Supplement to

Asymptotic Inference about Predictive Accuracy using High Frequency Data*

Jia Li
Department of Economics
Duke University

Andrew J. Patton
Department of Economics
Duke University

This Version: July 6, 2013

Abstract

This supplement contains three appendices. Appendix S.A contains proofs of results in the main text. Appendix S.B provides details for the stepwise procedures discussed in Section 4 of the main text. Appendix S.C contains some additional simulation results.

*Contact address: Department of Economics, Duke University, 213 Social Sciences Building, Box 90097, Durham NC 27708-0097, USA. Email: jl410@duke.edu and andrew.patton@duke.edu.

Appendix S.A Proofs of main results

In this appendix, we prove the results in the main text. Results in Sections 2 and 3 are proved in Appendices S.A.1 and S.A.2, respectively. Technical lemmas that are used in these proofs are proved in Appendix S.A.3. Below, we use K to denote a generic constant, which may change from line to line but does not depend on t.

S.A.1 Proofs in Section 2

PROOF OF PROPOSITION 2.1. (a) Under H_0, $\mathbb{E}[\bar{f}_T^*] = \chi$. By Assumption A1, $(a_T(\bar{f}_T - \chi), a_T' S_T) \xrightarrow{d} (\xi, S)$. By the continuous mapping theorem and Assumption A2, $\varphi_T \xrightarrow{d} \varphi(\xi, S)$. By Assumption A3, $\mathbb{E}\phi_T \to \alpha$.

Now consider H_{1a}, so Assumption B1(b) is in force. Under H_{1a}, the nonrandom sequence $a_T(\mathbb{E}[\bar{f}_{j,T}^*] - \chi_j)$ diverges to $+\infty$. Hence, by Assumption A1 and Assumption B1(b), φ_T diverges to $+\infty$ in probability. (To see this, one can use the almost sure representation of the weak convergence in Assumption A1, and then show the pathwise divergence towards $+\infty$ by using Assumption B1(b).) Since the law of (ξ, S) is tight, the law of $\varphi(\xi, S)$ is also tight by Assumption A2. Therefore, $z_{T,1-\alpha} = O_p(1)$. It is then easy to see $\mathbb{E}\phi_T \to 1$ under H_{1a}.

The case with H_{2a} can be proved similarly.

(b) Under H_0, $a_T(\bar{f}_T - \chi) \leq a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^*])$. Let $\tilde{\phi}_T = \mathbf{1}\{\varphi\left(a_T(\bar{f}_T - \mathbb{E}\left[\bar{f}_T^*\right]), a_T' S_T\right) > z_{T,1-\alpha}\}$. By monotonicity (Assumption B1(a)), $\phi_T \leq \tilde{\phi}_T$. Following a similar argument as in part (a), $\mathbb{E}\tilde{\phi}_T \to \alpha$. Then $\limsup_{T\to\infty} \mathbb{E}\phi_T \leq \alpha$ readily follows. The case under H_a follows a similar argument as in part (a). *Q.E.D.*

PROOF OF THEOREM 2.1. By Assumptions A1 and C1, $(a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^\dagger]), a_T' S_T) \xrightarrow{d} (\xi, S)$. With this convergence replacing that in Assumption A1, we use the same argument as in Proposition 2.1 to prove Theorem 2.1. The details are omitted. *Q.E.D.*

PROOF OF LEMMA 2.1. We observe that

$$
\begin{aligned}
a_T \|\mathbb{E}[\bar{f}_T^*] - \mathbb{E}[\bar{f}_T^\dagger]\| &\leq (a_T/P) \sum_{t=R}^{T} \mathbb{E}\|f_{t+\tau}^* - f_{t+\tau}^\dagger\| \\
&\leq (a_T/P) \sum_{t=R}^{T} \mathbb{E}\left[m_{t+\tau}\left\|Y_{t+\tau} - Y_{t+\tau}^\dagger\right\|\right] \\
&\leq K(a_T/P) \sum_{t=R}^{T} \|m_{t+\tau}\|_{p/(p-1)} \left\|Y_{t+\tau} - Y_{t+\tau}^\dagger\right\|_p \\
&\leq K(a_T/P) \sum_{t=R}^{T} d_{t+\tau}^\theta \to 0,
\end{aligned}
$$

where the first inequality is due to the triangle inequality; the second inequality is by Assumption C3(a); the third inequality is by Hölder's inequality; the fourth inequality is by Assumptions C2 and C3(a); the convergence follows from Assumption C3(b). Hence, $a_T(\mathbb{E}[\bar{f}_T^*] - \mathbb{E}[\bar{f}_T^\dagger]) \to 0$ as claimed. *Q.E.D.*

S.A.2 Proofs in Section 3

Throughout this section, we denote

$$X_t' = X_0 + \int_0^t b_s' ds + \int_0^t \sigma_s dW_s, \quad X_t'' = X_t - X_t', \tag{S.A.1}$$

where the process b_s' is defined in Assumption HF(d). Below, for any process Z, we denote the ith return of Z in day t by $\Delta_{t,i} Z = Z_{\tau(t,i)} - Z_{\tau(t,i-1)}$.

PROOF OF PROPOSITION 3.1. Denote $\beta_{t,i} = \sigma_{\tau(t,i-1)} \Delta_{t,i} W / d_{t,i}^{1/2}$. Observe that for $m = 2/p$ and $m' = 2/(2-p)$,

$$\begin{aligned} &\mathbb{E} \left| g(\Delta_{t,i} X / d_{t,i}^{1/2}) - g(\beta_{t,i}) \right|^p \\ &\quad \le K \mathbb{E} \left[\left(1 + \|\beta_{t,i}\|^{pq} + \|\Delta_{t,i} X / d_{t,i}^{1/2}\|^{pq} \right) \|\Delta_{t,i} X / d_{t,i}^{1/2} - \beta_{t,i}\|^p \right] \\ &\quad \le K \left(\mathbb{E} \left[\left(1 + \|\beta_{t,i}\|^{pqm'} + \|\Delta_{t,i} X / d_{t,i}^{1/2}\|^{pqm'} \right) \right] \right)^{1/m'} \left(\mathbb{E} \|\Delta_{t,i} X / d_{t,i}^{1/2} - \beta_{t,i}\|^{pm} \right)^{1/m} \\ &\quad \le K d_{t,i}^{p/2}, \end{aligned}$$

where the first inequality follows the mean-value theorem, the Cauchy-Schwarz inequality and condition (ii); the second inequality is due to Hölder's inequality; the third inequality holds because of condition (iii) and $\mathbb{E}\|\Delta_{t,i} X / d_{t,i}^{1/2} - \beta_{t,i}\|^2 \le K d_{t,i}$. Hence, $\|g(\Delta_{t,i} X / d_{t,i}^{1/2}) - g(\beta_{t,i})\|_p \le K d_{t,i}^{1/2}$, which further implies

$$\left\| \widehat{\mathcal{I}}_t(g) - \sum_{i=1}^{n_t} g\left(\beta_{t,i}\right) d_{t,i} \right\|_p \le K d_t^{1/2}. \tag{S.A.2}$$

Below, we write $\rho(\cdot)$ in place of $\rho(\,\cdot\,; g)$ for the sake of notational simplicity. Let $\zeta_{t,i} = g(\beta_{t,i}) - \rho(c_{\tau(t,i-1)})$. By construction, $\zeta_{t,i}$ forms a martingale difference sequence. By condition (iv), for all i, $\mathbb{E}[(\zeta_{t,i})^2] \le \mathbb{E}[\rho(c_{\tau(t,i-1)}; g^2)] \le K$. Hence, $\mathbb{E}|\sum_{i=1}^{n_t} \zeta_{t,i} d_{t,i}|^2 = \sum_{i=1}^{n_t} \mathbb{E}[(\zeta_{t,i})^2] d_{t,i}^2 \le K d_t$, yielding

$$\left\| \sum_{i=1}^{n_t} \zeta_{t,i} d_{t,i} \right\|_p \le \left\| \sum_{i=1}^{n_t} \zeta_{t,i} d_{t,i} \right\|_2 \le K d_t^{1/2}. \tag{S.A.3}$$

In view of (S.A.2) and (S.A.3), it remains to show

$$\left\| \int_{t-1}^{t} \rho(c_s) ds - \sum_{i=1}^{n_t} \rho\left(c_{\tau(t,i-1)}\right) d_{t,i} \right\|_p \leq K d_t^{1/2}. \tag{S.A.4}$$

First note that

$$\int_{t-1}^{t} \rho\left(c_s\right) ds - \sum_{i=1}^{n_t} \rho\left(c_{\tau(t,i-1)}\right) d_{t,i} = \sum_{i=1}^{n_t} \int_{\tau(t,i-1)}^{\tau(t,i)} \left(\rho\left(c_s\right) - \rho\left(c_{\tau(t,i-1)}\right)\right) ds.$$

