<table>
<thead>
<tr>
<th><strong>Statistica Sinica Preprint No:</strong> SS-2015-0215</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title</strong></td>
</tr>
<tr>
<td><strong>Manuscript ID</strong></td>
</tr>
<tr>
<td><strong>URL</strong></td>
</tr>
<tr>
<td><strong>DOI</strong></td>
</tr>
</tbody>
</table>
| **Complete List of Authors** | Liang-Ching Lin  
Sangyeol Lee and  
Meihui Guo |
| **Corresponding Author** | Meihui Guo |
| **E-mail** | guomh@math.nsysu.edu.tw |

Notice: Accepted version subject to English editing.
Goodness-of-fit Test for the SVM Based on Noisy Observations

Liang-Ching Lin\textsuperscript{1,3}, Sangyeol Lee\textsuperscript{2} and Meihui Guo\textsuperscript{3}

\textsuperscript{1}National Cheng Kung University and \textsuperscript{2}Seoul National University and \textsuperscript{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 asymptotically follows a weighted sum of products of centered normal random variables. To evaluate the proposed test, we conduct a simulation study using a bootstrap method. A real 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 in studying various 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, chapter 5). It is conventionally assumed that, instead of observing the efficient log-price $p$ at transaction time $t_i$, we observe $p$ with noise:

\[ \tilde{p}_{t_i} = p_{t_i} + \eta_{t_i}, \]

where $\{\eta_{t_i}\}$ are i.i.d. noises with mean zero and variance $\sigma^2_\eta$ and are independent of the process $p$. The noise term $\eta$ represents a microstructure contamination owing to imperfections of trading processes. See, for instance, Aït-Sahalia et al. (2005) and Bandi and Russell (2006). This microstructure noise results from either the information or non-information related factors such 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 create 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 et al. (2005), Zhang et al. (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) as follows:

\[
\begin{align*}
 dp_t &= \sigma_t dW_t, \\
 d\sigma_t^2 &= b(\sigma_t^2)dt + \sqrt{v(\sigma_t^2)}dB_t, \quad \sigma_0^2 = \zeta,
\end{align*}
\]

where \((B_t, W_t)_{t>0}\) is a two dimensional standard Brownian motion, \(\sigma_t^2\) is the instantaneous volatility at time \(t\), and \(\zeta\) 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 a case, the jump component can be smoothed by a wavelet method as in Fan and Wang (2007). Thus, in this study, we focus on Model (1.1) with no price jumps.

Modeling of the SVM (1.1) emphasizes on the specification of the diffusion coefficient \(v\) of the volatility process \(\{\sigma_t^2\}\), which 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 \(\sigma_t^2\): see Aït-Sahalia (1996a, b), Bibby et al. (2005) and Chen et al. (2008), it can be well specified through a goodness of fit test for the stationary distribution of \(\sigma_t^2\): see the hypothesis testing problem in (2.1) below. 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. In this study, we aim to extend the method of Lin, Lee and Guo (2013, 2014) to the observed price \(\tilde{p}\) of SVM (1.1) with microstructure noises \(\eta\). Specifically, we use the goodness of fit
test based on measuring deviations 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. The current issue is much more challenging than our previous study, since not only the volatility process is latent but also the price process is contaminated with noises.

The organization of this 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 popular SVMs for illustration. In Section 4, we study the noise parameter estimation and discuss the performances of model and noise parameter estimation. In Section 5, simulation and empirical studies are conducted. Concluding remarks are provided in Section 6. Proofs are given in the Appendix.

2. Main Result

We begin this section with presenting the regularity conditions for $b$ and $v$.

(A1) The functions $b(x)$ and $v(x)$ defined on $(0, \infty)$ satisfy

$$b(x) \in C^1, \quad v(x) \in C^2 \quad \text{for all } x > 0,$$

and there exists $K > 0$ such that for all $x > 0$,

$$|b(x)| \leq K(1 + |x|) \quad \text{and} \quad v(x) \leq K(1 + x^2).$$

(A2) The scale and speed densities of the diffusion $\sigma_t^2$

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

satisfy

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

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

Next, we impose some conditions on the stationary density of $\sigma_t^2$ given by

$$f_{\sigma, \theta}(x) = \frac{1}{M} m(x) 1_{[x > 0]},$$
where $\theta$ denotes the true parameter.

(A3) The initial random variable $\sigma_0^2 = \zeta$ has the density function $f_{\sigma, \theta}$ and
\[
\int_0^\infty |v|^\nu f_{\sigma, \theta}(v)dv < \infty \quad \text{for some } \nu \geq 2.
\]

(A4) For all $q \geq 1$, there exist constants $C_q > 0$ such that for
\[
E_{\theta}|\sigma_s - \sigma_t|^{2q} \leq C_q|t - s|^q.
\]

(A1) and (A2) ensure that the unique solution of $\sigma_t^2$ 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 et al. (1998, 1999, 2000).

We assume $\tilde{p}_t$ are observed at the equispaced time points $(t_1, t_2, \cdots, t_n)$, where $t_i = ik_n$ with $k_n \rightarrow 0$, $nk_n \rightarrow \infty$ and $nk_n^2 \rightarrow 0$ as $n \rightarrow \infty$. In this case, the observed log return at time $t_i$ is
\[
\tilde{r}_i = \tilde{p}_t - \tilde{p}_{t-1} = r_i + \varepsilon_i,
\]
where $r_i = p_t - p_{t-1}$ denotes the nominal return and $\varepsilon_i = \eta_t - \eta_{t-1}$. Since $\eta_i$'s are i.i.d. random variables with variance $\sigma_\eta^2$, the noise process $\{\varepsilon_i\}$ is an MA(1) process with variance
\[
Var(\varepsilon_i) = \sigma_\varepsilon^2 = 2\sigma_\eta^2.
\]

The distribution of $\varepsilon_i$ can be obtained from the marginal distribution of $\eta$ using a convolution method. We further assume that $\eta_t$ has a stationary density $f_{\eta, \beta}$ wherein $\beta$ 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 two families of density functions and suppose that one wishes to test the hypotheses
\[
\mathcal{H}_0 : \sigma_t^2 \sim f_{\sigma, \theta} \quad \text{and} \quad \eta_t \sim f_{\eta, \beta} \quad \text{for some } \theta \in \Theta, \ \beta \in B \quad \text{vs} \quad \mathcal{H}_1 : \text{not } \mathcal{H}_0. \quad (2.1)
\]
Set
\[
\xi_i = \sigma_{t_i-1}(W_{t_i} - W_{t_{i-1}})/k_{n}^{1/2}
\]
and
\[ \hat{\xi}_i = (p_{t_i} - p_{t_{i-1}})/k_n^{1/2} = \int_{t_{i-1}}^{t_i} \sigma_s dW_s/k_n^{1/2} = \xi_i + \Delta_{ni} \] (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),
\[ E_\theta|\Delta_{ni}|^{2q} = O(k_n^q) \text{ for any } q = 1, 2, \ldots \]
(cf. Lee, 2010).

