Supplementary Materials of "Information Ratio Test for Model Misspecification on Parametric Structures in Stochastic Diffusion Models"

This document is prepared to provide the detailed and complete proof of the lemma in section 2.2. The proof of the lemma is sketched in the appendix due to the space limitations. The main theorem follows immediately from the lemma.

**Proposition 1.** If conditions (a) and (b) in the appendix are satisfied, then

\[
\hat{S}(\hat{\theta}_n, \hat{\gamma}) \overset{pr}{\to} S(\theta_0, \gamma_0), \quad n \to \infty,
\]

\[
\hat{V}(\hat{\theta}_n, \hat{\gamma}) \overset{pr}{\to} V(\theta_0, \gamma_0), \quad n \to \infty.
\]

**Proof of Proposition 1.** Rewrite \(\hat{S}(\theta, \gamma)\) and \(\hat{V}(\theta, \gamma)\) as follows:

\[
\hat{S}(\theta, \gamma) \triangleq \frac{1}{n} \sum_{k=1}^{n} g_s(X_{(k-1)\Delta}, X_{k\Delta}; \theta, \gamma),
\]

\[
\hat{V}(\theta, \gamma) \triangleq \frac{1}{n} \sum_{k=1}^{n} g_v(X_{(k-1)\Delta}, X_{k\Delta}; \theta, \gamma).
\]

By conditions (a) and (b), we have \(\hat{\theta} \overset{pr}{\to} \theta_0\) and \(\hat{\gamma} \overset{pr}{\to} \gamma_0\). Applying the uniform law of large number (Theorem 4.1 in Wooldridge (1994)), the conclusions of proposition 1 are proved.

Furthermore, note that under the null hypothesis of correct model specification, \(S(\theta_0, \gamma_0) = V(\theta_0, \gamma_0)\). Thus, by condition (d) and Slutsky’s theorem, we have

\[
tr \left( S^{-1}(\hat{\theta}, \hat{\gamma})V(\hat{\theta}, \hat{\gamma}) \right) \overset{pr}{\to} p, \quad n \to \infty.
\]

Now, we provide a detailed and complete proof the lemma.
Proof of Lemma. In this proof, we apply the uniform law of large number theorem multiple times, and readers are refer to Theorem 4.1 of Wooldridge (1994) for the detail.

We consider the case that the null hypothesis $H_0$ is true. Given that $\hat{S}(\hat{\theta}, \hat{\gamma})$ converges to the positive definite sensitivity matrix $S(\theta_0, \gamma_0)$ in probability, we can obtain the following exact expansion:

$$
\hat{S}^{-1}(\hat{\theta}, \hat{\gamma}) = S^{-1}(\theta_0, \gamma_0) + \hat{S}^{-1}(\hat{\theta}, \hat{\gamma}) - S^{-1}(\theta_0, \gamma_0)
$$

$$
= S^{-1}(\theta_0, \gamma_0) + S^{-1}(\theta_0, \gamma_0) \left(S(\theta_0, \gamma_0) - \hat{S}(\hat{\theta}, \hat{\gamma})\right) \hat{S}^{-1}(\hat{\theta}, \hat{\gamma})
$$

$$
= S^{-1}(\theta_0, \gamma_0) + S^{-1}(\theta_0, \gamma_0) \left(S(\theta_0, \gamma_0) - \hat{S}(\hat{\theta}, \hat{\gamma})\right) S^{-1}(\theta_0, \gamma_0)
$$

$$
+ \{S^{-1}(\theta_0, \gamma_0)(S(\theta_0, \gamma_0) - \hat{S}(\hat{\theta}, \hat{\gamma}))\}^2 S^{-1}(\theta, \gamma)
$$

It follows that the information ratio statistic $R_n$ can be represented as follows:

$$
\sqrt{n} \left\{ R_n - p \right\} = \sqrt{n} \frac{\text{tr}}{\text{tr}} \left\{ \hat{S}^{-1}(\hat{\theta}, \hat{\gamma}) V(\hat{\theta}, \hat{\gamma}) - I_p \right\}
$$

$$
= \sqrt{n} \frac{\text{tr}}{\text{tr}} \left\{ \hat{S}^{-1}(\hat{\theta}, \hat{\gamma}) V(\hat{\theta}, \hat{\gamma}) - S^{-1}(\theta_0, \gamma_0) V(\theta_0, \gamma_0) \right\}
$$

$$
= \text{tr} \left\{ S^{-1}(\theta_0, \gamma_0) \sqrt{n} \left( V(\hat{\theta}, \hat{\gamma}) - V(\theta_0, \gamma_0) \right) \right\}
$$

$$
+ \text{tr} \left\{ \hat{S}^{-1}(\hat{\theta}, \hat{\gamma}) V(\hat{\theta}, \hat{\gamma}) S^{-1}(\theta_0, \gamma_0) \sqrt{n} \left( S(\theta_0, \gamma_0) - \hat{S}(\hat{\theta}, \hat{\gamma}) \right) \right\}
$$

$$
+ \text{tr} \left\{ \hat{S}^{-1}(\hat{\theta}, \hat{\gamma}) V(\hat{\theta}, \hat{\gamma}) S^{-2}(\theta_0, \gamma_0) \sqrt{n} \left( S(\theta_0, \gamma_0) - \hat{S}(\hat{\theta}, \hat{\gamma}) \right) \right\}.
$$