We then observe that, for all $s \in [\tau(t, i-1), \tau(t, i)]$ and with $m = 2/p$ and $m' = 2/\left(2-p\right)$,

$$\begin{aligned} &\left\|\rho\left(c_s\right) - \rho\left(c_{\tau(t,i-1)}\right)\right\|_p \\ &\leq K\left(\mathbb{E}\left[\left(1 + \|c_s\|^{pq/2} + \|c_{\tau(t,i-1)}\|^{pq/2}\right)\|c_s - c_{\tau(t,i-1)}\|^p\right]\right)^{1/p} \\ &\leq K\left(\mathbb{E}\left[\left(1 + \|\sigma_s\|^{pqm'} + \|\sigma_{\tau(t,i-1)}\|^{pqm'}\right)\right]\right)^{1/pm'}\left(\mathbb{E}\|c_s - c_{\tau(t,i-1)}\|^{pm}\right)^{1/pm} \\ &\leq K d_{t,i}^{1/2}, \end{aligned}$$

where the first inequality follows from the mean-value theorem, the Cauchy-Schwarz inequality and condition (ii); the second inequality is due to Hölder's inequality; the third inequality follows from condition (iii) and the standard estimate $\mathbb{E}\|c_s - c_{\tau(t,i-1)}\|^2 \leq K d_{t,i}$. From here (S.A.4) follows. This finishes the proof. *Q.E.D.*

Proof of Proposition 3.2. Step 1. For $x, y \in \mathbb{R}^d$, we set

$$\begin{cases} k\left(y,x\right) = g\left(y+x\right) - g\left(x\right) - g\left(y\right) \\ h\left(y,x\right) = g\left(y+x\right) - g\left(x\right) - g\left(y\right) - \partial g\left(y\right)^{\intercal} x 1_{\{\|x\| \leq 1\}}. \end{cases} \tag{S.A.5}$$

By Taylor's theorem and condition (ii),

$$\begin{cases} \left|k\left(y,x\right)\right| \leq K \sum_{j=1}^{2} \left(\|y\|^{q_j-1}\|x\| + \|x\|^{q_j-1}\|y\|\right), \\ \left|h\left(y,x\right)\right| \leq K \sum_{j=1}^{2} \left(\|y\|^{q_j-2}\|x\|^2 + \|x\|^{q_j-1}\|y\| + \|y\|^{q_j-1}\|x\| 1_{\{\|x\|>1\}}\right). \end{cases} \tag{S.A.6}$$

We consider a process $(Z_s)_{s\in[t-1,t]}$ that is given by $Z_s = X_s - X_{\tau(t,i-1)}$ when $s \in [\tau(t,i-1), \tau(t,i))$. We define Z'_s similarly but with X' replacing X; recall that X' is defined in (S.A.1). We then set

$Z''_s = Z_s - Z'_s$. Under Assumption HF, we have

$$\begin{cases} v \in [0,k] \Rightarrow \mathbb{E}[\sup_{s\in[\tau(t,i-1),\tau(t,i)]} \|Z'_s\|^v] \leq K d_{t,i}^{v/2}, \\ v \in [2,k] \Rightarrow \mathbb{E}[\sup_{s\in[\tau(t,i-1),\tau(t,i)]} \|Z''_s\|^v | \mathcal{F}_{\tau(t,i-1)}] \leq K d_{t,i}, \end{cases} \tag{S.A.7}$$

where the first line follows from a classical estimate for continuous Itô semimartingales, and the second line is derived by using Lemmas 2.1.5 and 2.1.7 in Jacod and Protter (2012).

By Itô's formula, we decompose

$$\begin{aligned} &\widehat{\mathcal{J}}_t(g) - \mathcal{J}_t(g) \\ &\quad = \int_{t-1}^{t} \partial g(Z_{s-})^\intercal b_s ds + \frac{1}{2}\sum_{j,l=1}^{d}\int_{t-1}^{t} \partial^2_{j,l} g(Z_{s-}) c_{jl,s} ds \\ &\quad\quad + \int_{t-1}^{t} ds \int_{\mathbb{R}} h(Z_{s-}, \delta(s,z))\lambda(dz) + \int_{t-1}^{t} \partial g(Z_{s-})^\intercal \sigma_s dW_s \\ &\quad\quad + \int_{t-1}^{t}\int_{\mathbb{R}} k(Z_{s-},\delta(s,z))\tilde{\mu}(ds,dz). \end{aligned} \tag{S.A.8}$$

Below, we study each component in the above decomposition separately.

Step 2. In this step, we show

$$\left\| \int_{t-1}^{t} \partial g(Z_{s-})^\intercal b_s ds \right\|_p \leq K d_t^{1/2}. \tag{S.A.9}$$

Let $m = 2/p$ and $m' = 2/(2-p)$. Observe that, for all $s \in [\tau(t,i-1),\tau(t,i)]$,

$$\begin{aligned} \|\partial g(Z_{s-})^\intercal b_s\|_p &\leq K\left(\mathbb{E}\left|\sum_{j=1}^{2}\|Z_{s-}\|^{q_j-1}\|b_s\|\right|^p\right)^{1/p} \\ &\leq K\sum_{j=1}^{2}\left(\mathbb{E}\|Z_{s-}\|^{(q_j-1)pm}\right)^{1/pm}\left(\mathbb{E}\|b_s\|^{pm'}\right)^{1/pm'} \\ &\leq K\sum_{j=1}^{2}\left(\mathbb{E}\|Z_{s-}\|^{2(q_j-1)}\right)^{1/2}\left(\mathbb{E}\|b_s\|^{pm'}\right)^{1/pm'} \\ &\leq K d_{t,i}^{1/2}, \end{aligned}$$

where the first inequality is due to condition (ii) and the Cauchy-Schwarz inequality; the second inequality is due to Hölder's inequality; the third inequality follows from our choice of m; the last inequality follows from (S.A.7). The claim (S.A.9) then readily follows.

Step 3. In this step, we show

$$\left\| \frac{1}{2} \sum_{j,l=1}^{d} \int_{t-1}^{t} \partial_{j,l}^{2} g\left(Z_{s-}\right) c_{jl,s} ds \right\|_{p} \leq K d_{t}^{1/2}. \tag{S.A.10}$$

By a component-wise argument, we can assume that $d = 1$ without loss of generality and suppress the component subscripts in our notation below. Let $m' = 2/(2-p)$. We observe

$$\begin{aligned} \left\|\partial^{2} g\left(Z_{s-}\right) c_{s}\right\|_{p} &\leq K \sum_{j=1}^{2} \left(\mathbb{E}\left[\left|Z_{s-}\right|^{p(q_{j}-2)} \left|c_{s}\right|^{p}\right]\right)^{1/p} \\ &\leq K \sum_{j=1}^{2} (\mathbb{E}\left|Z_{s-}\right|^{2(q_{j}-2)})^{1/2} \left(\mathbb{E}\left|c_{s}\right|^{pm'}\right)^{1/pm'} \\ &\leq K d_{t,i}^{1/2}, \end{aligned}$$

where the first inequality follows from condition (ii); the second inequality is due to Hölder's inequality and our choice of m'; the last inequality follows from (S.A.7). The claim (S.A.10) is then obvious.

Step 4. In this step, we show

$$\left\| \int_{t-1}^{t} ds \int_{\mathbb{R}} h\left(Z_{s-}, \delta\left(s, z\right)\right) \lambda\left(dz\right) \right\|_{p} \leq K d_{t}^{1/2}. \tag{S.A.11}$$

By (S.A.6) and $\|\delta\left(s,z\right)\| \leq \Gamma\left(z\right)$,

$$\left|h\left(Z_{s-}, \delta\left(s, z\right)\right)\right| \leq K \sum_{j=1}^{2} \left(\left\|Z_{s-}\right\|^{q_{j}-2} \Gamma\left(z\right)^{2} + \Gamma\left(z\right)^{q_{j}-1} \left\|Z_{s-}\right\| + \left\|Z_{s-}\right\|^{q_{j}-1} \Gamma\left(z\right) 1_{\{\Gamma(z)>1\}} \right).$$

Hence, by condition (iii),

$$\left| \int_{\mathbb{R}} h\left(Z_{s-}, \delta\left(s, z\right)\right) \lambda\left(dz\right) \right| \leq K \sum_{j=1}^{2} \left(\left\|Z_{s-}\right\|^{q_{j}-2} + \left\|Z_{s-}\right\| + \left\|Z_{s-}\right\|^{q_{j}-1} \right).$$

By (S.A.7), for any $s \in [\tau(t,i-1), \tau(t,i)]$,

$$
\begin{aligned}
&\left\| \int_{\mathbb{R}} h\left(Z_{s-}, \delta\left(s,z\right)\right) \lambda\left(dz\right) \right\|_2 \\
&\leq K \sum_{j=1}^{2} \left(\left(\mathbb{E} \left\| Z_{s-} \right\|^{2(q_j-2)} \right)^{1/2} + \left(\mathbb{E} \left\| Z_{s-} \right\|^{2} \right)^{1/2} + \left(\mathbb{E} \left\| Z_{s-} \right\|^{2(q_j-1)} \right)^{1/2} \right) \\
&\leq K d_{t,i}^{1/2}.
\end{aligned}
$$

The claim (S.A.11) then readily follows.