Concerning the estimation of \( \theta \) and \( \beta \), we assume the following conditions.

(A5) Let \( \psi : \Theta \times R \to R^d \) be a vector-valued function and define \( U_n(\theta) = \sum_{j=1}^n \psi(\theta, \xi_j) \) with \( \psi(\theta, x) = (\psi_1(\theta, x), \ldots, \psi_d(\theta, x))' \). Further, 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 \( i = 1, \ldots, d \),
\[ \sup_t \left\{ \frac{1}{n} \sum_{j=1}^n \left( \frac{\partial \psi_i}{\partial \theta_i}(t, \hat{\xi}_j) - \frac{\partial \psi_i}{\partial \theta_i}(\theta, \hat{\xi}_j) \right) : |t - \theta_i| \leq a_n \} \xrightarrow{P} 0 \]
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
\[ \hat{\theta} = \theta + h(\theta, \hat{r}) + o_p \left( n_k^{-1/2} \right), \] (2.3)
\[ h(\theta, \hat{r}) = \frac{1}{n} \sum_{j=1}^n g(\theta, \hat{\xi}_j) + o_p \left( n_k^{-1/2} \right), \] (2.4)
where \( \hat{r} = (\hat{r}_1, \ldots, \hat{r}_n)' \) and \( n_k = nk_n \).

Further, 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 \( \nu > (2 - \kappa)/(1 - \kappa) > 1 \), and
\[ |g_r(\theta, x_1) - g_r(\theta, x_2)| \leq w_r(x_1, x_2, \theta)|x_1 - x_2| \]
for some real-valued continuous functions \( w_r \geq 0 \) satisfying
\[ \sup_{j,k \in \mathbb{N}} E_\theta \sup_{a \in [-A,A]} \left[ w_r(\xi_j + a, \xi_k, \theta)|^{2\nu} + |w_r(\xi_1 + a, \xi_1, \theta)|^{2\nu} \right] < \infty \]
for any independent copy \( \xi_1 \) of \( \xi_1 \) for some \( A > 0 \): we denote \( W = (w_1, \ldots, w_d) \).
(A6) $\hat{\beta} = \beta + O_p(n^{-1/2})$.

**Remark 2.1** Below, we illustrate that the method of moment estimator satisfies (2.3) and (2.4). For example, consider the Heston model defined in (3.2). $S_L(a^*)$ and $\hat{m}_4$ denote the estimators of integrated volatility and quarticity, respectively, described in Section 3. Set

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

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

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

Then, by (2.5), we have

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

Further, we can readily see that

$$\frac{1}{n} \sum_{j=1}^{n} E\psi(\hat{\theta}, \xi_j) = \left( \frac{1}{nk_n} \int_0^{nk_n} \sigma^2_z ds - \frac{\alpha}{\lambda} \right),$$

and thus,

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

$$y(\theta, \tilde{r}) - \frac{1}{n} \sum_{j=1}^{n} E\psi(\theta, \xi_j) = O_p(n^{-1/4}).$$

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

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

where $A(\theta^*) = [a_{ik}], a_{ik} = \frac{\partial U_i}{\partial \theta_k}(\theta^*)$ and $\theta^*$ lies on the line segment determined by $\hat{\theta}$ and $\theta$, i.e., $|\theta^* - \theta| \leq |\theta - \theta|$. If $A$ is invertible, by plugging (2.9) to (2.6), we have

$$\hat{\theta} = \theta + h(\theta, \tilde{r}) + o_p(n_k^{-1/2}),$$
where \( h(\theta, \hat{r}) = A^{-1}y(\theta, \hat{r}) \) and (2.3) holds. Then, multiplying (2.7) and (2.8) with \( A^{-1} \), we get

\[
\frac{1}{n} \sum_{j=1}^{n} g(\theta, \hat{\xi}_j) - \frac{1}{n} \sum_{j=1}^{n} E g(\theta, \hat{\xi}_j) = O_p(\sqrt{k_n}),
\]

(2.10)

\[
h(\theta, \hat{r}) - \frac{1}{n} \sum_{j=1}^{n} E g(\theta, \hat{\xi}_j) = O_p(n^{-1/4}),
\]

(2.11)

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

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

(2.4)

where we have used the fact that \( O_p(n^{-1/4}) = o_p(n^{-1/2}) \) and \( O_p(\sqrt{k_n}) = o_p(n^{1/2}) \), which is because \( k_n \to 0, nk_n \to \infty \) and \( nk_{2n} \to 0 \) as \( n \to \infty \). Hence, (2.4) holds.

Let

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

be the empirical characteristic function (e.c.f.) based on the observed log returns,

\[
\hat{\phi}_\xi(t) = E_{\theta}(e^{it\xi_1})_{\theta=\hat{\theta}}
\]

be the characteristic function (c.f.) of \( \xi_1 \) with \( \theta \) replaced by its estimator \( \hat{\theta} \), and

\[
\hat{\phi}_\eta(t) = E_{\beta}(e^{it\eta_1})_{\beta=\hat{\beta}}
\]

be the c.f. of \( \eta_1 \) with \( \beta \) replaced by its estimator \( \hat{\beta} \). Further, we assume the conditions:

(A7) \( \hat{\phi}_\xi(t) = E_{\theta}(e^{it\xi_1}) \) is continuously differentiable with respect to \( \theta \) and

\[
\nabla \hat{\phi}_\xi(t) = (\partial \hat{\phi}_\xi(t)/\partial \theta_1, \ldots, \partial \hat{\phi}_\xi(t)/\partial \theta_d)' \]

satisfies

\[
\int \left\| \frac{\partial \hat{\phi}_\xi(t)}{\partial \theta_1} \right\|^{2\nu} dG(t) < \infty.
\]

(A8) (i) The characteristic function of \( \eta \) satisfies

\[
|\phi_\eta(t)| = e^{-o_0(\beta)|t|^{\alpha_1}+R(\beta,t)} \quad \text{as } t \to \infty,
\]
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 $\eta$ 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 2.2 Condition (A8)(i) is related to the supersmoothness case of Fan (1991) which includes the $t$ distribution and the generalized error distribution. Condition (A8)(ii) is related to the ordinary smoothness case of Fan (1991) which includes the exponential distribution and the gamma distribution.

