}{n}\sum_{j=1}^n e^{it\tilde{r}_j/k_n^{1/2}} - \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j}\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} \right.$$

$$\left. + \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j} - \hat{\phi}_\xi(t) \right|^2 dG(t),$$

we get

$$
\begin{aligned}
|\hat{T}_n - \hat{T}_n^*| &\le nk_n \int \left| \frac{\frac{1}{n}\sum_{j=1}^n e^{it\tilde{r}_j/k_n^{1/2}} - \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j}\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} \right|^2 dG(t) \\
&= nk_n \int \left| \frac{\frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j}\left[e^{it(\eta_j-\eta_{j-1})/k_n^{1/2}} - \hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})\right]}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} \right|^2 dG(t) \\
&\le 2nk_n \int \left| \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j}\left[\frac{e^{it(\eta_j-\eta_{j-1})/k_n^{1/2}} - \phi_\eta(t/k_n^{1/2})\phi_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})}\right] \right|^2 dG(t) \\
&\quad +2nk_n \int \left| \frac{\phi_\eta(t/k_n^{1/2})\phi_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2})\hat{\phi}_\eta(-t/k_n^{1/2})} - 1 \right|^2 dG(t) \\
&= 2*(J1) + 2*(J2). \qquad\qquad (A.1)
\end{aligned}
$$

We verify that both $(J1)$ and $(J2)$ are asymptotically negligible. By using Taylor's theorem and (A6), we can write

$$\hat{\phi}_\eta(t) = \phi_\eta(t) + (\nabla_\beta \phi_\eta(t))'(\hat{\beta} - \beta),$$

and thus,

$$(J2) = nk_n \int \left| \frac{1}{1 + h(t) + h(-t) + h(t)h(-t)} - 1 \right|^2 dG(t),$$

where $h(t) = \left[\nabla_\beta \log(\phi_\eta(t/\sqrt{k_n}))\right]' (\hat{\beta} - \beta)$. Further, owing to (A8)(i), we get

$$|\nabla_\beta \log \phi_\eta(\frac{t}{\sqrt{k_n}})| = -\nabla_\beta \alpha_0(\beta) \left|\frac{t}{\sqrt{k_n}}\right|^{\alpha_1} + \nabla_\beta R(\beta, \frac{t}{\sqrt{k_n}}).$$

This together with (A6) implies

$$
\begin{aligned}
|h(t)| &= -(\nabla_\beta \alpha_0(\beta))' \left|\frac{t}{\sqrt{k_n}}\right|^{\alpha_1} (\hat{\beta} - \beta) + \left(\nabla_\beta R(\beta, \frac{t}{\sqrt{k_n}})\right)' (\hat{\beta} - \beta) \\
&= O_p\left(|t|^{\alpha_1} k_n^{-\alpha_1/2} n^{-1/2}\right) + o_p\left(|t|^{\alpha_1} k_n^{-\alpha_1/2} n^{-1/2}\right) \\
&= O_p\left(|t|^{\alpha_1} (nk_n^{\alpha_1})^{-1/2}\right). \qquad\qquad (A.2)
\end{aligned}
$$

Subsequently, since $\alpha_1 < 1$,

$$(J2) = nk_n \int O_p\left(|t|^{2\alpha_1}(nk_n^{\alpha_1})^{-1}\right) dG(t) = O_p\left(k_n^{1-\alpha_1} \int |t|^{2\alpha_1} dG(t)\right) = o_p(1).$$