Step 5. In this step, we show

$$
\left\| \int_{t-1}^{t} \partial g\left(Z_{s-}\right)^{\intercal} \sigma_s dW_s \right\|_2 \leq K d_t^{1/2}. \tag{S.A.12}
$$

By the Burkholder–Davis–Gundy inequality,

$$
\begin{aligned}
&\mathbb{E} \left| \int_{t-1}^{t} \partial g\left(Z_{s-}\right)^{\intercal} \sigma_s dW_s \right|^2 \\
&\leq K \mathbb{E} \left[\int_{t-1}^{t} \left\| \partial g\left(Z_s\right) \right\|^2 \left\| \sigma_s \right\|^2 ds \right] \\
&\leq K \mathbb{E} \left[\int_{t-1}^{t} \left\| \partial g\left(Z_s'\right) \right\|^2 \left\| \sigma_s \right\|^2 ds \right] \\
&\quad + K \mathbb{E} \left[\sum_{i=1}^{n_t} \int_{\tau(t,i-1)}^{\tau(t,i)} \left\| \partial g\left(Z_s\right) - \partial g\left(Z_s'\right) \right\|^2 \left\| \sigma_{\tau(t,i-1)} \right\|^2 ds \right] \\
&\quad + K \mathbb{E} \left[\sum_{i=1}^{n_t} \int_{\tau(t,i-1)}^{\tau(t,i)} \left\| \partial g\left(Z_s\right) - \partial g\left(Z_s'\right) \right\|^2 \left\| \sigma_s - \sigma_{\tau(t,i-1)} \right\|^2 ds \right].
\end{aligned} \tag{S.A.13}
$$

We first consider the first term on the majorant side of (S.A.13). By Hölder's inequality, we have, for $s \in [\tau(t,i-1), \tau(t,i)]$,

$$
\mathbb{E} \left[\left\| \partial g\left(Z_s'\right) \right\|^2 \left\| \sigma_s \right\|^2 \right] \leq K \sum_{j=1}^{2} \left(\mathbb{E} \left\| Z_s' \right\|^{2q_j} \right)^{(q_j-1)/q_j} \left(\mathbb{E} \left\| \sigma_s \right\|^{2q_j} \right)^{1/q_j} \leq K d_t^{q_1-1},
$$

where the second inequality is due to (S.A.7). This estimate implies

$$
\mathbb{E} \left[\int_{t-1}^{t} \left\| \partial g\left(Z_s'\right) \right\|^2 \left\| \sigma_s \right\|^2 ds \right] \leq K d_t. \tag{S.A.14}
$$

Now turn to the second term on the majorant side of (S.A.13). Observe that

$$\begin{aligned}
&\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|\partial g\left(Z_s\right)-\partial g\left(Z_s'\right)\right\|^2\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right]\\
&\quad\leq K\sum_{j=1}^{2}\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left(\left\|Z_s'\right\|^{2(q_j-2)}+\left\|Z_s''\right\|^{2(q_j-2)}\right)\left\|Z_s''\right\|^2\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right]\\
&\quad\leq K\sum_{j=1}^{2}\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|Z_s'\right\|^{2(q_j-2)}\left\|Z_s''\right\|^2\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right]\\
&\qquad+K\sum_{j=1}^{2}\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|Z_s''\right\|^{2(q_j-1)}\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right].
\end{aligned}\tag{S.A.15}$$

By repeated conditioning and (S.A.7), we have

$$\sum_{j=1}^{2}\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|Z_s''\right\|^{2(q_j-1)}\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right]\leq Kd_t. \tag{S.A.16}$$

Moreover, by Hölder's inequality and (S.A.7), for $s\in[\tau(t,i-1),\tau(t,i)]$,

$$\begin{aligned}
&\mathbb{E}\left[\left\|Z_s'\right\|^{2(q_j-2)}\left\|Z_s''\right\|^2\left\|\sigma_{\tau(t,i-1)}\right\|^2\right]\\
&\quad\leq\left(\mathbb{E}\left\|Z_s'\right\|^{2q_j}\right)^{(q_j-2)/q_j}\left(\mathbb{E}\left\|Z_s''\right\|^{2q_j}\right)^{1/q_j}\left(\mathbb{E}\left\|\sigma_{\tau(t,i-1)}\right\|^{2q_j}\right)^{1/q_j}\\
&\quad\leq Kd_{t,i}^{q_j-2}d_{t,i}^{1/q_j}.
\end{aligned}$$

Therefore,

$$\sum_{j=1}^{2}\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|Z_s'\right\|^{2(q_j-2)}\left\|Z_s''\right\|^2\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right]\leq Kd_t. \tag{S.A.17}$$

Combining (S.A.15)–(S.A.17), we have

$$\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|\partial g\left(Z_s\right)-\partial g\left(Z_s'\right)\right\|^2\left\|\sigma_{\tau(t,i-1)}\right\|^2 ds\right]\leq Kd_t. \tag{S.A.18}$$

We now consider the third term on the majorant side of (S.A.13). By the mean-value theorem and condition (ii),

$$\begin{aligned}
&\mathbb{E}\left[\left\|\partial g\left(Z_s\right)-\partial g\left(Z_s'\right)\right\|^2\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2\right] \\
&\quad\leq K\sum_{j=1}^{2}\mathbb{E}\left[\left\|Z_s'\right\|^{2(q_j-2)}\left\|Z_s''\right\|^2\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2\right] \\
&\qquad+K\sum_{j=1}^{2}\mathbb{E}\left[\left\|Z_s''\right\|^{2(q_j-1)}\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2\right].
\end{aligned} \tag{S.A.19}$$

By Hölder's inequality and (S.A.7),

$$\begin{aligned}
&\mathbb{E}\left[\left\|Z_s''\right\|^{2(q_j-1)}\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2\right] \\
&\quad\leq\left(\mathbb{E}\left\|Z_s''\right\|^{2q_j}\right)^{(q_j-1)/q_j}\left(\mathbb{E}\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^{2q_j}\right)^{1/q_j} \\
&\quad\leq Kd_{t,i}^{(q_j-1)/q_j}d_{t,i}^{1/q_j}\leq Kd_{t,i}.
\end{aligned} \tag{S.A.20}$$

Similarly,

$$\begin{aligned}
&\mathbb{E}\left[\left\|Z_s'\right\|^{2(q_j-2)}\left\|Z_s''\right\|^2\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2\right] \\
&\quad\leq\left(\mathbb{E}\left\|Z_s'\right\|^{2q_j}\right)^{(q_j-2)/q_j}\left(\mathbb{E}\left\|Z_s''\right\|^{2q_j}\right)^{1/q_j}\left(\mathbb{E}\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^{2q_j}\right)^{1/q_j} \\
&\quad\leq Kd_{t,i}^{q_j-2}d_{t,i}^{1/q_j}d_{t,i}^{1/q_j}\leq Kd_t.
\end{aligned} \tag{S.A.21}$$

Combining (S.A.19)–(S.A.21), we have

$$\mathbb{E}\left[\left\|\partial g\left(Z_s\right)-\partial g\left(Z_s'\right)\right\|^2\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2\right]\leq Kd_t.$$

Hence,

$$\mathbb{E}\left[\sum_{i=1}^{n_t}\int_{\tau(t,i-1)}^{\tau(t,i)}\left\|\partial g\left(Z_{s-}\right)-\partial g\left(Z_s'\right)\right\|^2\left\|\sigma_s-\sigma_{\tau(t,i-1)}\right\|^2ds\right]\leq Kd_t. \tag{S.A.22}$$

We have shown that each term on the majorant side of (S.A.13) is bounded by Kd_t; see (S.A.14), (S.A.18) and (S.A.22). The estimate (S.A.12) is now obvious.