Consider the characteristic function based test statistic:

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

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

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

which is the characteristic function based test statistic for the noiseless case and whose limiting null distribution can be seen in Lin, Lee and Guo (2013). In fact, similarly to Section 3.1 of their paper, we can get

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

where $\hat{g}(\theta, \hat{\xi}_j) = g(\theta, \hat{\xi}_j) - E_{\theta} g(\theta, \hat{\xi}_1)$ 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 the statistic:

$$n_k \int \left| \hat{\phi}_n(t) - E_{\theta} e^{it \hat{\xi}_1} - (\nabla \phi_\xi(t))' \frac{1}{n} \sum_{j=1}^{n} \hat{g}(\theta, \hat{\xi}_j) \right|^2 dG(t),$$
which is also named as $\hat{T}_n^*$ without any confusion. The following 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 2.1** Suppose that (A5) holds. Then, $\hat{T}_n^*$ is a degree-2 degenerate $V$-statistic, that is,

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

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

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

with

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

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

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

Next, we decompose formula (2.13) via using wavelet functions. By (A5) and (2.14), it holds that

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

where $\tilde{F}(x; \theta)$ denotes the stationary distribution of $\hat{\xi}_1$. Let $\Phi$ be a Lipschitz continuous scale function and $\Psi$ 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 by

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

which is an orthonormal basis of the $L_2$-space satisfying

$$\int \Phi_{j,l}(x) \Phi_{j',l'}(x) dx = \begin{cases} 1, & \text{if } j = j' \text{ and } l = l' \\ 0, & \text{otherwise}. \end{cases}$$
Owing to (2.15), the kernel function $k$ has the following 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

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

and

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

Then, we make the two additional assumptions on the scale density of $(\sigma_t^2)$ in (A1) as follows:

(B1) $\int_0^{+\infty} s(v) \left[ \int_0^v f_{\sigma, \theta}(u) du \right]^2 dv < \infty;$

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

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

**Theorem 2.1** Let $\{p_t\}$ and $\{\sigma_t^2\}$ be the ones in (1.1). Suppose that (A1)–(A5), (A7), (B1), and (B2) hold and the distribution function $G$ satisfies

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

Moreover, assume the $\alpha$-mixing coefficients of $\{\sigma_t\}$ satisfy $\alpha_\sigma(m) = O(e^{-am})$ for some $a > 0$. Then, under $H_0$, we have that as $n \to \infty$,

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

(2.17)

where $Z_{j; k}, j = 0, 1, 2, \ldots, k = 0, \pm1, \pm2, \ldots$, are correlated centered normally distributed random variables and $\lambda_{j; k_1, k_2}$ are the wavelet coefficients of the kernel function $\hat{h}(c)$ in (2.14) (cf. (2.16)).

Below, we verify that $\hat{T}_n$ has the same asymptotic distribution as $\hat{T}_n^*$. 
Theorem 2.2 Under the same assumptions in Theorem 2.1 and (A6) and (A8)(i), we have that under $\mathcal{H}_0$,

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

(ii) $\hat{T}_n$ has the same limiting distribution as shown in (2.17).

Theorem 2.3 Under the same assumptions in Theorem 2.1 and (A6) and (A8)(ii), we have that under $\mathcal{H}_0$,

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

(ii) $\hat{T}_n$ has the same limiting distribution as shown in (2.17).

3. Volatility parameter estimation

In this section, we consider two popular SVM examples for illustration. 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 denoted by $m_2$ (a.k.a. the integrated volatility) and $m_4$ (a.k.a. the quarticity), respectively, that is,

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

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 the following SDE:

$$\left\{ \begin{array}{l}
dp_t = \sigma_t dW_t \\
da\sigma_t^2 = -\rho(\sigma_t^2 - \mu)dt + \omega \sqrt{\sigma_t^2} dB_t,
\end{array} \right.$$

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

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

Further, the two moment equations for the parameters $\alpha$ and $\lambda$ are given by

$$m_2 = E\left(\int_0^{nk_n} \sigma_s^2 ds\right) = nk_n \frac{\alpha}{\lambda},$$
$$m_4 = E\left(\int_0^{nk_n} \sigma_s^4 ds\right) = nk_n \left(\frac{\alpha^2}{\lambda^2} + \frac{\alpha^2}{\lambda}\right).$$
Then, the moment estimators of $\alpha$ and $\lambda$ are obtained as

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

**Example 2** (Stein and Stein model)

The process $\{p_t\}$ of the Stein and Stein (1991) model satisfies the following SDE:

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

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

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

Then, the moment estimators of the parameters $\mu$ and $\tau^2$ are obtained from the two moment equations:

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

Thus, the moment estimators of $\mu$ and $\tau$ are obtained as

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

In this case, however, $\{\bar{r}_i\}$ are unobservable, and further, the observed returns $\{	ilde{r}_i\}$ are contaminated by microstructure noises. 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 on p.2256) to obtain

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

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

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

In the literature, various methods are proposed to estimate \( \int_0^1 \sigma_s^2 ds \), known as the integrated volatility, in the fixed-span in-fill setting \( (k_n \to 0 \text{ and } nk_n \to \text{constant}) \). For example, we can mention the two-scaled estimator of Zhang et al. (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 better than state-of-the-art methods in the estimation of integrated volatility.