Meanwhile, notice that

$$(J1) = nk_n \int \left| \frac{1}{n}\sum_{j=1}^n e^{it\hat{\xi}_j}\left[\frac{e^{it(\eta_j-\eta_{j-1})/\sqrt{k_n}}/(\phi_\eta(\frac{t}{\sqrt{k_n}})\phi_\eta(-\frac{t}{\sqrt{k_n}})) - 1}{\hat{\phi}_\eta(\frac{t}{\sqrt{k_n}})\hat{\phi}_\eta(-\frac{t}{\sqrt{k_n}})/(\phi_\eta(\frac{t}{\sqrt{k_n}})\phi_\eta(-\frac{t}{\sqrt{k_n}}))}\right] \right|^2 dG(t)$$

$$= nk_n \int \left| \frac{1}{n} \sum_{j=1}^{n} e^{it\hat{\xi}_j} \left[\frac{e^{it(\eta_j - \eta_{j-1})/\sqrt{k_n}} / (\phi_\eta(t/\sqrt{k_n})\phi_\eta(-t/\sqrt{k_n})) - 1}{1 + h(t) + h(-t) + h(t)h(-t)} \right] \right|^2 dG(t).$$

By (A.2), we have $h(t) = O_p\left(|t|^{\alpha_1}(nk_n^{\alpha_1})^{-1/2}\right) = o_p(1)$ since $nk_n^{\alpha_1} \to \infty$ as $n \to \infty$. Therefore,

$$(J1) = nk_n \int \left| \frac{1}{n} \sum_{j=1}^{n} Y_j \right|^2 dG(t),$$

where

$$Y_j = e^{it\hat{\xi}_j} \Big[\frac{e^{it(\eta_j - \eta_{j-1})/\sqrt{k_n}}}{\phi_\eta(t/\sqrt{k_n})\phi_\eta(-t/\sqrt{k_n})} - 1 \Big].$$

We derive the mean and the variance of $\bar{Y} = \sum_{j=1}^{n} Y_j$. Since $\{\eta_t\}$ is a white noise process and independent of $\{p_t\}$, it is immediate that

$$E\Big[\frac{1}{n} \sum_{j=1}^{n} Y_j\Big] = 0.$$

To handle the variance, let Y_j^* be the complex conjugate of Y_j. Then, since Y_j has only one-step correlation, we have

$$\begin{aligned} Var(\bar{Y}) &= E\Big[\Big(\frac{1}{n}\sum_{j=1}^{n} Y_j\Big)\Big(\frac{1}{n}\sum_{j=1}^{n} Y_j^*\Big)\Big] \\ &= \frac{1}{n^2}\sum_{j=1}^{n} E(Y_j Y_j^*) + \frac{1}{n^2}\sum_{j=1}^{n-1} E(Y_j Y_{j+1}^*) + \frac{1}{n^2}\sum_{j=1}^{n-1} E(Y_{j+1} Y_j^*) \\ &= (J1-1) + (J1-2) + (J1-3). \end{aligned} \tag{A.3}$$

By simple algebra, we can check that

$$(J1-1) = \frac{1}{n}\Big[\frac{1}{\phi_\eta^2(t/\sqrt{k_n})\phi_\eta^2(-t/\sqrt{k_n})} - 1\Big]. \tag{A.4}$$

Owing to (A8)(i),

$$\left|\phi_\eta^2(\frac{t}{\sqrt{k_n}})\phi_\eta^2(-\frac{t}{\sqrt{k_n}})\right| = e^{-4\alpha_0(\beta)|t|^{\alpha_1}k_n^{-\alpha_1/2} + R'(\beta, t/\sqrt{k_n})},$$

where $R'(\beta, t/\sqrt{k_n}) = 2R(\beta, t/\sqrt{k_n}) + 2R(\beta, -t/\sqrt{k_n})$ is of order $o(|t/\sqrt{k_n}|^{\alpha_1})$. Thus, by (A.4), we have

$$|(J1-1)| = \left|\frac{1}{n}\left[e^{4\alpha_0(\beta)|t|^{\alpha_1}k_n^{-\alpha_1/2} - R'(\beta, t/\sqrt{k_n})} - 1\right]\right|$$

$$\leq \left| \frac{1}{n} \left(4\alpha_0(\beta)|t|^{\alpha_1} k_n^{-\alpha_1/2} - R'(\beta, \frac{t}{\sqrt{k_n}}) \right) \right| \tag{A.5}$$

$$= O\left(|t|^{\alpha_1} (n k_n^{\alpha_1/2})^{-1} \right), \tag{A.6}$$

where the inequality in (A.5) holds due to the fact that $|e^x - 1| \leq |x| \; \forall x$.

