Econometrica Supplementary Material

SUPPLEMENT TO: "ASSESSMENT OF UNCERTAINTY IN HIGH FREQUENCY DATA: THE OBSERVED ASYMPTOTIC VARIANCE": PROOFS AND TECHNICAL ISSUES

(*Econometrica*, Vol. 85, No. 1, January 2017, 197–231)

BY PER A. MYKLAND AND LAN ZHANG

APPENDIX A: GENERAL RESULTS ON THE TRIANGULAR ARRAY CONVERGENCE OF THE QUADRATIC VARIATION OF SEMIMARTINGALES

DEFINITION 5—Orders in Probability: For a sequence $\alpha_t^{(n)}$ of semimartingales, we say that $(\alpha_t^{(n)}) = O_p(1)$ if the sequence is tight, with respect to convergence in law relative to the Skorokhod topology on $\mathbb{D}$ (Jacod and Shiryaev (2003, Theorem VI.3.21, p. 350)), and also P-UT (ibid., Chapter VI.3.b, and Definition VI.6.1, p. 377). For scalar random quantities, $O_p(\cdot)$ and $o_p(\cdot)$ are defined as usual; see, for example, Pollard (1984, Appendix A).

CONDITION 4: Let $\alpha_t^{(n)}$ and $\beta_t^{(n)}$ be sequences (in n) of semimartingales. Each of these sequences is (separately) assumed to be $O_p(1)$.

DEFINITION 6—Notation: The symbol $\mathbb{F}$ will refer to a collection of nonrandom functions $f_\cdot^{(l,n)} \in \mathbb{D}[0, \mathcal{T}]$, $n \in \mathbb{N}$, and $l = 1, \dots, 2K_n$ satisfying

$$\text{(A.1)} \qquad \left|f_t^{(l,n)}\right| \leq 1 \quad \text{for all } t, l, \text{ and } n.$$

Similarly, $\mathbb{G}$ will refer to a collection $g_t^{(l,n)}$ with the same size and properties.

Given $\mathbb{F}$ and $\mathbb{G}$, set

$$\text{(A.2)} \qquad \alpha_t^{(l,n)} = \int_0^t f_{s-}^{(l,n)}\, d\alpha_s^{(n)} \quad \text{and} \quad \beta_t^{(l,n)} = \int_0^t g_{s-}^{(l,n)}\, d\beta_s^{(n)} \quad \text{for} \quad l = 1, \dots, 2K_n.$$

Also,

$$\text{(A.3)} \qquad i \equiv L[2K] \quad \text{means that} \quad i = 2Kj + L, \quad \text{where} \quad j \text{ is an integer.}$$

DEFINITION 7—Decomposition of $\mathbb{F}$ and $\mathbb{G}$ by Block: Recall that B_n is the set of basic blocks, and that $\Delta T_n = \mathcal{T}/B_n$. With reference to the collection $\mathbb{F}$: For given (l, n), the function $f_t^{(l,n)}$ is allowed to jump at times $T_{K_n j + l}$ but must otherwise satisfy certain compactness properties.

Specifically, for each $n \in \mathbb{N}$, and $l = 1, \dots, 2K_n$, define, for $j \in \mathbb{N} \cap [1, (B_n - l)/(K_n + 1)]$,

$$\text{(A.4)} \qquad f_t^{(l,j,n)} = \begin{cases} f_{T_{K_n j+l}}^{(l,n)} & \text{for } t \in [0, T_{K_n j+l}), \\ f_t^{(l,n)} & \text{for } t \in [T_{K_n j+l}, T_{(K_n+1)j+l}), \\ \lim\limits_{t \uparrow T_{(K_n+1)j+l}} f_t^{(l,n)} & \text{for } t \in [T_{(K_n+1)j+l)}, \mathcal{T}]. \end{cases}$$

The set of such $f_t^{(l,j,n)}$ will be denoted $\mathbb{F}'$. $\mathbb{G}'$ is defined similarly.

THEOREM 7—Consistency of Triangular Array Rolling Quadratic Variation: *Under Condition* 4, *assume* (A.1), *and that the sets* $\mathbb{F}'$ *and* $\mathbb{G}'$ (*from Definition* 7) *are relatively*

 DOI: 10.3982/ECTA12501

compact for the Skorokhod topology.[42] *Also suppose that $K_n \Delta T_n \to 0$ as $n \to \infty$. Then,*

$$\frac{1}{2K_n} \sum_{l=1}^{2K_n} \sum_{K_n \le i \le B_n - K_n, i \equiv l[2K_n]} \left(\alpha^{(l,n)}_{T_{i+K_n}} - \alpha^{(l,n)}_{T_{i-K_n}}\right)^2 = \frac{1}{2K_n} \sum_{l=1}^{2K_n} \left[\alpha^{(l,n)}, \alpha^{(l,n)}\right]_{\mathcal{T}} + o_p(1), \tag{A.5}$$

and similarly for β. Also,

$$\begin{aligned} &\frac{1}{2K_n} \sum_{l=1}^{2K_n} \sum_{K_n \le i \le B_n - K_n, i \equiv l[2K_n]} \left(\alpha^{(l,n)}_{T_{i+K_n}} - \alpha^{(l,n)}_{T_{i-K_n}}\right)\left(\beta^{(l,n)}_{T_{i+K_n}} - \beta^{(l,n)}_{T_{i-K_n}}\right) \\ &\quad = \frac{1}{2K_n} \sum_{l=1}^{2K} \left[\alpha^{(l,n)}, \beta^{(l,n)}\right]_{\mathcal{T}} + o_p(1). \end{aligned} \tag{A.6}$$

REMARK 15—Uniformity in ΔT: Theorem 7 does not impose any requirement on ΔT_n, except that $\Delta T_n > 0$ and $K_n \Delta T_n \to 0$. See the final comment in the proof of the theorem.

Before proving our results, we recall the following useful concept.

DEFINITION 8—The Canonical Decomposition of α: We shall be using the *canonical decomposition* of α_t (Jacod and Shiryaev (2003, Chap. II.2a pp. 75–76)), which is defined for a general semimartingale (ibid. Definition I.4.21, p. 43), by writing

$$\alpha_t = \alpha_0 + \alpha(h)_t + B(h)_t + \check{\alpha}(h)_t. \tag{A.7}$$

h is called the truncation function. Compared to the notation in our reference work, their X is our α, their $M(h)$ is our $\alpha(h)$, while their $B(h)$ is the same as ours. Also, let $\tilde{C}_t = \langle \alpha(h), \alpha(h) \rangle$. This is the "second modified characteristic" (ibid., Definition II.2.16, p. 79). For the case of no truncation function, α can similarly be decomposed into a local martingale and a finite variation process A_t. See also ibid., p. 84, for further clarification of the relationship between the untruncated and the truncated processes. We let TV denote total variation,[43] and set

$$D(\alpha)(h)_t = \mathrm{TV}(\check{\alpha})_t + \mathrm{TV}\big(B(h)\big)_t. \tag{A.8}$$

Similar notation applies to $\alpha^{(n)}$, $\beta^{(n)}$, etc.