Note that the quadratic estimator of Lin and Guo (2015) takes the following form:

\[
S_L(a) = a_0 \sum_{j=1}^{n+1} \tilde{r}_j^2 + a_1 \sum_{j=1}^{n} \tilde{r}_j \tilde{r}_{j+1} + \ldots + a_\ell \sum_{j=1}^{n+1-\ell} \tilde{r}_j \tilde{r}_{j+\ell},
\]

where \( L_h = \sum \tilde{r}_j \tilde{r}_{j+h} \) denotes the lag \( h \) sample autocovariance. The optimal weights \( a = (a_0, a_1, \ldots, a_\ell) \) are chosen to satisfy the unbiasedness and minimum variance conditions of the estimator. 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, \( a^* = (a^*_2, \ldots, a^*_\ell) \), are chosen to minimize the finite sample variance and to satisfy the following system of equations (for details, see (16) of Lin and Guo, 2015):

\[
\begin{align*}
\mu_2 a_2 + \rho_3 a_3 + \gamma_4 a_4 &= -\gamma_2 - 2\rho_2 \\
\rho_3 a_2 + \mu_3 a_3 + \rho_4 a_4 + \gamma_5 a_5 &= -2\gamma_3 \\
\gamma_{h+2} a_h + \rho_{h+2} a_{h+1} + \mu_{h+2} a_{h+2} + \rho_{h+3} a_{h+3} + \gamma_{h+4} a_{h+4} &= 0
\end{align*}
\]

for \( 2 \leq h \leq \ell - 2 \), where \( \mu_h = E[\text{Var}_G(L_h)] \), \( \rho_h = E[\text{Cov}_G(L_{h-1}, L_h)] \) and \( \gamma_h = E[\text{Cov}_G(L_{h-2}, L_h)] \), and \( G \) is the \( \sigma \)-field generated by \( \{\sigma_t, \ t \geq 0\} \). To apply the quadratic estimator of Lin and Guo (2015) to our setting \( (nk_n \to \infty) \), we extend Lemma 2 of Lin and Guo (2015) as follows:

**Lemma 3.2** Assume that the log price \( \tilde{p}_t \)'s are observed at the 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
\[ n \to \infty. \text{ Let } K = E(\eta^4_n)/(E\eta^2_n)^2. \text{ Then,} \]
\[
\begin{align*}
\mu_0 &= 2k_n E \left( \int_0^{nk_n} \sigma_s^4 ds \right) + 4\sigma^2 E \left( \int_0^{nk_n} \sigma_s^2 ds \right) + (nK - 1)\sigma^4 + o(nk_n^3), \\
\mu_1 &= k_n E \left( \int_0^{nk_n} \sigma_s^4 ds \right) + 2\sigma^2 E \left( \int_0^{nk_n} \sigma_s^2 ds \right) + 2A_1 + \frac{(K + 4)n - 6}{4} \sigma^4 + o(nk_n^3), \\
\rho_1 &= -2\sigma^2 E \left( \int_0^{nk_n} \sigma_s^2 ds \right) - 2A_1 - \frac{(K + 1)n - 2}{2} \sigma^4, \\
\mu_h &= k_n E \left( \int_0^{nk_n} \sigma_s^4 ds \right) + 2\sigma^2 E \left( \int_0^{nk_n} \sigma_s^2 ds \right) + A_h + B_h + \frac{3n - 3h}{2} \sigma^4 + o(nk_n^3), \\
\rho_h &= -\sigma^2 E \left( \int_0^{nk_n} \sigma_s^2 ds \right) - \frac{1}{2} A_h - \frac{1}{2} B_h - \frac{2n - 2h + 1}{2} \sigma^4, \quad 2 \leq h \leq \ell
\end{align*}
\]

where \( A_h = \sigma^2 E \left( \int_{(n-h)K}^{nk_n} \sigma_s^2 ds \right) \) and \( B_h = \sigma^2 E \left( \int_0^{hK} \sigma_s^2 ds \right) \), \( 1 \leq h \leq \ell \).

The proof of Lemma 3.2 is given in the Appendix. The results for \( \gamma_h \)'s are actually the same as those in Lemma 2 of Lin and Guo (2015).

By dividing both sides of (3.4) by \( n\sigma^4 \) and ignoring the \( O(\ell k_n/n) \) terms, we can obtain the new system of equations:

\[
\begin{align*}
\dot{\mu}_2a_2 + \dot{\rho}_3a_3 + \dot{\gamma}_4a_4 + \dot{\gamma}_2 + 2\dot{\rho}_2 &= 0 \\
\dot{\rho}_3a_2 + \dot{\mu}_3a_3 + \dot{\rho}_4a_4 + \dot{\gamma}_5a_5 + 2\dot{\gamma}_3 &= 0 \\
\dot{\gamma}_{h+2}a_h + \dot{\rho}_{h+2}a_{h+1} + \dot{\mu}_{h+2}a_{h+2} + \dot{\rho}_{h+3}a_{h+3} + \dot{\gamma}_{h+4}a_{h+4} + \dot{\gamma}_h &= 0
\end{align*}
\]

(3.5)

where \( 2 \leq h \leq \ell - 2 \) and

\[
\begin{align*}
\dot{\mu}_h &= S_{nr}^2 + 2S_{nr} + \frac{3n - 3h}{2n}, & \dot{\rho}_h &= -S_{nr} - \frac{2n - 2h + 1}{2n}, \\
\dot{\gamma}_2 &= \frac{n - 1}{2n}, & \dot{\gamma}_h &= \frac{n - h + 1}{4n}, & S_{nr} &= E \int_0^{nk_n} \sigma_s^2 ds/n\sigma^2.
\end{align*}
\]

for \( 2 \leq h \leq \ell \), and then, use (3.5) to solve the optimal weights \( \mathbf{a}^* \). Note that the numerator of \( S_{nr} \) is slightly modified from that of Lin and Guo (2015) for our setting and the new 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, who employ the Newton-Raphson and Gauss-Seidel methods to solve \( \mathbf{a}^* \) and derive the asymptotic distribution of \( S_L(\mathbf{a}^*) \) via following an approach similar to Lin and Guo (2015), Theorem 1.