Step 6. We now show

$$\left\|\int_{t-1}^{t}\int_{\mathbb{R}}k\left(Z_{s-},\delta\left(s,z\right)\right)\tilde{\mu}\left(ds,dz\right)\right\|_2\leq Kd_t^{1/2}. \tag{S.A.23}$$

By Lemma 2.1.5 in Jacod and Protter (2012), (S.A.6) and Assumption HF,

$$\begin{aligned}
&\mathbb{E}\left|\int_{t-1}^{t}\int_{\mathbb{R}} k\left(Z_{s-},\delta\left(s,z\right)\right)\tilde{\mu}\left(ds,dz\right)\right|^{2} \\
&\leq K\sum_{j=1}^{2}\mathbb{E}\left[\int_{t-1}^{t} ds\int_{\mathbb{R}}\left(\|Z_{s-}\|^{q_j-1}\|\delta\left(s,z\right)\|+\|Z_{s-}\|\|\delta\left(s,z\right)\|^{q_j-1}\right)^{2}\lambda\left(dz\right)\right] \\
&\leq K\sum_{j=1}^{2}\mathbb{E}\left[\int_{t-1}^{t} ds\int_{\mathbb{R}}\left(\|Z_{s}\|^{2(q_j-1)}\Gamma\left(z\right)^{2}+\|Z_{s}\|^{2}\Gamma\left(z\right)^{2(q_j-1)}\right)\lambda\left(dz\right)\right] \\
&\leq Kd_t,
\end{aligned}$$

which implies (S.A.23).

Step 7. Combining the estimates in Steps 2–6 with the decomposition (S.A.8), we derive $\|\widehat{\mathcal{J}}_t(g)-\mathcal{J}_t\left(g\right)\|_p \leq Kd_t^{1/2}$ as wanted. *Q.E.D.*

We now turn to the proof of Proposition 3.3. Recalling (S.A.1), we set

$$\hat{c}'_{\tau(t,i)} = \frac{1}{k_t}\sum_{j=1}^{k_t} d_{t,i+j}^{-1}(\Delta_{t,i+j}X')(\Delta_{t,i+j}X')^{\intercal}.$$

The proof of Proposition 3.3 relies on the following technical lemmas that are proved in Appendix S.A.3.

Lemma S.A.1: Let $w \geq 2$ and $v \geq 1$. Suppose (i) Assumption HF holds for some $k \geq 2wv$ and (ii) $k_t \asymp d_t^{-1/2}$ as $t \to \infty$. Then

$$\left\|\mathbb{E}\left[\left\|\hat{c}'_{\tau(t,i)}-c_{\tau(t,i)}\right\|^{w}\middle|\mathcal{F}_{\tau(t,i)}\right]\right\|_{v} \leq \begin{cases} Kd_t^{1/2} & \text{in general,} \\ Kd_t^{w/4} & \text{if } \sigma_t \text{ is continuous.}\end{cases}$$

Lemma S.A.2: Let $w \geq 1$ and $v \geq 1$. Suppose (i) Assumption HF holds for some $k \geq 2wv$ and (ii) $k_t \asymp d_t^{-1/2}$ as $t \to \infty$. Then

$$\left\|\left\|\mathbb{E}\left[\hat{c}'_{\tau(t,i)}-c_{\tau(t,i)}\middle|\mathcal{F}_{\tau(t,i)}\right]\right\|^{w}\right\|_{v} \leq Kd_t^{w/2}.$$

Lemma S.A.3: Let $w \geq 1$. Suppose Assumption HF hold with $k \geq 2w$. We have $\mathbb{E}\|\hat{c}_{\tau(t,i)}-\hat{c}'_{\tau(t,i)}\|^{w} \leq Kd_t^{\bar{\theta}(k,w,\varpi,r)}$, where

$$\begin{aligned}
&\bar{\theta}\left(k,w,\varpi,r\right) \\
&\quad= \min\left\{k/2-\varpi\left(k-2w\right)-w,\right. \\
&\qquad\left. 1-\varpi r+w(2\varpi-1), w(\varpi-1/2)+(1-\varpi r)\min\{w/r,(k-w)/k\}\right\}.
\end{aligned}$$

Proof of Proposition 3.3. Step 1. Throughout the proof, we denote $\mathbb{E}[\,\cdot\,|\mathcal{F}_{\tau(t,i)}]$ by $\mathbb{E}_i[\,\cdot\,]$. Consider the decomposition: $\widehat{\mathcal{I}}_t^\star(g) - \mathcal{I}_t^\star(g) = \sum_{j=1}^4 R_j$, where

$$
\begin{aligned}
R_1 &= \sum_{i=0}^{n_t-k_t} \Big(g(\hat{c}'_{\tau(t,i)}) - g(c_{\tau(t,i)}) - \partial g(c_{\tau(t,i)})^\intercal (\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}) \Big) d_{t,i} \\
R_2 &= \sum_{i=0}^{n_t-k_t} \partial g(c_{\tau(t,i)})^\intercal (\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}) d_{t,i} \\
R_3 &= \sum_{i=0}^{n_t-k_t} g(c_{\tau(t,i)}) d_{t,i} - \int_{t-1}^{t} g(c_s) ds \\
R_4 &= \sum_{i=0}^{n_t-k_t} (g(\hat{c}_{\tau(t,i)}) - g(\hat{c}'_{\tau(t,i)})) d_{t,i};
\end{aligned}
$$

note that in the first two lines of the above display, we have treated $\hat{c}'_{\tau(t,i)}$ and $c_{\tau(t,i)}$ as their vectorized versions so as to simplify notations. In this step, we show that

$$
\|R_1\|_p \leq \begin{cases} K d_t^{1/(2p)} & \text{in general,} \\ K d_t^{1/2} & \text{if } \sigma_t \text{ is continuous.} \end{cases} \tag{S.A.24}
$$

By Taylor's expansion and condition (i),

$$
\begin{aligned}
|R_1| &\leq K \sum_{i=0}^{n_t-k_t} d_{t,i} (1 + \|c_{\tau(t,i)}\|^{q-2} + \|\hat{c}'_{\tau(t,i)}\|^{q-2}) \|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^2 \\
&\leq K \sum_{i=0}^{n_t-k_t} d_{t,i} \Big((1 + \|c_{\tau(t,i)}\|^{q-2}) \|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^2 + \|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^q \Big).
\end{aligned} \tag{S.A.25}
$$

Let $v = q/2$ and $v' = q/(q-2)$. Notice that

$$
\begin{aligned}
&\mathbb{E}\Big[(1 + \|c_{\tau(t,i)}\|^{q-2})^p \|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^{2p} \Big] \\
&\leq K \left\| (1 + \|c_{\tau(t,i)}\|^{q-2})^p \right\|_{v'} \left\| \mathbb{E}_i \|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^{2p} \right\|_{v} \\
&\leq \begin{cases} K d_t^{1/2} & \text{in general,} \\ K d_t^{p/2} & \text{when } \sigma_t \text{ is continuous,} \end{cases}
\end{aligned}
$$

where the first inequality follows from repeated conditioning and Hölder's inequality, and the second inequality is derived by using Lemma S.A.1 with $w = 2p$. Applying Lemma S.A.1 again (with $w = qp$ and $v = 1$), we derive $\mathbb{E}\|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^{qp} \leq K d_t^{1/2}$ and, when σ_t is continuous, the bound can be improved as $K d_t^{qp/4} \leq K d_t^{p/2}$. The claim (S.A.24) then follows from (S.A.25).

Step 2. In this step, we show that

$$\|R_2\|_p \leq K d_t^{1/2}. \tag{S.A.26}$$

Denote $\zeta_i = \partial g(c_{\tau(t,i)})^\intercal(\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)})$, $\zeta'_i = \mathbb{E}_i[\zeta_i]$ and $\zeta''_i = \zeta_i - \zeta'_i$. Notice that $\zeta'_i = \partial g(c_{\tau(t,i)})^\intercal \mathbb{E}_i[\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}]$. By condition (i) and the Cauchy–Schwarz inequality, $|\zeta'_i| \leq K(1 + \|c_{\tau(t,i)}\|^{q-1})\|\mathbb{E}_i[\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}]\|$. Observe that, with $v = q$ and $v' = q/(q-1)$,

$$\begin{aligned}
\mathbb{E}\left|\zeta'_i\right|^p &\leq K\left\|1 + \|c_{\tau(t,i)}\|^{p(q-1)}\right\|_{v'}\left\|\|\mathbb{E}_i[\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}]\|^p\right\|_v \\
&\leq K d_t^{p/2},
\end{aligned}$$

where the first inequality is by Hölder's inequality, and the second inequality is derived by using Lemma S.A.2 (with $w = p$). Hence,

$$\left\|\sum_{i=0}^{n_t-k_t} \zeta'_i d_{t,i}\right\|_p \leq K d_t^{1/2}. \tag{S.A.27}$$

Next consider ζ''_i. First notice that

$$\begin{aligned}
\mathbb{E}\left|\zeta''_i\right|^2 &\leq K\mathbb{E}\left|\zeta_i\right|^2 \\
&\leq K\mathbb{E}\left[(1 + \|c_{\tau(t,i)}\|^{q-1})^2\|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^2\right] \\
&\leq K\left\|1 + \|c_{\tau(t,i)}\|^{2(q-1)}\right\|_{v'}\left\|\mathbb{E}_i\|\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)}\|^2\right\|_v \\
&\leq K d_t^{1/2},
\end{aligned}$$

where the first inequality is obvious; the second inequality follows from condition (i) and the Cauchy–Schwarz inequality; the third inequality is by repeated conditioning and Hölder's inequality; the fourth inequality is derived by applying Lemma S.A.1 (with $w = 2$). Further notice that ζ''_i and ζ''_l are uncorrelated whenever $|i - l| \geq k_t$. By the Cauchy–Schwarz inequality and the above estimate, as well as condition (ii),

$$\mathbb{E}\left|\sum_{i=0}^{n_t-k_t} \zeta''_i d_{t,i}\right|^2 \leq K k_t \sum_{i=0}^{n_t-k_t} \mathbb{E}|\zeta''_i|^2 d_{t,i}^2 \leq K d_t.$$

Therefore, $\|\sum_{i=0}^{n_t-k_t} \zeta''_i d_{t,i}\|_2 \leq K d_t^{1/2}$. This estimate, together with (S.A.27), implies (S.A.26).