PROOF OF THEOREM 7: We prove (A.5). The result (A.6) is obtained similarly but with longer notation. For (A.6), we specifically need that $\alpha^{(n)}$ and $\beta^{(n)}$ be tight, which is assumed, and that $D(\alpha^{(n)})(h)_{\mathcal{T}}$, $D(\beta^{(n)})(h)_{\mathcal{T}}$, $\langle \alpha^{(n)}(h), \alpha^{(n)}(h) \rangle_{\mathcal{T}}$, $\langle \beta^{(n)}(h), \beta^{(n)}(h) \rangle_{\mathcal{T}}$, and $\langle \beta^{(n)}(h), \alpha^{(n)}(h) \rangle_{\mathcal{T}}$ be tight. The first four of these follow from the P-UT property of $\alpha^{(n)}$ and $\beta^{(n)}$ (Jacod and Shiryaev (2003, Theorem VI.6.15)), the final one since $|\langle \beta^{(n)}(h), \alpha^{(n)}(h) \rangle_{\mathcal{T}}| \le (\langle \alpha^{(n)}(h), \alpha^{(n)}(h) \rangle_{\mathcal{T}} + \langle \beta^{(n)}(h), \beta^{(n)}(h) \rangle_{\mathcal{T}})/2$.

[42]A criterion can be found in Jacod and Shiryaev (2003, Theorem VI.1.14(b), p. 328). The condition is satisfied in all our applications (B.1), (C.6), and (C.11).

[43] Jacod and Shiryaev denoted the total variation by Var.

In analogy with (A.2), define $\alpha_t^{(l,,j,n)} = \int_0^t f_{s-}^{(l,j,n)} d\alpha_s^{(n)}$. Also, define

$$Z_{n,l}(t) = \sum_{T_{i+K_n} \le t, i \equiv l[2K]} \left(\alpha_{T_{i+K_n}}^{(l,n)} - \alpha_{T_{i-K_n}}^{(l,n)}\right)^2 + \left(\alpha_t^{(l,n)} - \alpha_{T_{*,L}}^{(l,n)}\right)^2 - \left[\alpha^{(l,n)}, \alpha^{(l,n)}\right]_t, \tag{A.9}$$

where $T_{*,L} = \max\{T_i : T_{i+K_n} \le t, i \equiv L[2K_n]\}$, so that

$$dZ_{n,l}(t) = 2\left(\alpha_{t-}^{(l,n)} - \alpha_{T_{*,l}}^{(l,n)}\right) d\alpha_t^{(l,n)}. \tag{A.10}$$

For given truncation function h, define the processes $\alpha_t^{(l,n)}(h) = \int_0^t f_s^{(l,n)} d\alpha^{(n)}(h)_s$, $\breve{\alpha}_t^{(l,n)}(h) = \int_0^t f_s^{(l.n)} d\breve{\alpha}(h)_s$, etc. (The truncation is done on the original jumps, those of $\alpha_t^{(n)}$, and not starting with the process $\alpha_t^{(l,n)}$. This assures uniformity in the following argument.) Similarly, define $dZ_{l,n}(h)(t) = 2(\alpha_{t-}^{(l,n)} - \alpha_{T_{*,l}}^{(l,n)}) d\alpha^{(l,n)}(h)_t$, starting at $Z_{l,n}(h)(0) = Z_{l,n}(0) = 0$. Also set

$$Z_n(t) = \frac{1}{2K_n} \sum_{l=1}^{2K_n} Z_{l,n}(t) \quad \text{and} \quad Z_n(h)(t) = \frac{1}{2K_n} \sum_{L=1}^{2K_n} Z_{l,n}(h)(t). \tag{A.11}$$

Observe that $Z_n(\mathcal{T}) =$ the difference between the explicit terms on left- and right-hand sides of (A.5).

To bound the difference between $Z_n(t)$ and $Z_n(h)(t)$, note that

$$\left|Z_{l,n}(h)(t) - Z_{l,n}(t)\right| \le 2 \int_0^t \left|\alpha_{s-}^{(l,n)} - \alpha_{T_{*,l}}^{(l,n)}\right| dD^{(n)}(h)_t, \tag{A.12}$$

where $D^{(n)}(h)$ is defined as in (A.8), and with the original $\alpha^{(n)}$. Also, in the notation of Jacod and Shiryaev (2003, VI.1.8, p. 326), it follows from (A.1) that, for all $t \in [0, \mathcal{T}]$ and all $s \in [T_{*,L}, t]$,

$$\begin{aligned} \left|\alpha_{s-}^{(l,n)} - \alpha_{T_{*,l}}^{(l,n)}\right| &\le 2 \max_j w'_{\mathcal{T}}\left(\alpha^{(l,j,n)}, K_n \Delta T_n\right) + \sup_{T_{*,L} < s < t} \left|\Delta\alpha_s^{(n)}\right| \\ &\le 2 \max_j w'_{\mathcal{T}}\left(\alpha^{(l,j,n)}, K_n \Delta T_n\right) + v_n(t-), \end{aligned} \tag{A.13}$$

where $v_n(t-) = \sup_{T_{**} < s < t} |\Delta\alpha_s^{(n)}|$, with $T_{**} = \max\{T_i : T_{i+2K_n} \le t, \}$, so that

$$\begin{aligned} \sup_{0 \le t \le \mathcal{T}} \left|Z_n(h)(t) - Z_n(t)\right| &\le 4 \max_{l,j} w'_{\mathcal{T}}\left(\alpha^{(l,j,n)}(h), K\Delta T\right) D^{(n)}(h)(\mathcal{T}) \\ &\quad + 2 \int_0^{\mathcal{T}} v_n(t-)\, dD^{(n)}(h)_t. \end{aligned} \tag{A.14}$$

This is because the right-hand side bounds $\sup_{0 \le t \le \mathcal{T}} |Z_{l,n}(h)(t) - Z_{l,n}(t)|$ for each l, and thus the average.

Meanwhile, to assess the size of $Z_n(h)_t$, by similar argument,

$$\left\langle Z_n(h), Z_n(h)\right\rangle_{\mathcal{T}} \le 8\left(4 \max_{l,j} w'_{\mathcal{T}}\left(\alpha^{(l,j,n)}, K\Delta T\right)^2 \tilde{C}_T^{(n)} + \int_0^{\mathcal{T}} v_n^2(t-)\, d\tilde{C}_t^{(n)}\right). \tag{A.15}$$

This is because the same bound applies to each $\langle Z_{n,l_1}(h), Z_{n,l_2}(h)\rangle_{\mathcal{T}}$.

We now seek to describe the asymptotic behavior of $\max_{l,j} w'_{\mathcal{T}}(\alpha^{(l,j,n)}(h), K_n\Delta T_n)$ and $v_n(t-)$ so as to control the asymptotic behavior of (A.14)–(A.15).