4. Noise parameter estimation
In this section, we discuss on the parameter estimators of the noise distribution by using the method of moment estimator. By the rule of thumb, we suggest to use the lag-1 sample autocovariance and the fourth moment of the observed returns to obtain the moment estimators. That is, we employ the following two moment equations:

\[
E\left(\frac{1}{n} \sum_{j=1}^{n} \tilde{r}_j \tilde{r}_{j+1}\right) = -E(\eta_t^2),
\]

\[
E\left(\frac{1}{n} \sum_{j=1}^{n} \tilde{r}_j^4\right) = 3k_s^2m_4 + 12k_nm_2E(\eta_t^2) + 2E(\eta_t^4) + 6(E(\eta_t^2))^2,
\]

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

For example, if \(\eta_t\) is assumed to follow a scald \(t\) distribution, the moment estimators of \(s\) and \(\nu\) can be solved from the two moment equations:

\[
\frac{1}{n} \sum_{j=1}^{n} \tilde{r}_j \tilde{r}_{j+1} = \frac{-s^2 \nu}{\nu - 2},
\]

\[
\frac{1}{n} \sum_{j=1}^{n} \tilde{r}_j^4 = 3k_s^2m_4 + 12k_nm_2 s^2 \nu + 12 \frac{s^4 \nu^2 (\nu - 3)}{\nu - 4)(\nu - 2)^2},
\]

For another example, if \(\eta_t \sim \mathcal{E}(\beta) - \beta^{-1}\), where \(\mathcal{E}(\beta)\) denotes the exponential distribution with expected value \(\beta^{-1}\), \(\varepsilon_t = \eta_t - \eta_{t-1}\) follows a double exponential distribution with parameter \(\beta\). Hence, the moment estimator of \(\beta\) can be solved from the equation:

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

Below, we investigate the accuracy of the moment estimators for both the volatility and noise distributions. Recall that \(S_{nr}\) is the ratio of the integrated volatility to the \(n\)-folds variance of microstructure noise, that is,

\[
S_{nr} = \frac{E\left(\int_0^{nk\tau} \sigma_s^2 ds\right)}{2n\sigma_{\eta}^2},
\]

named the signal-to-noise ratio. For the Heston model in (3.2), the parameters are set to be \(\rho = 10\), \(\mu = 0.00048\) and \(\omega = \sqrt{\rho \mu}/2\). For the Stein and Stein
In these settings, the numerator of $S_{nr}$ is $4.8 \times 10^{-4}$. The sample size is set to be $n = 2 \times 10^5$. To investigate the performance of the moment estimators, we conduct a simulation study for three different values of $S_{nr}$. The parameters of noise distribution are $\nu = 5$, $s = 0.00035$ ($S_{nr} = 0.12$), $\nu = 8$, $s = 0.0002$ ($S_{nr} = 0.45$) and $\nu = 10$, $s = 0.0001$ ($S_{nr} = 1.9$) for the scaled $t$ distribution, and further, $\beta = 2200$ ($S_{nr} = 0.12$), $\beta = 4300$ ($S_{nr} = 0.45$) and $\beta = 9000$ ($S_{nr} = 1.9$) for the exponential distribution.

The relative errors of $(\hat{\alpha}, \hat{\lambda})$ and $(\hat{\mu}, \hat{\tau})$ are listed in Table 4.1. The result demonstrates 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 benchmark relative errors, we consider the moment estimators for noise free SVMs, that is, $\epsilon_i = 0$ for all $i$. In the same parameter settings as above, 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 4.1, which are served 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 $(\hat{\mu}_o, \hat{\tau}_o)$.

Table 4.1: Relative errors of parameter estimation

<table>
<thead>
<tr>
<th></th>
<th>Heston Model</th>
<th></th>
<th>Stein and Stein Model</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_{nr} = 0.12$</td>
<td>$S_{nr} = 0.45$</td>
<td>$S_{nr} = 1.9$</td>
<td>$S_{nr} = 0.12$</td>
</tr>
<tr>
<td>$\hat{\alpha}_o$</td>
<td>0.1466</td>
<td>0.1537</td>
<td>0.1524</td>
<td>$\hat{\mu}_o$</td>
</tr>
<tr>
<td>$\hat{\lambda}_o$</td>
<td>0.1571</td>
<td>0.1637</td>
<td>0.1641</td>
<td>$\hat{\tau}_o$</td>
</tr>
<tr>
<td>$\hat{\alpha}$</td>
<td>0.2247</td>
<td>0.1584</td>
<td>0.1569</td>
<td>$\hat{\mu}$</td>
</tr>
<tr>
<td>$\hat{\lambda}$</td>
<td>0.2265</td>
<td>0.1656</td>
<td>0.1659</td>
<td>$\hat{\tau}$</td>
</tr>
<tr>
<td>$\hat{\nu}$</td>
<td>0.3586</td>
<td>0.2409</td>
<td>0.4141</td>
<td>$\hat{\nu}$</td>
</tr>
<tr>
<td>$\hat{s}$</td>
<td>0.1534</td>
<td>0.1249</td>
<td>0.3958</td>
<td>$\hat{s}$</td>
</tr>
<tr>
<td>$\hat{\beta}$</td>
<td>0.0304</td>
<td>0.1074</td>
<td>0.2457</td>
<td>$\hat{\beta}$</td>
</tr>
</tbody>
</table>
5. Simulation and Empirical Results

5.1. Simulation Results

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

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

(5.1)

with

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

$$
\phi_{\theta_n}(t) = \mathbb{E}(e^{it\xi_1}) \mathbb{E}(e^{it\eta_1/k_n^{1/2}}) \mathbb{E}(e^{-it\eta_1/k_n^{1/2}}) |_{\theta=\hat{\theta}, \beta=\hat{\beta}},
$$

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

We consider the following null hypotheses $H_{0,ij}$: $\sigma_t \sim f_i$ and $\eta_t \sim f_\eta^{(j)}$ for $i = 1, 2$ and $j = 1, 2$, where $f_1 \sim \sqrt{\text{Gamma}}$ (the Heston SVM, see (3.2)), $f_2 \sim \text{Normal}$ (the Stein and Stein SVM, see (3.3)), $f_\eta^{(1)} \sim \text{st}$ and $f_\eta^{(2)} \sim \mathcal{E}(\beta) - \beta^{-1}$.

In the next section, we will discuss the choice of microstructure noise distribution based on empirical data. Under $H_{0,i1}$, we have

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

where $B_{\nu/2}$ denotes a modified Bessel function of the third kind with index $\nu/2$.

Under $H_{0,i2}$, we have

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

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

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

Our simulation scheme is similar to that of Lin, Lee and Guo (2013) and the key steps are as follows:
1. Simulate a sample \( \{ p_i \}_{1 \leq i \leq n} \) from a hypothesized SVM and \( \{ \eta_i \}_{1 \leq i \leq n} \) with the corresponding true parameters \( \theta \) 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 \( H_{0,1} \) and (4.2) and (3.4) for \( H_{0,2} \). 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 1000 times to obtain the sizes and powers.