Similarly, we have

$$|(J1-2)| = \Big| \frac{1}{n^2} \sum_{j=1}^{n-1} E\left[e^{it(\hat{\xi}_{j+1} - \hat{\xi}_j)} \right] \left[\frac{\phi_\eta(-2t/\sqrt{k_n})}{\phi_\eta(t/\sqrt{k_n})\phi_\eta(-t/\sqrt{k_n})} - 1 \right] \Big|$$

$$\leq \left| \frac{1}{n} \left[(2 - 2^{\alpha_1})\alpha_0(\beta) \Big| \frac{t}{\sqrt{k_n}} \Big|^{\alpha_1} \right] \right|$$

$$= O\left(|t|^{\alpha_1} (n k_n^{\alpha_1/2})^{-1} \right), \tag{A.7}$$

$$|(J1-3)| = O\left(|t|^{\alpha_1} (n k_n^{\alpha_1/2})^{-1} \right). \tag{A.8}$$

Then, combining (A.3), (A.6), (A.7), and (A.8), we have

$$|(J1)| \leq n k_n \int O_p\left(|t|^{\alpha_1} (n k_n^{\alpha_1/2})^{-1} \right) dG(t) = O_p\left(k_n^{1-\alpha_1/2} \int |t|^{\alpha_1} dG(t) \right) = o_p(1).$$

This validates the theorem.

The proof of Theorem 3. We follow the lines in the proof of Theorem 2, so highlight only some key steps. First, we show that $(J1)$ and $(J2)$ defined in (A.1) are $o_p(1)$. Recall $h(t) = \left[\nabla_\beta \log(\phi_\eta(t/\sqrt{k_n})) \right]' (\hat{\beta} - \beta)$. Owing to (A8)(ii), we have

$$\left| \nabla_\beta \log(\phi_\eta(\frac{t}{\sqrt{k_n}})) \right| = \frac{\nabla_\beta \alpha_0(\beta)}{\alpha_0(\beta)} - \nabla_\beta \alpha_1(\beta) \log |\frac{t}{\sqrt{k_n}}| + \frac{\nabla_\beta R(\beta, t/\sqrt{k_n})}{R(\beta, t/\sqrt{k_n})}.$$

This together with (A6) implies

$$|h(t)| = O_p\left((-\log|t| + 2^{-1}\log k_n) n^{-1/2} \right) + O_p\left(|t|^{-\alpha_3} k_n^{\alpha_3/2} n^{-1/2} \right)$$

$$= O_p\left(n^{-1/2} (\log|t| + \log k_n) \right).$$

Therefore,

$$(J2) = n k_n \int O_p\left((\log|t| + \log k_n)^2 n^{-1} \right) dG(t)$$

$$= O_p\left(k_n \int (\log|t|)^2 dG(t) + k_n (\log k_n)^2 \right) = o_p(1).$$

To show that $(J1) = o_p(1)$, we deduce the orders of the three terms in (A.3). By (A8)(ii), we have

$$\left| \phi_\eta^2(\frac{t}{\sqrt{k_n}}) \phi_\eta^2(-\frac{t}{\sqrt{k_n}}) \right| = O\left(|t|^{-2\alpha_1(\beta)} k_n^{\alpha_1(\beta)} \right).$$

Using (A.4) and the fact that $|e^x - 1| \leq |x|\ \forall x$, we get

$$
\begin{aligned}
|(J1-1)| &= \left| \frac{1}{n} O\left(|t|^{2\alpha_1(\beta)} k_n^{-\alpha_1(\beta)} - 1 \right) \right| \\
&= \left| \frac{1}{n} O\left(e^{2\alpha_1(\beta)\log|t| - \alpha_1(\beta)\log k_n} - 1 \right) \right| \\
&= O\left(n^{-1}\log|t| + n^{-1}\log k_n \right).
\end{aligned}
$$