Step 3. Consider R_3 in this step. Let $v = 2/p$ and $v' = 2/(2-p)$. Notice that for $s \in$

$[\tau(t,i-1),\tau(t,i)]$,

$$\begin{aligned}\mathbb{E}|g(c_s)-g(c_{\tau(t,i-1)})|^p &\leq K\mathbb{E}\left[(1+\left\|c_{\tau(t,i)}\right\|^{p(q-1)}+\|c_s\|^{p(q-1)})\left\|c_s-c_{\tau(t,i-1)}\right\|^p\right]\\ &\leq K\left\|1+\left\|c_{\tau(t,i)}\right\|^{p(q-1)}+\|c_s\|^{p(q-1)}\right\|_{v'}\left\|\left\|c_s-c_{\tau(t,i-1)}\right\|^p\right\|_v\\ &\leq Kd_{t,i}^{p/2}.\end{aligned}$$

Hence, $\left\|g(c_s)-g(c_{\tau(t,i-1)})\right\|_p\leq Kd_{t,i}^{1/2}$. This estimate further implies

$$\|R_3\|_p\leq Kd_t^{1/2}. \tag{S.A.28}$$

Step 4. By a mean-value expansion and condition (i),

$$|g(\hat{c}_{\tau(t,i)})-g(\hat{c}'_{\tau(t,i)})|\leq K(1+\|\hat{c}'_{\tau(t,i)}\|^{q-1})\|\hat{c}_{\tau(t,i)}-\hat{c}'_{\tau(t,i)}\|+K\|\hat{c}_{\tau(t,i)}-\hat{c}'_{\tau(t,i)}\|^q.$$

By Lemma S.A.3,

$$\mathbb{E}\|\hat{c}_{\tau(t,i)}-\hat{c}'_{\tau(t,i)}\|^q\leq Kd_t^{\bar{\theta}(k,q,\varpi,r)}.$$

Let $m'=k/2(q-1)$ and $m=k/(k-2(q-1))$. By Hölder's inequality and Lemma S.A.3,

$$\begin{aligned}&\mathbb{E}\left[(1+\|\hat{c}'_{\tau(t,i)}\|^{q-1})\|\hat{c}_{\tau(t,i)}-\hat{c}'_{\tau(t,i)}\|\right]\\ &\leq\left\|(1+\|\hat{c}'_{\tau(t,i)}\|^{q-1})\right\|_{m'}\left\|\hat{c}_{\tau(t,i)}-\hat{c}'_{\tau(t,i)}\right\|_m\\ &\leq Kd_t^{\bar{\theta}(k,m,\varpi,r)/m}.\end{aligned}$$

Therefore, we have

$$\mathbb{E}|R_4|\leq Kd_t^{\min\{\bar{\theta}(k,q,\varpi,r),\bar{\theta}(k,m,\varpi,r)/m\}}. \tag{S.A.29}$$

We now simplify the bound in (S.A.29). Note that the condition $k\geq(1-\varpi r)/(1/2-\varpi)$ implies, for any $w\geq 1$,

$$\begin{cases}k/2-\varpi\,(k-2w)-w\geq 1-\varpi r+w(2\varpi-1),\\ w(\varpi-1/2)+(1-\varpi r)\,(k-w)/k\geq 1-\varpi r+w(2\varpi-1),\end{cases} \tag{S.A.30}$$

and, recalling $m=k/(k-2(q-1))$,

$$(1-\varpi r+m(2\varpi-1))\,/m\geq 1-\varpi r+q(2\varpi-1). \tag{S.A.31}$$

Using (S.A.30) and $q\geq 2\geq r$, we simplify $\bar{\theta}\,(k,q,\varpi,r)=1-\varpi r+q(2\varpi-1)$; similarly, $\bar{\theta}\,(k,m,\varpi,r)=\min\{1-\varpi r+m(2\varpi-1),m\,(1/r-1/2)\}$. We then use (S.A.31) to simplify

(S.A.29) as

$$\mathbb{E}|R_4| \leq K d_t^{\min\{1-\varpi r+q(2\varpi-1),1/r-1/2\}}. \tag{S.A.32}$$

Combining (S.A.24), (S.A.26), (S.A.28) and (S.A.32), we readily derive the assertion of the proposition. *Q.E.D.*

PROOF OF PROPOSITION 3.4. Define Z_s as in the proof of Proposition 3.2. By applying Itô's formula to $(\Delta_{t,i}X)(\Delta_{t,i}X)^\intercal$ for each i, we have the following decomposition:

$$\begin{aligned} RV_t - QV_t &= 2\int_{t-1}^{t} Z_{s-} b_s^\intercal ds \\ &+2\int_{t-1}^{t} ds \int_{\mathbb{R}} Z_{s-}\delta(s,z)^\intercal 1_{\{\|\delta(s,z)\|>1\}}\lambda(dz) \\ &+2\int_{t-1}^{t} Z_{s-}(\sigma_s dW_s)^\intercal + 2\int_{t-1}^{t}\int_{\mathbb{R}} Z_{s-}\delta(s,z)^\intercal \tilde{\mu}(ds,dz). \end{aligned} \tag{S.A.33}$$

Recognizing the similarity between (S.A.33) and (S.A.8), we can use a similar (but simpler) argument as in the proof of Proposition 3.2 to show that the L_p norm of each component on the right-hand side of (S.A.33) is bounded by $Kd_t^{1/2}$. The assertion of the proposition readily follows. *Q.E.D.*

PROOF OF PROPOSITION 3.5. Step 1. Recall (S.A.1). We introduce some notation

$$\begin{cases} BV_t' = \frac{n_t}{n_t-1}\frac{\pi}{2}\sum_{i=1}^{n_t-1}|d_{t,i}^{-1/2}\Delta_{t,i}X'||d_{t,i+1}^{-1/2}\Delta_{t,i+1}X'|d_{t,i}, \\ \zeta_{1,i} = |d_{t,i}^{-1/2}\Delta_{t,i}X|\,|d_{t,i+1}^{-1/2}\Delta_{t,i+1}X''|, \quad \zeta_{2,i} = |d_{t,i}^{-1/2}\Delta_{t,i}X''|\,|d_{t,i+1}^{-1/2}\Delta_{t,i+1}X'|, \\ R_1 = \sum_{i=1}^{n_t-1}\zeta_{1,i}d_{t,i}, \quad R_2 = \sum_{i=1}^{n_t-1}\zeta_{2,i}d_{t,i}. \end{cases}$$

It is easy to see that $|BV_t - BV_t'| \leq K(R_1+R_2)$. By Lemmas 2.1.5 and 2.1.7 in Jacod and Protter (2012), $\mathbb{E}[|d_{t,i+1}^{-1/2}\Delta_{t,i+1}X''|^p|\mathcal{F}_{\tau(t,i)}] \leq Kd_t^{(p/r)\wedge 1-p/2}$. Moreover, note that

$$\mathbb{E}|d_{t,i}^{-1/2}\Delta_{t,i}X|^p \leq K\mathbb{E}|d_{t,i}^{-1/2}\Delta_{t,i}X'|^p + K\mathbb{E}|d_{t,i}^{-1/2}\Delta_{t,i}X''|^p \leq K.$$

By repeated conditioning, we deduce $\|\zeta_{i,1}\|_p \leq Kd_t^{(1/r)\wedge(1/p)-1/2}$, which further yields $\|R_1\|_p \leq Kd_t^{(1/r)\wedge(1/p)-1/2}$.