The parameter settings are 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 statistic \( \hat{T}_n \) for \( H_{0,ij}, i = 1, 2, j = 1, 2 \), versus five \( H_1 \) (corresponding to the five volatility distributions: \( \sqrt{\text{Gamma}}, \text{Normal}, \text{Uniform}, \text{F} \) and Inverse Gamma (IG)), are presented in Table 5.2 \( (H_{0,1i}, i = 1, 2) \) and Table 5.3 \( (H_{0,2i}, i = 1, 2) \). The obtained results support the validity of our test.

Table 5.2: The sizes and powers (in percentage) of \( \hat{T}_n \) for \( H_{0,1i}, i = 1, 2 \) versus five \( H_1 \).

<table>
<thead>
<tr>
<th>( H_1 )</th>
<th>( H_{0,11} : \sqrt{\text{Gamma}} )</th>
<th>( H_{0,21} : \text{Normal} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \sqrt{\text{Gamma}} )</td>
<td>( S_{nr} = 0.12 )</td>
<td>5.2</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 0.45 )</td>
<td>5.3</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 1.9 )</td>
<td>4.6</td>
</tr>
<tr>
<td>Normal</td>
<td>( S_{nr} = 0.12 )</td>
<td>48.1</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 0.45 )</td>
<td>35.5</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 1.9 )</td>
<td>76.2</td>
</tr>
<tr>
<td>Uniform</td>
<td>( S_{nr} = 0.12 )</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 0.45 )</td>
<td>99.9</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 1.9 )</td>
<td>100.0</td>
</tr>
<tr>
<td>F</td>
<td>( S_{nr} = 0.12 )</td>
<td>82.5</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 0.45 )</td>
<td>70.8</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 1.9 )</td>
<td>90.2</td>
</tr>
<tr>
<td>IG</td>
<td>( S_{nr} = 0.12 )</td>
<td>33.2</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 0.45 )</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>( S_{nr} = 1.9 )</td>
<td>93</td>
</tr>
</tbody>
</table>
Table 5.3: The sizes and powers (in percentage) of $\hat{H}_n$ for $H_{0,i2}$, $i = 1, 2$ versus five $H_1$.

<table>
<thead>
<tr>
<th>$H_1$</th>
<th>$H_{0,12} : \sqrt{\text{Gamma}}$</th>
<th>$H_{0,22} : \text{Normal}$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$S_{nr} = 0.12$ $S_{nr} = 0.45$ $S_{nr} = 1.9$</td>
<td>$S_{nr} = 0.12$ $S_{nr} = 0.45$ $S_{nr} = 1.9$</td>
</tr>
<tr>
<td>$\sqrt{\text{Gamma}}$</td>
<td>3.6</td>
<td>4.7</td>
</tr>
<tr>
<td>Normal</td>
<td>82.3</td>
<td>93.8</td>
</tr>
<tr>
<td>Uniform</td>
<td>95.9</td>
<td>88.9</td>
</tr>
<tr>
<td>F</td>
<td>70.9</td>
<td>99.8</td>
</tr>
<tr>
<td>IG</td>
<td>68.7</td>
<td>82.7</td>
</tr>
</tbody>
</table>

5.2. Real Data Analysis

To showcase an empirical application, we consider the ultra high frequency tick-by-tick data of 13 stocks listed on the NYSE (New York Stock Exchange): 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 $\tilde{p}_t$’s for each stock. To preprocess the suspicious jumps, we apply the wavelet method of Fan and Wang (2007). The following analysis are based on the log returns after the jumps are smoothed.

Before performing the hypothesis testing, we discuss on appropriate microstructure noise distributions through a real high frequency data analysis. To this task, we consider the three stocks with different transaction frequencies: ABT (low frequency), GE (middle frequency) and JPM (high frequency). Recall that 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,

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

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

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

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 $\eta_t$’s are i.i.d., the distribution of $\varepsilon_t$ is can be obtained as the convolution of the density function of $\eta_t$. As for the candidate distributions of $\{\eta_t\}$, below we consider a scaled $t$ distribution, a normal distribution, a generalized error distribution, and an exponential distribution.

First, we consider the JPM case. In Fig. 5.1 (left panel), we plot the empirical characteristic function (in red) of the JPM and the fitted characteristic functions of $\varepsilon_t = \eta_t - \eta_{t-1}$ for $\eta_t$ following a normal (in yellow), a generalized error (in green), an exponential (in blue), and a scaled $t$ (in purple) distributions, respectively. The corresponding log empirical/fitted characteristic functions versus log($t$) are plotted in the right panel of Fig. 5.1. The parameters of these fitted characteristic functions are 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 observations, we select the scaled $t$ distribution as our candidate distribution for $\eta_t$.

Figure 5.1: The left panel includes the empirical characteristic function (red) of the JPM and the fitted characteristic functions of $\varepsilon_t$ for $\eta_t$ following normal (yellow), generalized error (green), scaled log uniform (blue) and scaled $t$ (purple) distributions. The right panel includes the corresponding log characteristic functions.

To investigate the microstructure noise effect on model testing, both the 2-minute and 5-minute returns are considered. We utilize the high frequency transaction data of the 21 trading days in the period 2002/01/02~2002/01/31. As with the setting of Lin, Lee and Guo (2013), we regard one hour as a time unit and overnight returns are ignored. For the 5-minute returns, we set $k_n = 7n^{-0.6}$. 
that is, the actual sampling time length is $60 \min \times k_n \approx 5 \min$. The normal trading hours of NYSE is 6.5 hours from 9:30 to 16:00, and thus, for each stock in the 21 trading days, we have sample size $n = 1638$. For the 2-minute returns, we set $k_n = 5n^{-0.6}$, that is, the actual sampling time length is $60 \min \times k_n \approx 2 \min$ and we have sample size $n = 4095$ for each stock in the 21 trading days. We utilize high frequency returns to test the two null hypotheses $H_{0,i}$: $\sigma_t \sim f_i$ and $\eta_t \sim$ scaled $t$, $i = 1, 2$, where $f_1 \sim \sqrt{\text{Gamma}}$ (Heston model) and $f_2 \sim \text{Normal}$ (Stein and Stein model) for each stock.