Similarly, owing to (A8)(ii), (A.7), and (A.8), $|(J1{-}2)| = O\left(n^{-1}\log|t| + n^{-1}\log k_n\right)$ and $|(J1-3)| = O\left(n^{-1}\log|t| + n^{-1}\log k_n\right)$. Therefore, we get

$$
\begin{aligned}
|(J1)| &\leq n k_n \int O_p\left(n^{-1}\log|t| + n^{-1}\log k_n \right) dG(t) \\
&= O_p\left(k_n \int \log|t| dG(t) + k_n \log k_n \right) = o_p(1),
\end{aligned}
$$

which proves the theorem.

The proof of Lemma 2. The second and fourth moments of the nominal return r_j are

$$
\begin{aligned}
\sum_{j=1}^{n+1} E\left(r_j^2\right) &= \sum_{j=1}^{n+1} E\Big(\int_{(j-1)k_n}^{jk_n} \sigma_s dW_s \Big)^2 = \sum_{j=1}^{n} E\Big(\int_{(j-1)k_n}^{jk_n} \sigma_s^2 ds \Big) \\
&= E\Big(\int_0^{nk_n} \sigma_s^2 ds \Big), \\
E\Big(\sum_{j=1}^{n+1} r_j^4 \Big) &= k_n^2 \sum_{j=1}^{n+1} E\Big(\frac{r_j}{k_n^{1/2}} \Big)^4 = k_n^2 \sum_{j=1}^{n+1} E\left(\xi_j^4 + 6\xi_j^2 \Delta_{nj}^2 + \Delta_{nj}^4 \right) \\
&= 3k_n^2 \sum_{j=1}^{n+1} E\left(\sigma_{t_{j-1}}^4 \right) + 6k_n^2 \sum_{j=1}^{n+1} E\left(\sigma_{t_{j-1}}^2 \right) E(\Delta_{nj}^2) + k_n^2 \sum_{j=1}^{n+1} E(\Delta_{nj}^4) \\
&= 3k_n E\left(\int_0^{nk_n} \sigma_s^4 ds \right) + O(nk_n^3).
\end{aligned}
$$

The remainder of the proof follows essentially the same lines as does in the proof of Lemma 2 of Lin and Guo (2015)

References

Aït-Sahalia, Y. (1996a). Nonparametric pricing of interest rate derivative securities. *Econometrica* **64**, 527-560.

Aït-Sahalia, Y. (1996b). Testing continuous-time models of the spot interest rate. *Rev. Finan. Studies* **9**, 385-426.

Aït-Sahalia, Y., Mykland, P. A. and Zhang, L. (2005). How often to sample a continuous-time process in the presence of market microstructure noise. *Rev. Finan. Studies* **18**, 351-416.

Andersen, T. G., Bollerslev, T., Diebold, F. X. and Ebens, H. (2001). The distribution of realized stock return volatility. *J. Finan. Econom.* **61**, 43-76.

Barndorff-Nielsen, O. E. and Shephard, N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *J. Roy. Statist. Soc. Ser. B* **64**, 253-280.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica* **76**, 1481-1536.

Barndorff-Nielsen, O. E., Hansen, P. R., Lunde, A. and Shephard, N. (2009). Realized kernels in practice: trades and quotes. *Econom. J.* **12**, C1-C32.

Bandi, F. M. and Russell, J. R. (2006). Separating microstructure noise from volatility. *J. Finan. Econom.* **79**, 655-692.

Bandi, F. M. and Russell, J. R. (2008). Microstructure noise, realized variance, and optimal sampling. *Rev. Econom. Studies* **75**, 339-369.