Now turn to R_2. Let $m = p'/p$ and $m' = p'/(p'-p)$. Since $pm' \leq k$ by assumption, we use Hölder's inequality and an argument similar to that above to derive

$$\left\|\zeta_{2,i}\right\|_p \leq \left(\mathbb{E}|d_{t,i}^{-1/2}\Delta_{t,i}X''|^{pm}\right)^{1/pm}\left(\mathbb{E}|d_{t,i+1}^{-1/2}\Delta_{t,i+1}X'|^{pm'}\right)^{1/pm'} \leq Kd_t^{(1/r)\wedge(1/p')-1/2}.$$

Hence, $\|R_2\|_p \leq K d_t^{(1/r)\wedge(1/p')-1/2}$. Combining these estimates, we deduce

$$\left\|BV_t - BV_t'\right\|_p \leq K\left\|R_1\right\|_p + K\left\|R_2\right\|_p \leq K d_t^{(1/r)\wedge(1/p')-1/2}. \tag{S.A.34}$$

Step 2. In this step, we show

$$\left\|BV_t' - \int_{t-1}^t c_s ds\right\|_p \leq K d_t^{1/2}. \tag{S.A.35}$$

For $j = 0$ or 1, we denote $\beta_{t,i,j} = \sigma_{\tau(t,i-1)} d_{t,i+j}^{-1/2} \Delta_{t,i+j} W$ and $\lambda_{t,i,j} = d_{t,i+j}^{-1/2} \Delta_{t,i+j} X' - \beta_{t,i,j}$. Observe that

$$\begin{aligned}
&\left|BV_t' - \frac{n_t}{n_t - 1}\frac{\pi}{2}\sum_{i=1}^{n_t-1} |\beta_{t,i,0}||\beta_{t,i,1}| d_{t,i}\right| \\
&\leq K \sum_{i=1}^{n_t-1} \left(|d_{t,i}^{-1/2}\Delta_{t,i}X'|\,|\lambda_{t,i,1}| + |\lambda_{t,i,0}|\,|\beta_{t,i,1}| \right) d_{t,i}.
\end{aligned}$$

Let $m = 2/p$ and $m' = 2/(2-p)$. By Hölder's inequality and Assumption HF,

$$\begin{aligned}
\left\| |d_{t,i}^{-1/2}\Delta_{t,i}X'|\,|\lambda_{t,i,1}| \right\|_p &\leq \left(\mathbb{E}|d_{t,i}^{-1/2}\Delta_{t,i}X'|^{pm'} \right)^{1/pm'} \left(\mathbb{E}|\lambda_{t,i,1}|^{pm}\right)^{1/pm} \\
&\leq K d_t^{1/2},
\end{aligned}$$

where the second inequality follows from $\mathbb{E}|d_{t,i}^{-1/2}\Delta_{t,i}X'|^q \leq K$ for each $q \in [0,k]$ and $\mathbb{E}|\lambda_{t,i,j}|^2 \leq K d_{t,i+j}$. Similarly, $\||\lambda_{t,i,0}|\,|\beta_{t,i,1}|\|_p \leq K d_t^{1/2}$. Combining these estimates, we have

$$\left\|BV_t' - \frac{n_t}{n_t - 1}\frac{\pi}{2}\sum_{i=1}^{n_t-1} |\beta_{t,i,0}||\beta_{t,i,1}| d_{t,i}\right\|_p \leq K d_t^{1/2}. \tag{S.A.36}$$

Let $\xi_i = (\pi/2)|\beta_{t,i,0}||\beta_{t,i,1}|$, $\xi_i' = \mathbb{E}\left[\xi_i|\mathcal{F}_{\tau(t,i-1)}\right]$ and $\xi_i'' = \xi_i - \xi_i'$. Under Assumption HF with $k \geq 4$, $\mathbb{E}|\xi_i''|^2 \leq \mathbb{E}|\xi_i|^2 \leq K$. Moreover, notice that ξ_i'' is $\mathcal{F}_{\tau(t,i+1)}$-measurable and $\mathbb{E}[\xi_i''|\mathcal{F}_{\tau(t,i-1)}] = 0$. Therefore, ξ_i'' is uncorrelated with ξ_l'' whenever $|i-l| \geq 2$. By the Cauchy-Schwarz inequality,

$$\mathbb{E}\left|\sum_{i=1}^{n_t-1} \xi_i'' d_{t,i}\right|^2 \leq K d_t \sum_{i=1}^{n_t-1} \mathbb{E}|\xi_i''|^2 d_{t,i} \leq K d_t. \tag{S.A.37}$$

By direct calculation, $\xi_i' = c_{\tau(t,i-1)}$. By a standard estimate, for any $s \in [\tau(t,i-1), \tau(t,i)]$, we have $\|c_s - c_{\tau(t,i-1)}\|_p \leq K d_t^{1/2}$ and, hence,

$$\left\| \sum_{i=1}^{n_t - 1} \xi_i' d_{t,i} - \int_{t-1}^{t} c_s ds \right\|_p \leq K d_t^{1/2}. \tag{S.A.38}$$

Combining (S.A.36)–(S.A.38), we derive (S.A.35).

Step 3. We now prove the assertions of the proposition. We prove part (a) by combining (S.A.34) and (S.A.35). In part (b), BV_t' coincides with BV_t because X is continuous. The assertion is simply (S.A.35). *Q.E.D.*

Proof of Proposition 3.6. We only consider $\widehat{SV}_t^+$ for brevity. To simplify notation, let $g(x) = \{x\}_+^2$, $x \in \mathbb{R}$. We also set $k(y,x) = g(y+x) - g(y) - g(x)$. It is elementary to see that $|k(y,x)| \leq K|x||y|$ for $x, y \in \mathbb{R}$. We consider the decomposition

$$\sum_{i=1}^{n_t} g(\Delta_{t,i} X) = \sum_{i=1}^{n_t} g\left(\Delta_{t,i} X'\right) + \sum_{i=1}^{n_t} g\left(\Delta_{t,i} X''\right) + \sum_{i=1}^{n_t} k(\Delta_{t,i} X', \Delta_{t,i} X''). \tag{S.A.39}$$

By Proposition 3.1 with $\mathcal{I}_t(g) \equiv \int_{t-1}^{t} \rho(c_s; g) ds = (1/2) \int_{t-1}^{t} c_s ds$, we deduce

$$\left\| \sum_{i=1}^{n_t} g\left(\Delta_{t,i} X'\right) - \mathcal{I}_t(g) \right\|_p \leq K d_t^{1/2}. \tag{S.A.40}$$

Hence, when X is continuous (so $X = X'$), the assertion of part (b) readily follows.

Now consider the second term on the right-hand side of (S.A.39). We define a process $(Z_s)_{s \in [t-1,t]}$ as follows: $Z_s = X_s'' - X_{\tau(t,i-1)}''$ when $s \in [\tau(t,i-1), \tau(t,i))$. Since $r \leq 1$ by assumption, Z is a finite-variational process. Observe that

$$\begin{aligned}
&\mathbb{E}\left| \sum_{i=1}^{n_t} g(\Delta_{t,i} X'') - \int_{t-1}^{t} \int_{\mathbb{R}} g\left(\delta(s,z)\right) \mu(ds,dz) \right|^p \\
&= \mathbb{E}\left| \int_{t-1}^{t} \int_{\mathbb{R}} k\left(Z_{s-}, \delta(s,z)\right) \mu(ds,dz) \right|^p \\
&\leq K\mathbb{E}\left| \int_{t-1}^{t} \int_{\mathbb{R}} |Z_{s-}| \Gamma(z) \mu(ds,dz) \right|^p \\
&\leq K\mathbb{E}\left[\int_{t-1}^{t} ds \int_{\mathbb{R}} |Z_{s-}|^p \Gamma(z)^p \lambda(dz) \right] + K\mathbb{E}\left[\left(\int_{t-1}^{t} ds \int_{\mathbb{R}} |Z_{s-}| \Gamma(z) \lambda(dz) \right)^p \right] \\
&\leq K d_t,
\end{aligned}$$

where the equality is by Itô's formula (Theorem II.31, Protter (2004)); the first inequality is due to $|k(y,z)| \leq K|x||y|$; the second and the third inequalities are derived by repeatedly using Lemma

2.1.7 of Jacod and Protter (2012). It then readily follows that

$$\left\|\sum_{i=1}^{n_t} g\left(\Delta_{t,i}X''\right) - \int_{t-1}^{t}\int_{\mathbb{R}} g\left(\delta\left(s,z\right)\right)\mu\left(ds,dz\right)\right\|_p \leq Kd_t^{1/p} \leq Kd_t^{1/2}. \tag{S.A.41}$$