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

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

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

We choose 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 mean time, 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) $\tilde{T}_n$ accepts $H_{0,11}$ and rejects $H_{0,21}$ 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 $H_{0,11}$ and $H_{0,21}$ 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, $\hat{T}_n$ accepts $H_{0,11}$ and rejects $H_{0,21}$ in the 5-minute return case, while it rejects both the hypotheses in 2-minute return case. For KO and NOK, though, $\hat{T}_n$ accepts both the $H_{0,11}$ and $H_{0,21}$ when 5-minute returns are used, while it rejects $H_{0,21}$ when 2-minute returns are used.

The summary in (ii) indicates that $\hat{T}_n$ yields more consistent results, reflects real situations 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 consider the following null hypothesis: $H_{0,12}: \sigma_t \sim \sqrt{\text{Gamma}}$ (Heston model) and $\eta_t \sim E(\beta) - \beta^{-1}$. Compared $H_{0,12}$ with $H_{0,11}$, we kept the same distribution assumption on $\sigma_t$ yet change the one on $\eta_t$. The test $\hat{T}_n$ rejects $H_{0,12}$ for all 13 stocks for the 2-minute return case. This result indicates that the proposed $\hat{T}_n$ has power in testing the the distribution assumption on $\eta_t$ and the scaled $t$ distribution is preferable to $E(\beta) - \beta^{-1}$ 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 is conducted to evaluate the proposed test. Our real 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.

Acknowledgment 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.

References


Liang-Ching Lin
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

Sangyeol Lee
Department of Statistics, Seoul National University, Seoul, 151-742, South Korea.
E-mail: sylee@states.snu.ac.kr

Meihui Guo
Department of Applied Mathematics, National Sun Yat-sen University, Kaohsiung, Taiwan, 804, R.O.C.
E-mail: guomh@math.nsyesu.edu.tw
Appendix: Proofs

The proof of Theorem 2.1

Proof. Since (2.12) is reexpressed as

\[
\begin{align*}
\hat{T}_n &= nk_n \int \left| \frac{\hat{\phi}_n(t) - \hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2})} \right|^2 dG(t) \\
&= nk_n \int \left| \frac{1}{n} \sum_{j=1}^{n} e^{i \tilde{t} j / k_n^{1/2}} - \frac{1}{n} \sum_{j=1}^{n} e^{i \xi_j} \hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2})} \right|^2 dG(t) \\
&= nk_n \int \left| \frac{1}{n} \sum_{j=1}^{n} e^{i \xi_j} \left[ e^{i (\eta_j - \eta_{j-1}) / k_n^{1/2}} - \hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2}) \right] \right|^2 dG(t) \\
&\leq 2nk_n \int \left| \frac{1}{n} \sum_{j=1}^{n} e^{i \xi_j} \left[ e^{i (\eta_j - \eta_{j-1}) / k_n^{1/2}} - \hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2}) \right] \right|^2 dG(t) \\
&\quad + 2nk_n \int \left| \frac{\hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2})}{\hat{\phi}_\eta(t/k_n^{1/2}) \hat{\phi}_\eta(-t/k_n^{1/2})} - 1 \right|^2 dG(t) \\
&= 2 \times (J1) + 2 \times (J2). \tag{7.1}
\end{align*}
\]

Below, 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

\[
|\nabla_\beta \log \phi_\eta(t/\sqrt{k_n})| = -\nabla_\beta \alpha_0(\beta) \left| t/\sqrt{k_n} \right|^{\alpha_1} + \nabla_\beta R(\beta, t/\sqrt{k_n}).
\]
This together with (A6) implies

\[ |h(t)| = - (\nabla_\beta a_0(\beta))^T \left| \frac{t}{\sqrt{k_n}} \right|^{\alpha_1} (\hat{\beta} - \beta) + \left( \nabla_\beta R(\beta, t/\sqrt{k_n}) \right)^T (\hat{\beta} - \beta) \]

\[ = \mathcal{O}_p \left( |t|^{\alpha_1} k_n^{-\alpha_1/2} n^{-1/2} \right) + \mathcal{O}_p \left( |t|^{\alpha_1} k_n^{-\alpha_1/2} n^{-1/2} \right) \]

\[ = \mathcal{O}_p \left( |t|^{\alpha_1} (nk_n^{-1})^{-1/2} \right). \tag{7.2} \]

Subsequently, since \( \alpha_1 < 1 \),

\[ (J2) = nk_n \int \mathcal{O}_p \left( |t|^{2\alpha_1} (nk_n^{-1})^{-1} \right) dG(t) = \mathcal{O}_p \left( k_n^{1-\alpha_1} \int |t|^{2\alpha_1} dG(t) \right) = o_p(1). \]

Meanwhile, notice that

\[ (J1) = nk_n \int \left| \sum_{j=1}^{\infty} e^{i\xi_j} \left[ \frac{e^{it(\eta_j-\eta_{j-1})/\sqrt{k_n}}/(\phi_{\eta}(t/\sqrt{k_n})\phi_{\eta}(-t/\sqrt{k_n})) - 1}{\phi_{\eta}(t/\sqrt{k_n})\phi_{\eta}(-t/\sqrt{k_n})} \right]^2 \right| dG(t) \]

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

By (7.2), we have \( h(t) = \mathcal{O}_p \left( |t|^{\alpha_1} (nk_n^{-1})^{-1/2} \right) = o_p(1) \) since \( nk_n^{\alpha_1} \to \infty \) as \( n \to \infty \). Therefore,

\[ (J1) = nk_n \int \left| \sum_{j=1}^{\infty} Y_j \right|^2 dG(t), \]

where

\[ Y_j = e^{i\xi_j} \left[ \frac{e^{it(\eta_j-\eta_{j-1})/\sqrt{k_n}}}{\phi_{\eta}(t/\sqrt{k_n})\phi_{\eta}(-t/\sqrt{k_n})} - 1 \right]. \]

Below, we derive the mean and the variance of \( \bar{Y} = \sum_{j=1}^{\infty} Y_j \). Since \( \{\eta_t\} \) is a white noise process and independent of \( \{p_t\} \), it is immediate to see that