On the one hand, since $\mathbb{F}'$ from Definition 7 is relatively compact for the Skorokhod topology (ex hypothesi), we obtain from Jacod and Shiryaev (2003, Theorem VI.3.21, p. 350, and Theorem VI.6.22, p. 383) that

$$\text{(A.16)} \quad \max_{l,j} w'_{\mathcal{T}}\big(\alpha^{(l,j,n)}(h), K_n\Delta T_n\big) \xrightarrow{p} 0 \quad \text{as} \quad n \to \infty.$$

On the other hand, we bound $v_n(t-)$ as follows. Let $\varepsilon > 0$ be arbitrary. Since $\alpha^{(n)}$ is tight, we shall without loss of generality be working with a convergent subsequence so that $\alpha^{(n)} \xrightarrow{\mathcal{L}} \alpha$. Redo the canonical decomposition (Definition 8) with a specific truncation function given by $h_\varepsilon(x) = x$ if $|x| \le \varepsilon$, and $= \varepsilon \operatorname{sgn}(x)$ otherwise:

$$\text{(A.17)} \quad \begin{aligned} &\alpha_t^{(n)} = \alpha_0^{(n)} + \alpha^{(n)}(h_\varepsilon)_t + B^{(n)}(h_\varepsilon)_t + \breve{\alpha}^{(n)}(h_\varepsilon)_t \quad \text{and} \\ &\alpha_t = \alpha_0 + \alpha(h_\varepsilon)_t + B(h_\varepsilon)_t + \breve{\alpha}(h_\varepsilon)_t. \end{aligned}$$

Set $v_{n,\varepsilon}(t-) = \sup_{T_{**}<s<t} |\Delta\breve{\alpha}^{(n)}(h_\varepsilon)_s|$ and observe that

$$\text{(A.18)} \quad v_n(t-) \le v_{n,\varepsilon}(t-) + \varepsilon.$$

Let $\tau_{n,i}$ be the ith jump time of $\breve{\alpha}^{(n)}(h_\varepsilon)_t$, with $\tau_{n,0} = 0$. Similarly, τ_i is the ith jump time of $\breve{\alpha}(h_\varepsilon)_t$. We note that, for given $t \in [0, \mathcal{T}]$, and for any $\delta > 0$,

$$\text{(A.19)} \quad \begin{aligned} \big\{v_{n,\varepsilon}(t-) = 0\big\} &\supseteq \bigcup_i \{\tau_{n,i} \ge t \ge \tau_{n,i-1} + 2K_n\Delta T_n\} \\ &\supseteq \bigcup_i \{\tau_{n,i} \ge t \ge \tau_{n,i-1} + \delta\} \end{aligned}$$

as soon as $\delta \ge 2K_n\Delta T_n$ (and this does happen eventually, by assumption). By invoking Jacod and Shiryaev (2003, Proposition VI.3.15, p. 349) with $\tau_{n,i}$ as $T_i(\breve{\alpha}^{(n)}(h_\varepsilon), \frac{\varepsilon}{2})$ and τ_i as $T_i(\breve{\alpha}(h_\varepsilon), \frac{\varepsilon}{2})$, the proposition yields that $(\tau_{n,1}, \dots, \tau_{n,k}) \xrightarrow{\mathcal{L}} (\tau_1, \dots, \tau_k)$ as $n \to \infty$ for any k. This is because the process $\breve{\alpha}^{(n)}(h_\varepsilon)$ converges in law to $\breve{\alpha}(h_\varepsilon)$ in view of ibid., Proposition VI.3.16, p. 349.

By approximating the indicator of the set $\{\tau_{n,i} \ge t \ge \tau_{n,i-1}\}$ by a continuous function, and then undoing the approximation, we obtain $P\{\tau_{n,i} \ge t \ge \tau_{n,i-1} + \delta\} \to P\{\tau_i \ge t \ge \tau_{i-1} + \delta\}$ as $n \to \infty$. Since the union (A.19) is disjoint, it follows that

$$\text{(A.20)} \quad \begin{aligned} \liminf_n P\big\{v_{n,\varepsilon}(t-) = 0\big\} &\ge \sum_{i=1}^k P\{\tau_i \ge t \ge \tau_{i-1} + \delta\} \\ &\to P\{\tau_k \ge t\} \quad \text{as} \quad \delta \downarrow 0 \\ &\to 1 \quad \text{as} \quad k \to \infty. \end{aligned}$$

Hence from (A.18), $P\{v_n(t-) \ge \varepsilon\} \to 0$. Since ε was arbitrary, we obtain

$$\text{(A.21)} \quad \forall t \in [0, \mathcal{T}]\colon \quad v_n(t-) \xrightarrow{p} 0 \quad \text{and} \quad \big|v_n(t-)\big| \le \sup_{0\le s\le \mathcal{T}} \big|\Delta\alpha_s^{(n)}\big|,$$

the latter statement assuring dominated convergence.

We can now combine (A.14)–(A.15) with (A.16) and (A.21) to obtain, as $n \to \infty$,

$$\text{(A.22)} \qquad \sup_{0\le t\le \mathcal{T}} \left| Z_n(h)(t) - Z_n(t) \right| \xrightarrow{p} 0 \quad \text{and}$$

$$\langle Z_n(h), Z_n(h) \rangle_{\mathcal{T}} \xrightarrow{p} 0.$$

The transition to (A.22) did not assume that $D^{(n)}(h)_t$ or $\tilde{C}^{(n)}_{\mathcal{T}}$ has a limit as $n \to \infty$. By the assumption that the $\alpha^{(n)}_t$ is $O_p(1)$ and hence P-UT, however, Jacod and Shiryaev (2003, Theorem VI.6.15, p. 380) yields that $D^{(n)}(h)_{\mathcal{T}}$ and $\tilde{C}^{(n)}_{\mathcal{T}}$ are tight.

From the second line in (A.22), by Lenglart's inequality (Jacod and Shiryaev (2003, Lemma I.3.30, p. 35)),

$$\text{(A.23)} \qquad \sup_{0\le t\le \mathcal{T}} \left| Z_n(h)(t) \right| \xrightarrow{p} 0.$$

Combining (A.23) with the first line of (A.22) yields the result of the theorem, since $Z_n(\mathcal{T}) =$ the left-hand side of (A.5). Since none of the bounds used depend on ΔT_n but only on $K_n \Delta T_n$, the result does not impose any requirement on ΔT_n, except that $\Delta T_n > 0$ and $K_n \Delta T_n \to 0$. *Q.E.D.*

APPENDIX B: RESULTS ON THE QUADRATIC VARIATION OF θ: TIGHTNESS AND CONVERGENCE PROPERTIES

PROOF OF THEOREM 1: Because we shall use Theorem 7, we here let all quantities depend on index n. Thus, unlike Definition 2 in Section 2, $K = K_n$, etc., though we shall often omit the subscript when the meaning is obvious. For the purposes of the current proof, one can simply take $n = B$, but this will no longer be the case in later appendices. Set

$$\text{(B.1)} \qquad f^{(l,n)}_t = \frac{1}{K\Delta T} \sum_{K\le i\le B-K;\, i\equiv l[2K]} \big((T_{i+K} - t) I\{T_{i+K} > t \ge T_i\} + (t - T_{i-K}) I\{T_i > t \ge T_{i-K}\} \big),$$

where $i \equiv l[2K]$ means that i is on the form $2Ki + l$. We note that $f^{(l)}_t = f^{(l,n)}_t$ depends on n through ΔT, K, and B. It is easy to see that the family $\mathbb{F} = \{f^{(l,n)}\}$ satisfies (A.1), and that the set $\mathbb{F}'$ (from Definition 7) is indeed relatively compact for the Skorokhod topology.