For $1 \leq i, j \leq p$, expanding the $(i, j)$-element of $\hat{S}(\hat{\theta}, \hat{\gamma})$ around $(\theta_0, \gamma_0)$, we obtain

$$
\sqrt{n} \left( \hat{S}(\hat{\theta}, \hat{\gamma})_{ij} - S(\theta_0, \gamma_0)_{ij} \right)
$$

$$
= \sqrt{n}(\hat{S}(\theta_0, \gamma_0)_{ij} - S(\theta_0, \gamma_0)_{ij}) + \frac{\partial}{\partial \theta_i} \hat{S}(\theta_0, \gamma_0)_{ij} \sqrt{n}(\hat{\theta} - \theta_0) + \frac{\partial}{\partial \gamma_j} \hat{S}(\theta_0, \gamma_0)_{ij} \sqrt{n}(\hat{\gamma} - \gamma_0) + o_p(1)
$$

By conditions (a)-(c) and the uniform law of large number theorem (Theorem 4.1 in Wooldridge (1994)), there exist matrices $M_{s,\theta}^{ij}$ and $M_{s,\gamma}^{ij}$ such that

$$
\frac{\partial}{\partial \theta_i} \hat{S}(\theta_0, \gamma_0)_{ij} \overset{p}{\to} M_{s,\theta}^{ij},
$$

$$
\frac{\partial}{\partial \gamma_j} \hat{S}(\theta_0, \gamma_0)_{ij} \overset{p}{\to} M_{s,\gamma}^{ij}.
$$

In the meanwhile, expanding $G_n(\hat{\theta}, \hat{\gamma})$ around $(\theta_0^T, \gamma_0^T)$, we have

$$
G_n(\hat{\theta}, \hat{\gamma}) = G_n(\theta_0, \gamma_0) + G_{n,\gamma}(\hat{\theta}, \hat{\gamma})(\hat{\gamma} - \gamma_0) + G_{n,\theta}(\hat{\theta}, \hat{\gamma})(\hat{\theta} - \theta_0),
$$
where \( \tilde{\theta} \) lies between \( \theta_0 \) and \( \hat{\theta} \) and \( \tilde{\gamma} \) lies between \( \gamma_0 \) and \( \bar{\gamma} \). Here \( G_{n,\tilde{\theta}}(\cdot) \) and \( G_{n,\tilde{\gamma}}(\cdot) \) are the first-order partial derivatives of \( G_n(\cdot) \) with respect to \( \theta \) and \( \gamma \), respectively. By conditions (a) and (b), applying the uniform law of large number, we have \( G_{n,\tilde{\theta}}(\tilde{\theta}, \bar{\gamma}) \) and \( G_{n,\tilde{\gamma}}(\tilde{\theta}, \bar{\gamma}) \) converge to some constant matrices, denoted by \( K_{g,1} \) and \( K_{g,2} \), respectively.

Now, denote martingale estimating equation (5) by \( H_n(\bar{\gamma}) = \frac{1}{n} \sum_{k=1}^{n} h(X_{(k-1)\Delta}, X_{k\Delta}; \gamma) \). Because \( \bar{\gamma} \) is the root of equation (5), the mean-value theorem implies,

\[
H_n(\bar{\gamma}) = H_n(\gamma_0) + H_{n,\gamma}(\bar{\gamma})(\bar{\gamma} - \gamma_0),
\]

where \( \gamma_0 \) and \( \gamma \) and \( H_{n,\gamma}(\cdot) \) is the first-order derivative of \( H_n \) w.r.t. \( \gamma \). Under conditions (a) and (b), applying the uniform law of large number, there exist a matrix \( K_h \), such that \( H_{n,\gamma}(\bar{\gamma}) \xrightarrow{p} K_h \). Thus, we have

\[
\sqrt{n} (\bar{\gamma} - \gamma_0) = -K_h^{-1} \frac{1}{\sqrt{n}} \sum_{k=1}^{n} h(X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + o_p(1)
\]