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

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

\[ \text{Var}(\bar{Y}) = E \left[ \left( \frac{1}{n} \sum_{j=1}^{n} Y_j \right) \left( \frac{1}{n} \sum_{j=1}^{n} Y_j^* \right) \right] \]

\[ = \frac{1}{n^2} \sum_{j=1}^{n} E(Y_j Y_j^*) + \frac{1}{n^2} \sum_{j=1}^{n-1} E(Y_j Y_{j+1}^*) + \frac{1}{n^2} \sum_{j=1}^{n-1} E(Y_{j+1} Y_j^*) \]

\[ = (J_1 - 1) + (J_1 - 2) + (J_1 - 3). \]  

(7.3)

By simple algebra, we can check that

\[ (J_1 - 1) = \frac{1}{n} \left[ \frac{1}{\phi_\eta^2(t/\sqrt{k_n})\phi_\eta^2(-t/\sqrt{k_n})} - 1 \right]. \]  

(7.4)

Note that owing to (A8)(i),

\[ |\phi_\eta^2(t/\sqrt{k_n})\phi_\eta^2(-t/\sqrt{k_n})| = e^{-4\alpha_0(\beta)|t|^\alpha k_n^{-\alpha/2} + R'(\beta, t/\sqrt{k_n})}, \]

where \( R'(\beta, t/\sqrt{k_n}) = 2R(\beta, t/\sqrt{k_n}) + 2R(\beta, -t/\sqrt{k_n}) \) has an order of \( o(|t/\sqrt{k_n}|^\alpha). \)

Thus, by (7.4), we have

\[ |(J_1 - 1)| = \left| \frac{1}{n} \left[ e^{4\alpha_0(\beta)|t|^\alpha k_n^{-\alpha/2} - R'(\beta, t/\sqrt{k_n})} - 1 \right] \right| \]

\[ \leq \left| \frac{1}{n} \left( 4\alpha_0(\beta)|t|^\alpha k_n^{-\alpha/2} - R'(\beta, t/\sqrt{k_n}) \right) \right| \]  

(7.5)

\[ = O \left( |t|^\alpha nk_n^{\alpha/2} \right)^{-1}, \]  

(7.6)

where the inequality in (7.5) holds due to the fact that \( |e^x - 1| \leq |x| \) \( \forall x. \)

Similarly, we can have

\[ |(J_1 - 2)| = \left| \frac{1}{n^2} \sum_{j=1}^{n-1} E \left[ e^{it(\xi_j + 1 - \xi_j)} \right] \left( \frac{\phi_\eta(-2t/\sqrt{k_n})}{\phi_\eta(t/\sqrt{k_n})\phi_\eta(-t/\sqrt{k_n})} - 1 \right) \right| \]

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

\[ = O \left( |t|^\alpha nk_n^{\alpha/2} \right)^{-1} \]  

(7.7)

and

\[ |(J_1 - 3)| = O \left( |t|^\alpha nk_n^{\alpha/2} \right)^{-1}. \]  

(7.8)
Then, combining (7.3), (7.6), (7.7) and (7.8), we have

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

This validates the theorem. □

The proof of Theorem 2.3

Proof. Basically, we follow the same lines in the proof of Theorem 2.2, so highlight some key steps. First, we show that \((J1)\) and \((J2)\) defined in (7.1) 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

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

This together with (A6) implies

\[ |h(t)| = O_p \left( (-\log |t| + 2^{-1} \log k_n)n^{-1/2} \right) + O_p \left( |t|^{-\alpha_1} k_n^{\alpha_3/2} n^{-1/2} \right) \]

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

Therefore,

\[ (J2) = nk_n \int O_p \left( (\log |t| + \log k_n) \right) dG(t) \]

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

To show that \((J1) = o_p(1)\), we deduce the orders of the three terms in (7.3).

By (A8)(ii), we have

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

Then, using (7.4) and the fact that \(|e^x - 1| \leq |x| \forall x\), we get

\[ |(J1-1)| = \left| \frac{1}{n} O \left( |t|^{2\alpha_1(\beta)} k_n^{-\alpha_1(\beta)} - 1 \right) \right| \]

\[ = \left| \frac{1}{n} O \left( e^{2\alpha_1(\beta)} \log \log k_n - 1 \right) \right| \]

\[ = O \left( n^{-1} \log |t| + n^{-1} \log k_n \right) \]

Similarly, it can be seen that owing to (A8)(ii), (7.7) and (7.8),

\[ |(J1-2)| = O \left( n^{-1} \log |t| + n^{-1} \log k_n \right) \]
and
\[ |(J1 - 3)| = O\left(n^{-1} \log |t| + n^{-1} \log k_n\right). \]

Therefore, we get
\[ |(J1)| \leq nk_n \int O_p\left(n^{-1} \log |t| + n^{-1} \log k_n\right) dG(t) \]
\[ = O_p\left(k_n \int \log |t| dG(t) + k_n \log k_n\right) = o_p(1), \]

which asserts the theorem. □

The proof of Lemma 3.2

Proof. The second and fourth moments of the nominal return \( r_j \) are
\[ \sum_{j=1}^{n+1} E\left(r_j^2\right) = \sum_{j=1}^{n+1} E\left(\int_{(j-1)k_n}^{jk_n} \sigma_s dW_s\right)^2 = \sum_{j=1}^{n} E\left(\int_{(j-1)k_n}^{jk_n} \sigma_s^2 ds\right) \]
\[ = E\left(\int_0^{nk_n} \sigma_s^2 ds\right) \]

and
\[ E\left(\sum_{j=1}^{n+1} r_j^4\right) = k_n^2 \sum_{j=1}^{n+1} E\left(\frac{r_j}{k_n^{1/2}}\right)^4 = k_n^2 \sum_{j=1}^{n+1} E\left(\xi_j^4 + 6 \xi_j^2 \Delta_{nj}^2 + \Delta_{nj}^4\right) \]
\[ = 3k_n^2 \sum_{j=1}^{n+1} E\left(\sigma_{t_j-1}^4\right) + 6k_n \sum_{j=1}^{n+1} E\left(\sigma_{t_j-1}^2\right) E(\Delta_{nj}^2) + k_n \sum_{j=1}^{n+1} E(\Delta_{nj}^4) \]
\[ = 3k_n E\left(\int_0^{nk_n} \sigma_s^4 ds\right) + O(nk_n^3). \]

The remaining part of the proof follows essentially the same lines as in the proof of Lemma 2 of Lin and Guo (2015). □