Define the processes $\theta^{(l,n)}_t = \int_0^t f^{(l,n)}_{s-}\, d\theta_s$. To motivate the following development, note from Theorem 2 in Section 2.3 that, for fixed $i \equiv l[2K]$,

$$\text{(B.2)} \qquad \frac{1}{K(\Delta T)} (\Theta_{(T_i, T_{i+K}]} - \Theta_{(T_{i-K}, T_i]}) = \frac{1}{K(\Delta T)} (\Theta'_{(T_i, T_{i+K}]} + \Theta''_{(T_{i-K}, T_i]})$$
$$= \int_{T_{i-K}}^{T_{i+K}} f^{(l,n)}_t \, d\theta$$
$$= \theta^{(l,n)}_{T_{i+K}} - \theta^{(l,n)}_{T_{i-K}},$$

whence

$$\text{(B.3)}\qquad \frac{1}{K^2(\Delta T)^2}\sum_{K\le i\le B-K,i\equiv l[2K]}(\Theta_{(T_i,T_{i+K}]}-\Theta_{(T_{i-K},T_i]})^2=\sum_{K\le i\le B-K,i\equiv l[2K]}\bigl(\theta^{(l,n)}_{T_{i+K}}-\theta^{(l,n)}_{T_{i-K}}\bigr)^2$$

and

$$\text{(B.4)}\qquad \frac{1}{2}\frac{1}{K^2(\Delta T)^2}\,\mathrm{QV}_{B,K}(\Theta)=\frac{1}{2K_n}\sum_{l=1}^{2K}\sum_{K\le i\le B-K,i\equiv l[2K]}\bigl(\theta^{(l,n)}_{T_{i+K}}-\theta^{(l,n)}_{T_{i-K}}\bigr)^2.$$

We now wish to show that

$$\text{(B.5)}\qquad \begin{aligned}\frac{1}{2K_n}\sum_{l=1}^{2K}\sum_{K\le i\le B-K,i\equiv l[2K]}\bigl(\theta^{(l)}_{T_{i+K}}-\theta^{(l)}_{T_{i-K}}\bigr)^2&=\frac{1}{2K}\sum_{l=1}^{2K}\bigl[\theta^{(l,n)},\theta^{(l,n)}\bigr]_{\mathcal{T}}+o_p(1)\\&=\int_0^{\mathcal{T}}f_t^{(n)}\,d[\theta,\theta]_t+o_p(1),\quad\text{where}\end{aligned}$$

$$\text{(B.6)}\qquad \begin{aligned}f_t^{(n)}&=\frac{1}{2K}\sum_{l=1}^{2K}\bigl(f_t^{(l,n)}\bigr)^2\\&=\frac{1}{2K^3(\Delta T)^2}\sum_{K\le i\le B-K}\bigl((T_{i+K}-t)^2I\{T_{i+K}\ge t>T_i\}\\&\quad+(t-T_{i-K})^2I\{T_i\ge t>T_{i-K}\}\bigr).\end{aligned}$$

If K is finite, this is a simple matter of checking that

$$\sum_{K\le i\le B-K,i\equiv l[2K]}\bigl(\theta^{(l)}_{T_{i+K}}-\theta^{(l)}_{T_{i-K}}\bigr)^2=\bigl[\theta^{(l,n)},\theta^{(l,n)}\bigr]_{\mathcal{T}}+o_p(1)\quad\text{for each } l=1,\dots,2K,$$

where we recall that $i\equiv l[2K]$ means that i is on the form $2Ki+l$. For the general case where K can be finite or infinite, we proceed as follows. The class of functions $f_t^{(l,n)}$ given by (B.1) satisfies the conditions of Theorem 7. So does $\alpha_t^{(n)}=\theta_t$; since the process does not move with n, it is both tight and P-UT. Theorem 7 therefore yields (B.5)-(B.6).

For $t\in(T_{j-1},T_j]\subseteq(T_K,T_{B-K}]$,

$$\text{(B.7)}\qquad \begin{aligned}f_t^{(n)}&=\frac{1}{2K^3(\Delta T)^2}\Biggl(\sum_{j-K\le i\le j-1}(T_{i+K}-t)^2+\sum_{j\le i\le j+K-1}(t-T_{i-K})^2\Biggr)\\&=\frac{1}{3}\Bigl(1-\frac{1}{K^2}\Bigr)+\frac{1}{2}\frac{1}{K^2}\Biggl(\Bigl(\frac{T_j-t}{\Delta T}\Bigr)^2+\Bigl(\frac{t-T_{j-1}}{\Delta T}\Bigr)^2\Biggr),\end{aligned}$$

hence, eventually, on all $[\delta,\mathcal{T}-\delta]$, for any $\delta>0$. Since, for all $t\in[0,\mathcal{T}]$, $0\le f_t^{(n)}\le 1$, and since $f_{\mathcal{T}}^{(n)}=0$, Theorem 1 follows. Remark 15 in Appendix A continues to apply, for the same reasons. *Q.E.D.*

PROOF OF THEOREM 2: By Itô's formula, $d(T+\delta-t)(\theta_t-\theta_T)=(T+\delta-t)\,d\theta_t-(\theta_t-\theta_T)\,dt$. Integrating from T to $T+\delta$ yields

$$\text{(B.8)}\qquad 0=\Theta'_{(T,T+\delta]}-\Theta_{(T,T+\delta]}+\theta_T\delta.$$

Similarly, $d(t-(T-\delta))(\theta_T-\theta_t) = -(t-(T-\delta))\,d\theta_t + (\theta_T-\theta_t)\,dt$. Integrating from $T-\delta$ to T yields

$$0 = -\Theta''_{(T-\delta,T]} - \Theta_{(T-\delta,T]} + \theta_T\delta. \tag{B.9}$$

Combining (B.8)–(B.9) yields the result. *Q.E.D.*

APPENDIX C: PROOF OF THEOREM 3, AND A MORE GENERAL RESULT

We here show a broader result of which Theorem 3 is a corollary. First of all, we replace the "omnibus" Condition 1 by the weaker and more precise Condition 5. Also, it shows what happens when one gives up on forcing negligibility in the form of conditions (22) and $\Delta T = o(n^{-\alpha})$. The former is conceptually important as it separates out what part of Condition 1 is required for the convergence of quadratic variations (as opposed to being a valid asymptotic variance). The latter is useful in case one were tempted to take K fixed in the discontinuous θ_t case. We first state and prove the more general Theorem 8, and then derive Theorem 3.

CONDITION 5—Relative Size of Semimartingale and Edge Effect in $\hat{\Theta}$ in (13): We assume that $M_{n,t}$ is a sequence of semimartingales. We assume that there is a rate $\alpha > 0$ (which need not be known) so that the sequence of semimartingales $(n^\alpha M_{n,t}) = O_p(1)$ in the sense of Definition 5 in Appendix A. We assume that $e_{n,T} = o_p(n^{-\alpha})$ and $\tilde{e}_{n,S} = o_p(n^{-\alpha})$ for any S and T.

THEOREM 8—More General Expansion of $\mathrm{QV}_{B,K}(\widehat{\Theta})$: *Assume that θ_t is a semimartingale on $[0,\mathcal{T}]$, and suppose that Condition 5 holds. Define*

$$\mathrm{QV}_{B,K}(\Theta, M) = \frac{1}{K}\sum_{i=K}^{B-K}(\Theta_{(T_i,T_{i+K}]} - \Theta_{(T_{i-K},T_i]})\big((M_{T_{i+K}} - M_{T_i}) - (M_{T_i} - M_{T_{i-K}})\big), \tag{C.1}$$

$$\mathrm{QV}_{B,K}(M) = \frac{1}{K}\sum_{i=K}^{B-K}\big((M_{T_{i+K}} - M_{T_i}) - (M_{T_i} - M_{T_{i-K}})\big)^2, \quad \text{and}$$

$$R_{n,K} = \frac{1}{K}\sum_{i=K}^{B-K}(\tilde{e}_{T_{i+K}} - e_{T_i} - \tilde{e}_{T_i} + e_{T_{i-K}})^2,$$

and also

$$\overline{\mathrm{QV}}_{B,K}(\hat{\Theta}) = \mathrm{QV}_{B,K}(\Theta) + 2\,\mathrm{QV}_{B,K}(\Theta, M) + \mathrm{QV}_{B,K}(M). \tag{C.2}$$