Bibby, B. M., Skovgaard, I. M. and Sørensen, M. (2005). Diffusion-type models with given marginal distribution and autocorrelation function. *Bernoulli* **11**, 191-220.

Chen, S. X., Gao, J. and Tang, C. Y. (2008). A test for model specification of diffusion processes. *Ann. Statist.* **36**, 167-198.

Dacorogna, M. M., Gencay, R., Müller, U. A., Olsen, R. B. and Pictet, O. V. (2001). *An Introduction to High-Frequency Finance.* Academic Press, San Diego, CA.

Daubechies, I. (2002). *Ten Lectures on Wavelets.* Society for Industrial and Applied Mathematics, Philadelphia, PA.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.* **19**, 1257-1272.

Fan, J. (2005). A selective overview of nonparametric methods in financial econometrics. *Statist. Sci.* **20**, 317-357.

Fan, J. and Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *J. Amer. Statist. Assoc.* **102**, 1349-1362.

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (1998). Limit theorems for discretely observed stochastic volatility models. *Bernoulli* **4**, 283-303.

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (1999). Parameter estimation for discretely observed stochastic volatility models. *Bernoulli* **5**, 855-872.

Genon-Catalot, V., Jeantheau, T. and Larédo, C. (2000). Stochastic volatility as hidden Markov models and statistical applications. *Bernoulli* **6**, 1051-1079.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Rev. Finan. Studies* **6**, 327-343.

Jacod, J., Li, Y., Mykland, P. A., Podolskij, M. and Vetter, M. (2009). Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Process. Appl.* **119**, 2249-2276.

Kessler, M. (2000). Simple and explicit estimating functions for a discretely observed diffusion process. *Scand. J Statist.* **27**, 65-82.

Lee, S. (2010). On the goodness of fit test for discretely observed sample from diffusion processes: divergence measure approach. *J. Korean Math. Soc.* **47**, 1137-1146.

Lin, L.-C. and Guo, M. (2015). Optimal restricted quadratic estimator of integrated volatility. *Ann. Inst. Stat. Math.* DOI: 10.1007/s10463-015-0507-z.

Lin, L.-C., Lee, S. and Guo, M. (2013). Goodness-of-fit test for stochastic volatility models. *J. Multivariate Anal.* **116**, 473-498.

Lin, L.-C., Lee, S. and Guo, M. (2014). The Bickel-Rosenblatt test for continuous time stochastic volatility models. *TEST* **23**, 195-218.

Prakasa Rao, B. L. S. (1999). *Statistical Inference for Diffusion Type Processes*. Arnold, London.

Reiss, M. (2011). Asymptotic equivalence for inference in the volatility from noisy observations. *Ann. Statist.* **39**, 772-802.

Schurz, H. (2000). Numerical analysis of stochastic differential equations without tears. In *Handbook of Stochastic Analysis and Applications* (D. Kannan and V. Laksh- mikantham, eds.), 237-359. Dekker, New York.

Stein, E. M. and Stein, J. C. (1991). Stock price distributions with stochastic volatility: an analytic approach. *Rev. Finan. Studies* **4**, 727-752.

Tsay, R. S. (2010). *Analysis of Financial Time Series*. 3rd edition. Wiley, New Jersey.

Zhang, L. (2006). Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach. *Bernoulli* **12**, 1019-1043.

Zhang, L., Mykland, P. A. and Aït-Sahalia, Y. (2005). A tale of two time scales: determining integrated volatility with noisy high-frequency data. *J. Amer. Statist. Assoc.* **100**, 1394-1411.

Department of Statistics, National Cheng Kung University, Tainan, 701, R.O.C.

Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan, 804, R.O.C.

E-mail: lclin@mail.ncku.edu.tw

Statistics, Seoul National University, Seoul, 151-742, South Korea.

E-mail: sylee@states.snu.ac.kr

Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan, 804, R.O.C.

E-mail: guomh@math.nsysu.edu.tw

(Received June 2015; accepted November 2015)