Using similar arguments to those given above, we obtain

\[
\sqrt{n} (\tilde{\theta} - \theta_0) = -G_{n,\tilde{\theta}}^{-1}(\tilde{\theta}, \bar{\gamma}) \sqrt{n} G_n(\theta_0, \gamma_0) + G_{n,\tilde{\theta}}^{-1}(\tilde{\theta}, \bar{\gamma}) G_{n,\tilde{\gamma}}(\tilde{\theta}, \bar{\gamma}) H_{n,\gamma}^{-1}(\bar{\gamma}) \sqrt{n} H_n(\gamma_0)
\]

\[
\triangleq \frac{1}{\sqrt{n}} \sum_{k=1}^{n} f_{\tilde{\theta}}(X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + o_p(1)
\]

Similarly, we have

\[
\sqrt{n} (\hat{S}(\tilde{\theta}, \bar{\gamma})_{ij} - S(\theta_0, \gamma_0)_{ij}) = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} g_s(X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) - S(\theta_0, \gamma_0)
\]

\[
+ \frac{1}{\sqrt{n}} \sum_{k=1}^{n} M_{s,\theta} f_{\theta}(X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0)
\]

\[
+ \frac{1}{\sqrt{n}} \sum_{k=1}^{n} M_{s,\gamma} f_{\gamma}(X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + o_p(1)
\]

\[
\triangleq \frac{1}{\sqrt{n}} \sum_{k=1}^{n} h_{\hat{S}}^{ij}(X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + o_p(1),
\]
Again, using similar argument, we have

\[ \sqrt{n} (\hat{\theta}, \hat{\gamma})_{ij} - V(\theta_0, \gamma_0)_{ij} \]

\[ = \sqrt{n} (\hat{\theta}, \hat{\gamma})_{ij} - V(\theta_0, \gamma_0)_{ij} + \frac{\partial}{\partial \theta^T} \hat{\theta} (\theta_0, \gamma_0)_{ij} \sqrt{n} (\hat{\theta} - \theta_0) + \frac{\partial}{\partial \gamma^T} \hat{\gamma} (\theta_0, \gamma_0)_{ij} \sqrt{n} (\hat{\gamma} - \gamma_0) + o_p(1) \]

\[ = \frac{1}{\sqrt{n}} \sum_{k=1}^{n} \mathcal{G}_v (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) - V(\theta_0, \gamma_0) \]

\[ + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} M_{ij}^{k, \theta} f_\theta (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) \]

\[ + \frac{1}{\sqrt{n}} \sum_{k=1}^{n} M_{ij}^{k, \gamma} f_\gamma (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + o_p(1) \]

\[ \Delta \frac{1}{\sqrt{n}} \sum_{k=1}^{n} h_{ij}^0 (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + o_p(1), \]

where \( M_{ij}^{k, \theta} \) and \( M_{ij}^{k, \gamma} \) are some matrices such that

\[ \frac{\partial}{\partial \theta^T} \hat{\theta} (\theta_0, \gamma_0)_{ij} \xrightarrow{pr} M_{ij}^{k, \theta}, \]

\[ \frac{\partial}{\partial \gamma^T} \hat{\gamma} (\theta_0, \gamma_0)_{ij} \xrightarrow{pr} M_{ij}^{k, \gamma}. \]

By conditional (a)-(c), applying the uniform law of large number and the martingale central limit theorem in Billingsley (1961), we have \( \sqrt{n} (\hat{\theta}, \hat{\gamma}) - (\theta_0, \gamma_0) = O_p(\frac{1}{\sqrt{n}}) \). Therefore, we can reach the following expression:

\[ \sqrt{n} \{ R_n - p \} = \text{tr} \left\{ \frac{1}{\sqrt{n}} \sum_{k=1}^{n} h_{ij}^0 (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) \right\} + o_p(1) \]

where \( h_{ij}^0 = S^{-1}(\theta_0, \gamma_0) \left\{ h_s (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) + h_V (X_{(k-1)\Delta}, X_{k\Delta}; \theta_0, \gamma_0) \right\}, h_s \) and \( h_V \) are \( p \times p \) matrices with element \( h_{ij}^s \) and \( h_{ij}^V \), respectively, \( i, j = 1, ..., p \). Using the Billingsley (1961)'s martingale central limit theorem, we obtain

\[ \sqrt{n} \{ R_n - p \} \xrightarrow{d} N(0, \sigma_R^2). \]

where \( \sigma_R^2 \) is the asymptotic variance, which can be consistently estimated by

\[ \hat{\sigma}_R^2 = \frac{1}{n} \sum_{k=1}^{n} \left( \sum_{i=1}^{p} h_{ij}^i \right)^2 \]

where \( h_{ij}^i \) is the \( i \)-th diagonal element of matrix \( h_{R_n} \).

\[ \blacksquare \]
References