Let $K = K_n$ be positive integers, and assume that $K_n\Delta T_n \to 0$. Then, in extension of (23),

$$\frac{1}{2K}\sum_{K\le i\le B-K}(\hat{\Theta}_{(T_{i-K},T_{i+K}]} - \Theta_{(T_{i-K},T_{i+K}]})^2 = [M_n, M_n]_{\mathcal{T}} + R_{n,K} + O_p\big(n^{-\alpha}R_{n,K}^{1/2}\big). \tag{C.3}$$

Also, in extension of (25),

$$\overline{\mathrm{QV}}_{B,K}(\hat{\Theta}) = 2[M_n, M_n]_{\mathcal{T}} + (K\Delta T)^2 \frac{2}{3}\left(1 - \frac{1}{K^2}\right)[\theta, \theta]_{\mathcal{T}-} + (\Delta T)^2 \int_0^{\mathcal{T}} \left(\left(\frac{t^* - t}{\Delta T}\right)^2 + \left(\frac{t - t_*}{\Delta T}\right)^2\right) d[\theta, \theta]_t + 2\Delta T \int_0^{\mathcal{T}} \left(1 - 2\frac{t - t_*}{\Delta T}\right) d[\theta, M_n]_t + o_p\big((K_n\Delta T)^2\big) + o_p\big(n^{-2\alpha}\big) \tag{C.4}$$

and

$$\mathrm{QV}_{B,K}(\hat{\Theta}) = \overline{\mathrm{QV}}_{B,K}(\hat{\Theta}) + R_{n,K} + O_p\big((K\Delta T + n^{-\alpha})R_{n,K}^{1/2}\big). \tag{C.5}$$

The convergence in probability is uniform in ΔT_n, so long as $\Delta T_n > 0$ and $K_n \Delta T_n \to 0$.

For the proofs, set $\alpha_t^{(n)} = \theta_t$, $\beta_t^{(n)} = n^\alpha M_{n,t}$. Let $f_t^{(l,n)}$ be given by (B.1) above. We shall use two different definitions of $g_t^{(l,n)}$. For both cases, let $\alpha_t^{(l,n)}$ and $\beta_t^{(l,n)}$ be as given by (A.2).

PROOF OF (C.3) (CASE 1 FOR $g_t^{(l,n)}$): Set

$$g_t^{(l,n)} = \sum_{K \le i \le B-K; i \equiv l[2K]} I\{T_{i+K} > t \ge T_{i-K}\}. \tag{C.6}$$

From Theorem 7,

$$\begin{aligned}\frac{1}{2K}\sum_{i=K}^{B-K}\big(\beta_{T_{i+K}}^{(n)} - \beta_{T_{i-K}}^{(n)}\big)^2 &= \frac{1}{2K_n}\sum_{l=1}^{2K}\sum_{K\le i\le B-K, i\equiv l[2K]}\big(\beta_{T_{i+K}}^{(l,n)} - \beta_{T_{i-K}}^{(l,n)}\big)\big(\beta_{T_{i+K}}^{(l,n)} - \beta_{T_{i-K}}^{(l,n)}\big) \\ &= \frac{1}{2K}\sum_{l=1}^{2K}\big[\beta^{(l,n)}, \beta^{(l,n)}\big]_{\mathcal{T}} + o_p(1) \\ &= \big[\beta^{(n)}, \beta^{(n)}\big]_{\mathcal{T}} + o_p(1).\end{aligned} \tag{C.7}$$

Thus, following (13), and using (C.7), write

$$\begin{aligned}&\frac{1}{2K}\sum_{i=K}^{B-K}(\hat{\Theta}_{(T_{i-K}, T_{i+K}]} - \Theta_{(T_{i-k}, T_{i+K}]})^2 \\ &\quad = \frac{1}{2K}\sum_{i=K}^{B-K}\big(n^{-\alpha}\big(\beta_{T_{i+K}}^{(n)} - \beta_{T_{i-K}}^{(n)}\big) + (\tilde{e}_{T_{i+K}} - e_{T_i})\big)^2 \\ &\quad = n^{-2\alpha}\big[\beta^{(n)}, \beta^{(n)}\big]_{\mathcal{T}} + R_{n,K} + O_p\big((K\Delta T + n^{-\alpha})R_{n,K}^{1/2}\big),\end{aligned} \tag{C.8}$$

by Cauchy–Schwarz. Since $n^{-2\alpha}[\beta^{(n)}, \beta^{(n)}]_{\mathcal{T}} = [M_n, M_n]_{\mathcal{T}}$, (C.3) is proved. Remark 15 in Appendix A remains valid for the same reasons, and also in view of Proof of Theorem 1. *Q.E.D.*

PROOF OF THE REST OF THEOREM 8 (CASE 2 FOR $g_t^{(l,n)}$): Recall that

$$\begin{aligned}
\text{(C.9)}\qquad & \hat{\Theta}_{(T_i,T_{i+K}]} - \hat{\Theta}_{(T_{i-K},T_i]} \\
&= \Theta_{(T_i,T_{i+K}]} - \Theta_{(T_{i-K},T_i]} \\
&\quad + (M_{T_{i+K}} - M_{T_i}) - (M_{T_i} - M_{T_{i-K}}) + (\tilde{e}_{T_{i+K}} - e_{T_i} - \tilde{e}_{T_i} + e_{T_{i-K}}).
\end{aligned}$$

We obtain from Cauchy–Schwarz that

$$\text{(C.10)}\qquad \mathrm{QV}_{B,K}(\hat{\Theta}) = \overline{\mathrm{QV}}_{B,K}(\hat{\Theta}) + R_{n,K} + O_p\bigl(\overline{\mathrm{QV}}_{B,K}(\hat{\Theta})^{1/2} R_{n,K}^{1/2}\bigr),$$

whence (C.5) follows from (C.4).

It remains to show (C.4). The first term in (C.2) is covered by Theorem 1 in Section 2.3. To handle the two remaining terms, we redefine

$$\text{(C.11)}\qquad g_t^{(l,n)} = \sum_{K\le i\le B-K;\, i\equiv l[2K]} \bigl(I\{T_{i+K} > t \ge T_i\} - I\{T_i > t \ge T_{i-K}\}\bigr),$$

but keep the rest of the notation from the beginning of this section (Appendix C). Note that $f_t^{(l,n)}$ is absolutely continuous, and that $g_t^{(l,n)} = -(K\Delta T)\, df_t^{(l,n)}/dt$ (except at discontinuities), whence by Fubini's theorem, where $f_t^{(n)}$ is given in equation (B.6),

$$\begin{aligned}
\text{(C.12)}\qquad \sum_{l=1}^{2K} g_t^{(l,n)} f_t^{(l,n)} &= -\frac{1}{2}(K\Delta T)\frac{d}{dt}\sum_{l=1}^{2K}\bigl(f_t^{(l,n)}\bigr)^2 \\
&= -\bigl(K^2\Delta T\bigr)\frac{d}{dt} f_t^{(n)} \\
&= 1 - 2\frac{t-t_*}{\Delta T}
\end{aligned}$$

eventually for all $t \in [\delta, \mathcal{T} - \delta]$, by (B.7). One can alternatively verify (C.12) directly.