Next, we consider the third term on the right-hand side of (S.A.39). Let $m' = p'/p$ and $m = p'/(p'-p)$. We have

$$\left\|k(\Delta_{t,i}X', \Delta_{t,i}X'')\right\|_p \leq K\left(\mathbb{E}\left|\Delta_{t,i}X'\right|^{pm}\right)^{1/pm}\left(\mathbb{E}\left|\Delta_{t,i}X''\right|^{pm'}\right)^{1/pm'} \leq Kd_t^{1/2+1/p'},$$

where the first inequality is due to $|k(y,x)| \leq K|x||y|$ and Hölder's inequality; the second inequality holds because Assumption HF holds for $k \geq pp'/(p'-p)$ and $\mathbb{E}\left|\Delta_{t,i}X''\right|^{p'} \leq Kd_t$. Hence,

$$\left\|\sum_{i=1}^{n_t} k(\Delta_{t,i}X', \Delta_{t,i}X'')\right\|_p \leq Kd_t^{1/p'-1/2}. \tag{S.A.42}$$

The assertion of part (a) readily follows from (S.A.39)–(S.A.42). *Q.E.D.*

S.A.3 Proofs of technical lemmas

Proof of Lemma S.A.1. Step 1. We outline the proof in this step. For notational simplicity, we denote $\mathbb{E}_i\xi = \mathbb{E}[\xi|\mathcal{F}_{\tau(t,i)}]$ for some generic random variable ξ; in particular, $\mathbb{E}_i|\xi|^w$ is understood as $\mathbb{E}_i[|\xi|^w]$. Let $\alpha_i = (\Delta_{t,i}X')(\Delta_{t,i}X')^{\intercal} - c_{\tau(t,i-1)}d_{t,i}$. We decompose $\hat{c}'_{\tau(t,i)} - c_{\tau(t,i)} = \zeta_{1,i} + \zeta_{2,i}$, where

$$\zeta_{1,i} = k_t^{-1}\sum_{j=1}^{k_t}(c_{\tau(t,i+j-1)} - c_{\tau(t,i)}), \quad \zeta_{2,i} = k_t^{-1}\sum_{j=1}^{k_t} d_{t,i+j}^{-1}\alpha_{i+j}. \tag{S.A.43}$$

In Steps 2 and 3 below, we show

$$\left\|\mathbb{E}_i\left\|\zeta_{1,i}\right\|^w\right\|_v \leq \begin{cases} Kd_t^{1/2} & \text{in general,} \\ Kd_t^{w/4} & \text{if } \sigma_t \text{ is continuous,} \end{cases} \tag{S.A.44}$$

$$\left\|\mathbb{E}_i\left\|\zeta_{2,i}\right\|^w\right\|_v \leq \begin{cases} Kd_t + Kk_t^{-w/2} & \text{in general,} \\ Kd_t^{w/2} + Kk_t^{-w/2} & \text{if } \sigma_t \text{ is continuous.} \end{cases} \tag{S.A.45}$$

The assertion of the lemma then readily follows from condition (ii) and $w \geq 2$.

Step 2. We show (S.A.44) in this step. Let $\bar{u} = \tau(t, i+k_t-1) - \tau(t,i)$. Since $\bar{u} = O(d_t^{1/2})$, we can assume $\bar{u} \leq 1$ without loss. By Itô's formula, c_t can be represented as

$$\begin{aligned} c_t &= c_0 + \int_0^t \bar{b}_s ds + \int_0^t \bar{\sigma}_s dW_s \\ &\quad + \int_0^t \int_{\mathbb{R}} 2\sigma_{s-} \tilde{\delta}(s,z)^\intercal \tilde{\mu}(ds,dz) + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s,z) \tilde{\delta}(s,z)^\intercal \mu(ds,dz), \end{aligned} \tag{S.A.46}$$

for some processes $\bar{b}_s$ and $\bar{\sigma}_s$ that, under condition (i), satisfy

$$\mathbb{E}\|\bar{b}_s\|^{wv} + \mathbb{E}\|\bar{\sigma}_s\|^{wv} \leq K. \tag{S.A.47}$$

By (S.A.46),

$$\|\zeta_{1,i}\|^w \leq \sup_{u\in[0,\bar{u}]} \|c_{\tau(t,i)+u} - c_{\tau(t,i)}\|^w \leq K \sum_{l=1}^4 \xi_{l,i}, \tag{S.A.48}$$

where

$$\begin{cases} \xi_{1,i} = \sup_{u\in[0,\bar{u}]} \left\| \int_{\tau(t,i)}^{\tau(t,i)+u} \bar{b}_s ds \right\|^w, \\ \xi_{2,i} = \sup_{u\in[0,\bar{u}]} \left\| \int_{\tau(t,i)}^{\tau(t,i)+u} \bar{\sigma}_s dW_s \right\|^w, \\ \xi_{3,i} = \sup_{u\in[0,\bar{u}]} \left\| \int_{\tau(t,i)}^{\tau(t,i)+u} \int_{\mathbb{R}} 2\sigma_{s-} \tilde{\delta}(s,z)^\intercal \tilde{\mu}(ds,dz) \right\|^w, \\ \xi_{4,i} = \sup_{u\in[0,\bar{u}]} \left\| \int_{\tau(t,i)}^{\tau(t,i)+u} \int_{\mathbb{R}} \tilde{\delta}(s,z) \tilde{\delta}(s,z)^\intercal \mu(ds,dz) \right\|^w. \end{cases}$$

By (S.A.47), it is easy to see that $\|\mathbb{E}_i[\xi_{1,i}]\|_v \leq \|\xi_{1,i}\|_v \leq K\bar{u}^w$. Moreover, $\|\mathbb{E}_i[\xi_{2,i}]\|_v \leq \|\xi_{2,i}\|_v \leq K\bar{u}^{w/2}$, where the second inequality is due to the Burkholder–David–Gundy inequality. By Lemma 2.1.5 in Jacod and Protter (2012) and condition (i),

$$\begin{aligned} \mathbb{E}_i[\xi_{3,i}] &\leq K\mathbb{E}_i\left[\int_{\tau(t,i)}^{\tau(t,i)+\bar{u}} \int_{\mathbb{R}} \|\sigma_{s-}\|^w \|\tilde{\delta}(s,z)\|^w \lambda(dz)\, ds\right] \\ &\quad + K\mathbb{E}_i\left[\left(\int_{\tau(t,i)}^{\tau(t,i)+\bar{u}} \int_{\mathbb{R}} \|\sigma_{s-}\|^2 \|\tilde{\delta}(s,z)\|^2 \lambda(dz)\, ds\right)^{w/2}\right] \\ &\leq K\mathbb{E}_i\left[\int_{\tau(t,i)}^{\tau(t,i)+\bar{u}} \|\sigma_{s-}\|^w ds\right] + K\mathbb{E}_i\left[\left(\int_{\tau(t,i)}^{\tau(t,i)+\bar{u}} \|\sigma_{s-}\|^2 ds\right)^{w/2}\right]. \end{aligned}$$

Hence, $\|\mathbb{E}_i[\xi_{3,i}]\|_v \leq K\bar{u}$. By Lemma 2.1.7 in Jacod and Protter (2012) and condition (i),

$$\begin{aligned} \mathbb{E}_i[\xi_{4,i}] &\leq K\mathbb{E}_i\left[\int_{\tau(t,i)}^{\tau(t,i)+\bar{u}} \int_{\mathbb{R}} \|\tilde{\delta}(s,z)\|^{2w} \lambda(dz)\, ds\right] \\ &\quad + K\mathbb{E}_i\left[\left(\int_{\tau(t,i)}^{\tau(t,i)+\bar{u}} \int_{\mathbb{R}} \|\tilde{\delta}(s,z)\|^2 \lambda(dz)\, ds\right)^w\right] \\ &\leq K\bar{u}. \end{aligned}$$

Hence, $\|\mathbb{E}_i[\xi_{4,i}]\|_v \leq K\bar{u}$. Combining these estimates with (S.A.48), we derive (S.A.44) in the general case as desired. Furthermore, when σ_t is continuous, we have $\xi_{3,i} = \xi_{4,i} = 0$ in (S.A.48). The assertion of (S.A.44) in the continuous case readily follows.