From Theorem 7,

$$\begin{aligned}
\text{(C.13)}\qquad \frac{1}{2} n^{2\alpha}\, \mathrm{QV}_{B,K}(M) &= \frac{1}{2K}\sum_{l=1}^{2K}\sum_{K\le i\le B-K,\, i\equiv l[2K]} \bigl(\beta_{T_{i+K}}^{(l,n)} - \beta_{T_{i-K}}^{(l,n)}\bigr)\bigl(\beta_{T_{i+K}}^{(l,n)} - \beta_{T_{i-K}}^{(l,n)}\bigr) \\
&= \frac{1}{2K}\sum_{l=1}^{2K}\bigl[\beta^{(l,n)}, \beta^{(l,n)}\bigr]_{\mathcal{T}} + o_p(1) \\
&= \bigl[\beta^{(n)}, \beta^{(n)}\bigr]_{\mathcal{T}} + o_p(1)
\end{aligned}$$

and

$$\begin{aligned}
\text{(C.14)}\qquad \frac{1}{2K\Delta T} n^{\alpha}\, \mathrm{QV}_{B,K}(\Theta, M) &= \frac{1}{2K}\sum_{l=1}^{2K}\sum_{K\le i\le B-K,\, i\equiv l[2K]} \bigl(\alpha_{T_{i+K}}^{(l,n)} - \alpha_{T_{i-K}}^{(l,n)}\bigr)\bigl(\beta_{T_{i+K}}^{(l,n)} - \beta_{T_{i-K}}^{(l,n)}\bigr) \\
&= \frac{1}{2K}\sum_{l=1}^{2K}\bigl[\alpha^{(l,n)}, \beta^{(l,n)}\bigr]_{\mathcal{T}} + o_p(1)
\end{aligned}$$

$$= \frac{1}{2K} \sum_{l=1}^{2K} \int_0^{\mathcal{T}} g_t^{(l,n)} f_t^{(l,n)} d\big[\theta, \beta^{(n)}\big]_t + o_p(1)$$

$$= \frac{1}{2K} \int_0^{\mathcal{T}} \left(1 - 2\frac{t - t_*}{\Delta T}\right) d\big[\theta, \beta^{(n)}\big]_t + o_p(1)$$

by (C.12). *Q.E.D.*

REMAINING PROOF OF THEOREM 3: Condition 1 implies Condition 5. Equation (22) is the same as requiring that $\sum_i e_{T_i}^2 = o_p(K_n n^{-2\alpha})$ and $\sum_i \tilde{e}_{T_i}^2 = o_p(K_n n^{-2\alpha})$, whence $R_{n,K} = o_p(n^{-2\alpha})$. Expressions (23) and (25) then follow directly from Theorem 8 when assuming Condition 1. This is because of (18) in Proposition 1. For expression (18), we also have invoked the assumption (24). *Q.E.D.*

REMARK 16—AVAR versus AMSE: There are situations of interest when Condition 5 is satisfied, but the additional assumptions of Condition 1 are not. Most notably, consider the situation where $[L, L]_{\mathcal{T}}$ is not $\mathcal{G}$-measurable but instead just integrable. For simplicity, assume that $L_{n,t} = n^{-\alpha} M_{n,t}$ converges in law to L_t relative to the Skorokhod metric on $\mathbb{D}$ (as oppsed to just being tight). In this case, (15) needs to be replaced by

$$\text{(C.15)} \qquad \text{AMSE}(\hat{\Theta} - \Theta) = n^{-2\alpha}[L, L]_{\mathcal{T}} + o_p\big(n^{-2\alpha}\big),$$

where AMSE is the asymptotic mean squared error. This situation arises, for example, in the case of endogenous sampling times for realized volatility (Li, Mykland, Renault, Zhang, and Zheng (2014)). The same phenomenon occurs under direct estimation of skewness (Kinnebrock and Podolskij (2008, Example 6); Mykland and Zhang (2009, Example 3, pp. 1414–1416)).

APPENDIX D: STABLE CONVERGENCE AND OF THE P-UT CONDITION

D.1. *Concepts*

Stable Convergence (Definition 3 in Section 3.1) allows you to take the information from the data (represented by sigma-field $\mathcal{G}$) into the asymptotic distribution. Most commonly, this information is the quadratic variation $[L, L]_{\mathcal{T}}$, which plays the role of variance in the asymptotic distribution, but which can be consistently estimated from the data by any consistent estimator of $n^{2\alpha}[M_n, M_n]_{\mathcal{T}}$. This is the content of Proposition 1.

General conditions for stable convergence to hold can be found in Hall and Heyde (1980), and have a quite general formulation in Jacod and Shiryaev (2003, Theorem VI.6.26, p. 384). Stable convergence of estimators has also been found in countless articles in specific situations, including in high frequency data. See also the book by Jacod and Protter (2012) and the review paper by Podolskij and Vetter (2010).

The amount of data $\mathcal{G}$ that one wishes to carry to asymptopia may vary. The theory described in this paper will work for any $\mathcal{G} \subseteq \mathcal{F}$, so long as $[L, L]_{\mathcal{T}}$ is $\mathcal{G}$-measurable. (This is true under minimal conditions; see Proposition 6 at the end of this section.) One may, however, wish to carry other information. First, for suitably chosen $\mathcal{G}$, stable convergence commutes with measure change (Mykland and Zhang (2009, Proposition 1, p. 1408)), and this can simplify analysis. Second, stable convergence can help weaken conditions with the assistance of *localization*; see, for example, Jacod and Protter (2012, Lemma 4.4.9, pp. 118–121), and Mykland and Zhang (2012, Section 2.4.5, pp. 160–161). In common practice, the information in $\mathcal{G}$ will include latent efficient prices X_t and parameter processes θ_t, but typically not information from the microstructure noise, if present in the model (Zhang, Mykland, and Aït-Sahalia (2005), Zhang (2006), Jacod, Li, Mykland, Podolskij, and Vetter (2009), Podolskij and Vetter (2009), Jacod and Protter (2012), and many others). Thus, $L_{n,t} = n^{\alpha} M_{n,t}$ may in some circumstances not be $\mathcal{G}$-measurable.

For general discussions of stable convergence, see Jacod and Protter (1998, Section 2, pp. 169–170; 2012, Chapter 2.2.1, pp. 46–50), Jacod and Shiryaev (2003, Chapter VIII.5c–d, pp. 512–519), and Mykland and Zhang (2012, Section 2.4, pp. 150–161). For further background on stable convergence, see Rényi (1963), Aldous and Eagleson (1978), Rootzén (1980), and Zhang (2001). Stable convergence was originally thought of as a form of conditional convergence (Jacod and Shiryaev (2003, top of p. 513)).

REMARK 17: In this paper, convergence in law for processes is relative to the Skorokhod topology on the space $\mathbb{D} = \mathbb{D}[0, \mathcal{T}]$ of càdlàg functions $[0, \mathcal{T}] \to \mathbb{R}$. In Definition 3, the pair (L_n, Y) converges in the product topology. In other words, $(L_n, Y) \xrightarrow{\mathcal{L}} (L, Y)$ means that $Ef(L_n)g(Y) \to Ef(L)g(Y)$, for all bounded continuous $f : \mathbb{D} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$. For more on the Skorokhod topology, see Jacod and Shiryaev (2003, Chapter VI.1–2, pp. 325–346). Note that $\mathcal{F}_t$ can depend on n; cf. the discretization discussion in Section 6 and Appendix D.2.

The Predictably Uniformly Tight (P-UT) Condition was described and studied in Jacod and Shiryaev (2003, Chapter VI.6, pp. 377–388). It is an additional regularity condition which avoids certain idiosyncrasies associated with regular process convergence. If the sequence of semimartingales L_n is tight in the Skorokhod topology, one can take as definition of P-UT that if H_n is a bounded family of predictable processes, then $\int_0^T H_{n,t}\,dL_{n,t}$ is tight for each T (ibid., Definition 6.1, p. 377, and Corollary 6.20, p. 381). Also, by ibid., Theorem VI.6.22 (p. 383), if $(H_{n,+}, L_n) \xrightarrow{\mathcal{L}} (H, L)$ (and subject to regularity conditions), then $\int H_{n,t}\,dL_{n,t} \xrightarrow{\mathcal{L}} \int H_t\,dL_t$. Also, and this is important for the current paper, $[L_n, L_n] \xrightarrow{\mathcal{L}} [L, L]$ (ibid., Theorem VI.6.26, p. 384). Finally, P-UT prevents the predictable finite variation part of L_n from turning into a different type of process (ibid., Theorem 6.15 (iii), p. 380, and Theorem VI.6.21, p. 382).

We have seen in Sections 6 and 7 that there is little additional burden in verifying the P-UT condition once one proves stable convergence. Also, a sufficient condition for a sequence of local martingales $L_{n,t}$ to be P-UT is that (Jacod and Shiryaev (2003, Corollary VI.6.30, p. 385))

$$\sup_n E \sup_{0 \le t \le \mathcal{T}} |\Delta L_{n,t}| < \infty. \tag{D.1}$$

The condition (D.1) is weaker than what is usually required for a central limit theorem,[44] and it does in particular not impose asymptotic negligibility. If (D.1) still seems too strong, the requirement can be localized using stable convergence, as described above in this section.

The following is an illustration of how stable convergence blends with P-UT.

[44]See, for example, Hall and Heyde (1980, conditions (3.18) and (3.20), p. 58).

PROOF OF PROPOSITION 1: Let $(\mathcal{F}_t^L)$ be the filtration generated by the process L_t, on the extension $(\tilde{\Omega}, \tilde{\mathcal{G}}, \tilde{P})$. Since, by assumption, L_t is a local martingale with respect to filtration $(\mathcal{G} \vee \mathcal{F}_t^L)$, then it follows that $L_t^2 - [L, L]_t$ is also a local martingale w.r.t. $\mathcal{G} \vee \mathcal{F}_t^L$, and hence $E(L_\mathcal{T} \mid \mathcal{G}) = 0$ and $E(L_\mathcal{T}^2 - [L, L]_\mathcal{T} \mid \mathcal{G}) = 0$. Hence, $\mathrm{Var}(L_\mathcal{T} | \mathcal{G}) = [L, L]_\mathcal{T}$. Set $L_{n,t} = n^\alpha M_{n,t}$. Since $L_{n,t}$ is P-UT, Jacod and Shiryaev (2003, Proposition VI.2.1, p. 377 and Theorem VI.6.26, p. 384) yields that $[L_n, L_n]_\mathcal{T} \xrightarrow{\mathcal{L}} [L, L]_\mathcal{T}$ stably in law as $n \to \infty$. However, since $[L, L]_\mathcal{T}$ is $\mathcal{G}$-measurable and hence defined on the original space, $[L_n, L_n] \xrightarrow{p} [L, L]_\mathcal{T}$ by Jacod and Protter (2012, eq. (2.2.7), p. 47). (It is enough for the "$\Leftarrow$" part of the cited result that the limiting random variable be $\mathcal{G}$-measurable.) *Q.E.D.*

We finish with a result on minimal stable convergence.[45]

PROPOSITION 6—Automatic Minimal Stable Convergence: *Assume that the sequence of semimartingales $L_n = n^\alpha M_n$ converges in law to L, and is P-UT. Also assume that $[L_n, L_n]_\mathcal{T}$ converges in probability. Call this limit V (so $[L_n, L_n]_\mathcal{T} \xrightarrow{p} V$). Let $\mathcal{G}$ be the sigma-field generated by V. Then there is an extension $(\tilde{\Omega}, \tilde{\mathcal{G}}, \tilde{P})$ of $(\Omega, \mathcal{G}, P)$ so that L_n converges stably in law with respect to $\mathcal{G}$ as $n \to \infty$. Also, on this extension, $[L, L]_\mathcal{T} = V$, and $\mathcal{F}_\mathcal{T}^L$ is conditionally independent of $\mathcal{F}$ given $\mathcal{G}$.*

D.2. *Proofs of Propositions 4, 5, and 6*

PROOF OF PROPOSITION 5 IN SECTION 6: The only modification that is required in our proofs is to replace our parameter process by $\theta_{n,t} = \theta_{t_{n,i}}$ for $t_{n,i} \le t < t_{n,i+1}$. Since (the original $(\mathcal{F}_t)$ adapted) θ_t is a semimartingale, then so is $\theta_{n,t}$. Also, $\theta_{n,t}$ converges in probability to θ_t in the Skorokhod topology (Jacod and Shiryaev (2003, Proposition VI.6.37, p. 387)) (and hence also in law). Also, $\theta_{n,t}$ is P-UT (ibid., Definition VI.6.1, p. 377) since the relevant predictable functions on filtration $\mathcal{F}_{t_{n,i}}$ are a subset of the corresponding predictable functions on filtration $\mathcal{F}_t$.

For example, the proof of Theorem 1 in Appendix B goes through with $\theta_{n,t}$ in lieu of θ_t, because Theorem 7 in Appendix A allows time varying $\theta_{n,t}$. The times $T_{n,i}$ are not changed in derivations that do not involve microstructure noise.

Arguments involving only $(e_{n,T_{n,i}}, \tilde{e}_{n,T_{n,i}})$ are directly converted to $(e_{n,T_{n,i,*}}, \tilde{e}_{n,T_{n,i,*}})$. *Q.E.D.*

PROOF OF PROPOSITION 4 IN SECTION 5.2: This is a corollary to Proposition 5. If Condition 1 is valid (in its original form) for $M_{n,t}$, it certainly also holds when discretized as in Condition 3, again using Jacod and Shiryaev (2003, Proposition VI.6.37, p. 387). This shows the result. *Q.E.D.*

PROOF OF PROPOSITION 6: Let $f : \mathbb{D} \to \mathbb{R}$ and $g : \mathbb{R} \to \mathbb{R}$ be bounded and continuous. Since $L_{n,t}$ is P-UT, Jacod and Shiryaev (2003, Proposition VI.2.1, p. 377 and Theorem VI.6.26, p. 384) yields that $(L_n, [L_n, L_n]_\mathcal{T}) \xrightarrow{\mathcal{L}} (L, [L, L]_\mathcal{T})$ (in the non-stable sense), that is, $Ef(L_n)g([L_n, L_n]_\mathcal{T}) = Ef(L)g([L, L]_\mathcal{T}) + o(1)$. On the other hand, by the assumed convergence in probability, $|Ef(L_n)g([L_n, L_n]_\mathcal{T}) - Ef(L_n)g(V)| \le \sup_x |f(x)| E|g([L_n, L_n]_\mathcal{T}) - g(V)| \to 0$.

[45]The following proposition is conceptually related to Hall and Heyde (1980, condition (3.19), p. 58).

We now construct our extension as in Jacod and Protter (2012, p. 36): $\tilde{\Omega} = \Omega \times \mathbb{D}[0, T]$ with product sigma-field, where the sigma-field on $\mathbb{D}[0, T]$ is derived from the Skorokhod topology (Jacod and Shiryaev (2003, Theorem VI.1.14c, p. 328)). The transition probability is given as the regular conditional probability $Q(L|V)$ (Ash (1972, Theorem 6.6.5, p. 265)), where Q is defined as the joint distribution of $(L, [L, L]_\mathcal{T})$ on $\mathbb{D}[0, T] \times \mathbb{R}$ (with corresponding product sigma-field).

With these definitions, $[L, L]_\mathcal{T} = V$, and hence, from the above,

$$\begin{aligned} Ef(L)g(V) &= Ef(L)g\big([L, L]_\mathcal{T}\big) = Ef(L_n)g\big([L_n, L_n]_\mathcal{T}\big) + o(1) \\ &= Ef(L_n)g(V) + o(1) \quad \text{as} \quad n \to \infty. \end{aligned}$$

Hence, the stable convergence follows. The remaining statements of the proposition hold by construction. *Q.E.D.*

D.3. *P-UT Property in Example 2 in Section 7*

P-UT Property for $n^{1/2}M_n$. We make the following assumptions: (i) μ_t is locally integrable and σ_t^2 is continuous,[46] and (ii)

$$\text{(D.2)} \qquad \sum_{j=1}^{n} |\Delta J_{t_{j-1}}||\Delta J_{t_j}| = O_p\big(n^{-1/2}\big).$$

In particular, (D.2) is satisfied when $J_t = J_t^{(1)} + J_t^{(2)}$, where $J^{(1)}$ has finitely many jumps and $J^{(2)}$ is a purely discontinuous Itô-semimartingale (see, e.g., Jacod and Protter (2012, Definition 2.1.1, p. 35; see also Theorem 2.1.2, p. 37)).

PROOF OF P-UT PROPERTY: Without changing either assumptions or conclusions, we absorb the $\mu_t\,dt$ term into dJ_t, so that $dX_t = \sigma_t\,dW_t + dJ_t$. $[J, J]_t$ is unchanged, and so is the statement (D.2). From (D.2) as well as Jacod and Shiryaev (2003, Definition VI.6.1 and the additivity VI.6.4, both p. 377), it follows that to verify P-UT of the original M_n, it is enough that the P-UT property holds on a modified $\tilde{M}_n$ which has the same form as (53) but with X replaced by X^c, where $dX_t^c = \sigma_t\,dW_t$. For this process, it is easy to verify P-UT under the contiguous sequence of measures Q_n from Mykland and Zhang (2009, Section 3, pp. 1416–1421), and using the big block–small block device (Mykland, Shephard, and Sheppard (2012, Appendix A.5, pp. 32–33)), again using Definition VI.6.1 from Jacod and Shiryaev (2003). But this definition is invariant to contiguous change of measure, and hence $\tilde{M}_n$ is P-UT under the original measure P. It follows that the original $n^{1/2}M_n$ is P-UT. *Q.E.D.*

REFERENCES

ALDOUS, D. J., AND G. K. EAGLESON (1978): "On Mixing and Stability of Limit Theorems," *The Annals of Probability*, 6, 325–331. [11]

ASH, R. B. (1972): *Real Analysis and Probability*. New York: Academic Press. [13]

HALL, P., AND C. C. HEYDE (1980): *Martingale Limit Theory and Its Application*. Boston: Academic Press. [10-12]

[46]The spot volatility is also a semimartingale since $\theta_t = \sigma_t^2$. The continuity assumption is merely for convenience and can be reduced to an assumption that σ_t^2 be locally bounded.

JACOD, J., AND P. PROTTER (1998): "Asymptotic Error Distributions for the Euler Method for Stochastic Differential Equations," *The Annals of Probability*, 26, 267–307. [11]

——— (2012): *Discretization of Processes* (First Ed.). New York: Springer-Verlag. [10-13]

JACOD, J., AND A. N. SHIRYAEV (2003): *Limit Theorems for Stochastic Processes* (Second Ed.). New York: Springer-Verlag. [1-5,10-13]

JACOD, J., Y. LI, P. A. MYKLAND, M. PODOLSKIJ, AND M. VETTER (2009): "Microstructure Noise in the Continuous Case: The Pre-Averaging Approach," *Stochastic Processes and Their Applications*, 119, 2249–2276. [11]

KINNEBROCK, S., AND M. PODOLSKIJ (2008): "A Note on the Central Limit Theorem for Bipower Variation of General Functions," *Stochastic Processes and Their Applications*, 118, 1056–1070. [10]

LI, Y., P. A. MYKLAND, E. RENAULT, L. ZHANG, AND X. ZHENG (2014): "Realized Volatility When Sampling Times Are Possibly Endogenous," *Econometric Theory*, 30, 580–605. [10]

MYKLAND, P. A., AND L. ZHANG (2009): "Inference for Continuous Semimartingales Observed at High Frequency," *Econometrica*, 77, 1403–1455. [10,13]

——— (2012): "The Econometrics of High Frequency Data," in *Statistical Methods for Stochastic Differential Equations*, ed. by M. Kessler, A. Lindner, and M. Sørensen. New York: Chapman and Hall/CRC Press, 109–190. [10,11]

MYKLAND, P. A., N. SHEPHARD, AND K. SHEPPARD (2012): "Efficient and Feasible Inference for the Components of Financial Variation Using Blocked Multipower Variation," Technical Report, University of Oxford, http://dx.doi.org/10.2139/ssrn.2008690. [13]

PODOLSKIJ, M., AND M. VETTER (2009): "Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps," *Bernoulli*, 15, 634–658. [11]

——— (2010): "Understanding Limit Theorems for Semimartingales: A Short Survey," *Statistica Neerlandica*, 64, 329–351. [10]

POLLARD, D. (1984): *Convergence of Stochastic Processes*. New York: Springer-Verlag. [1]

RÉNYI, A. (1963): "On Stable Sequences of Events," *Sankyā Series A*, 25, 293–302. [11]

ROOTZÉN, H. (1980): "Limit Distributions for the Error in Approximations of Stochastic Integrals," *The Annals of Probability*, 8, 241–251. [11]

ZHANG, L. (2001): "From Martingales to ANOVA: Implied and Realized Volatility," Ph.D. Thesis, The University of Chicago, Department of Statistics. [11]

——— (2006): "Efficient Estimation of Stochastic Volatility Using Noisy Observations: A Multi-Scale Approach," *Bernoulli*, 12, 1019–1043. [11]

ZHANG, L., P. A. MYKLAND, AND Y. AÏT-SAHALIA (2005): "A Tale of Two Time Scales: Determining Integrated Volatility With Noisy High-Frequency Data," *Journal of the American Statistical Association*, 100, 1394–1411. [11]

Dept. of Statistics, The University of Chicago, 5747 South Ellis Avenue, Chicago, IL 60637, U.S.A.; mykland@pascal.uchicago.edu

and

Department of Finance, University of Illinois at Chicago, 601 South Morgan Street, mc 168, Chicago, IL 60607, U.S.A.; lanzhang@uic.edu.

Co-editor Elie Tamer handled this manuscript.

Manuscript received 12 May, 2014; final version accepted 17 September, 2016; available online 28 September, 2016.