Step 3. In this step, we show (S.A.45). Let $\alpha_i' = \mathbb{E}_{i-1}[\alpha_i]$ and $\alpha_i'' = \alpha_i - \alpha_i'$. We can then decompose $\zeta_{2,i} = \zeta_{2,i}' + \zeta_{2,i}''$, where $\zeta_{2,i}' = k_t^{-1}\sum_{j=1}^{k_t} d_{t,i+j}^{-1}\alpha_{i+j}'$ and $\zeta_{2,i}'' = k_t^{-1}\sum_{j=1}^{k_t} d_{t,i+j}^{-1}\alpha_{i+j}''$. By Itô's formula, it is easy to see that

$$\begin{aligned} \|\alpha_{i+j}'\| \quad &\leq \quad K\left\|\mathbb{E}_{i+j-1}\left[\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (X_s' - X_{\tau(t,i+j-1)}')(b_s')^\intercal ds\right]\right\| \\ &\quad + \left\|\mathbb{E}_{i+j-1}\left[\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (c_s - c_{\tau(t,i+j-1)})ds\right]\right\|. \end{aligned} \tag{S.A.49}$$

By Jensen's inequality and repeated conditioning,

$$\begin{aligned} \mathbb{E}_i\|\alpha_{i+j}'\|^w \quad &\leq \quad K\mathbb{E}_i\left\|\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (X_s' - X_{\tau(t,i+j-1)}')(b_s')^\intercal ds\right\|^w \\ &\quad + K\mathbb{E}_i\left\|\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (c_s - c_{\tau(t,i+j-1)})ds\right\|^w. \end{aligned} \tag{S.A.50}$$

Since conditional expectations are contraction maps, we further have

$$\begin{aligned} \left\|\mathbb{E}_i\|\alpha_{i+j}'\|^w\right\|_v \quad &\leq \quad K\left\|\left\|\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (X_s' - X_{\tau(t,i+j-1)}')(b_s')^\intercal ds\right\|^w\right\|_v \\ &\quad + K\left\|\mathbb{E}_i\left\|\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (c_s - c_{\tau(t,i+j-1)})ds\right\|^w\right\|_v. \end{aligned} \tag{S.A.51}$$

By standard estimates, the first term on the majorant side of (S.A.51) is bounded by $Kd_{t,i+j}^{3w/2}$. Following a similar argument as in Step 2, we can bound the second term on the majorant side of (S.A.51) by $Kd_{t,i+j}^{w+1}$ in general and by $Kd_{t,i+j}^{3w/2}$ if σ_t is continuous. Hence, $\|\mathbb{E}_i\|\alpha_{i+j}'\|^w\|_v \leq Kd_{t,i+j}^{w+1}$, and the bound can be improved to be $Kd_{t,i+j}^{3w/2}$ when σ_t is continuous. By Hölder's inequality and the triangle inequality,

$$\left\|\mathbb{E}_i\left\|\zeta_{2,i}'\right\|^w\right\|_v \leq \begin{cases} Kd_t & \text{in general,} \\ Kd_t^{w/2} & \text{when } \sigma_t \text{ is continuous.} \end{cases} \tag{S.A.52}$$

Now consider $\zeta_{2,i}''$. Notice that $(\alpha_{i+j}'')_{1\leq j\leq k_t}$ forms a martingale difference sequence. Using the

Burkholder–Davis–Gundy inequality and then Hölder's inequality, we derive

$$\mathbb{E}_i \left\|\zeta''_{2,i}\right\|^w \le K k_t^{-w/2-1} \sum_{j=1}^{k_t} d_{t,i+j}^{-w} \mathbb{E}_i \|\alpha''_{i+j}\|^w.$$

Moreover, notice that $\|\mathbb{E}_i\|\alpha''_{i+j}\|^w\|_v \le \|\|\alpha''_{i+j}\|^w\|_v \le K d_{t,i+j}^w$. Hence, $\|\mathbb{E}_i\|\zeta''_{2,i}\|^w\|_v \le K k_t^{-w/2}$. Combining this estimate with (S.A.52), we have (S.A.45). This finishes the proof. *Q.E.D.*

Proof of Lemma S.A.2. Step 1. Recall the notation in Step 1 of the proof of Lemma S.A.1. In this step, we show that

$$\left\|\left\|\mathbb{E}_i \zeta_{1,i}\right\|^w\right\|_v \le K d_t^{w/2}. \tag{S.A.53}$$

By (S.A.46), for each $j \ge 1$,

$$\mathbb{E}_i \left[c_{\tau(t,i+j-1)} - c_{\tau(t,i)}\right] = \mathbb{E}_i \left[\int_{\tau(t,i)}^{\tau(t,i+j-1)} \bar{b}_s ds + \int_{\tau(t,i)}^{\tau(t,i+j-1)} \int_{\mathbb{R}} \tilde{\delta}(s,z) \tilde{\delta}(s,z)^\intercal \lambda(dz) ds\right]. \tag{S.A.54}$$

By conditions (i,ii) and Hölder's inequality, we have

$$\left\|\left\|\mathbb{E}_i \left[c_{\tau(t,i+j-1)} - c_{\tau(t,i)}\right]\right\|^w\right\|_v \le K (k_t d_t)^w \le K d_t^{w/2}. \tag{S.A.55}$$

We then use Hölder's inequality and Minkowski's inequality to derive (S.A.53).

Step 2. Similar to (S.A.49), we have

$$\begin{aligned} \|\mathbb{E}_i [\alpha_{i+j}]\| \;\; &\le \;\; K \left\|\mathbb{E}_i \left[\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (X'_s - X'_{\tau(t,i+j-1)})(b'_s)^\intercal ds\right]\right\| \\ &\quad + \left\|\mathbb{E}_i \left[\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (c_s - c_{\tau(t,i+j-1)}) ds\right]\right\|. \end{aligned}$$

Notice that

$$\begin{aligned} &\left\|\left\|\mathbb{E}_i \left[\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (X'_s - X'_{\tau(t,i+j-1)})(b'_s)^\intercal ds\right]\right\|^w\right\|_v \\ &\quad \le K \left\|\left\|\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (X'_s - X'_{\tau(t,i+j-1)})(b'_s)^\intercal ds\right\|^w\right\|_v \le K d_{t,i+j}^{3w/2}, \end{aligned} \tag{S.A.56}$$

where the first inequality is due to Jensen's inequality; the second inequality follows from standard estimates for continuous Itô semimartingales (use Hölder's inequality and the Burkholder–Davis–Gundy inequality). Similar to (S.A.55), we have $\left\|\left\|\mathbb{E}_i \left[c_s - c_{\tau(t,i+j-1)}\right]\right\|^w\right\|_v \le K d_{t,i+j}^w$ for $s \in$

$[\tau(t,i+j-1),\tau(t,i+j)]$. We then use Hölder's inequality to derive

$$\left\| \left\| \mathbb{E}_i \left[\int_{\tau(t,i+j-1)}^{\tau(t,i+j)} (c_s - c_{\tau(t,i+j-1)}) ds \right] \right\|^w \right\|_v \leq K d_{t,i+j}^{2w}. \tag{S.A.57}$$

Combining (S.A.56) and (S.A.57), we deduce $\| \| \mathbb{E}_i [\alpha_{i+j}] \|^w \|_v \leq K d_{t,i+j}^{3w/2}$. Hence, by Hölder's inequality, $\left\| \left\| \mathbb{E}_i [\zeta_{2,i}] \right\|^w \right\|_v \leq K d_t^{w/2}$. This estimate, together with (S.A.53), implies the assertion of the lemma. *Q.E.D.*

Proof of Lemma S.A.3. We denote $u_{t,i+j} = \bar{\alpha} d_{t,i+j}^{\varpi}$. We shall use the following elementary inequality: for all $x, y \in \mathbb{R}^d$ and $0 < u < 1$:

$$\begin{aligned} &\| (x+y)(x+y)^\intercal 1_{\{\|x+y\| \leq u\}} - x x^\intercal \| \\ &\quad \leq K(\|x\|^2 1_{\{\|x\| > u/2\}} + \|y\|^2 \wedge u^2 + \|x\|(\|y\| \wedge u)). \end{aligned} \tag{S.A.58}$$

Applying (S.A.58) with $x = \Delta_{t,i+j} X'$, $y = \Delta_{t,i+j} X''$ and $u = u_{t,i+j}$, we have $\|\hat{c}_{\tau(t,i)} - \hat{c}'_{\tau(t,i)}\| \leq K(\zeta_1 + \zeta_2 + \zeta_3)$, where

$$\begin{aligned} \zeta_1 &= k_t^{-1} \sum_{j=1}^{k_t} d_{t,i+j}^{-1} \|\Delta_{t,i+j} X'\|^2 1_{\{\|\Delta_{t,i+j} X'\| > u_{t,i+j}/2\}} \\ \zeta_2 &= k_t^{-1} \sum_{j=1}^{k_t} d_{t,i+j}^{-1} (\|\Delta_{t,i+j} X''\| \wedge u_{t,i+j})^2 \\ \zeta_3 &= k_t^{-1} \sum_{j=1}^{k_t} d_{t,i+j}^{-1} \|\Delta_{t,i+j} X'\| (\|\Delta_{t,i+j} X''\| \wedge u_{t,i+j}). \end{aligned}$$

Since $k \geq 2w$, by Markov's inequality and $\mathbb{E}\|\Delta_{t,i+j} X'\|^k \leq K d_{t,i+j}^{k/2}$, we have

$$\begin{aligned} &\mathbb{E} \left| \|\Delta_{t,i+j} X'\|^2 1_{\{\|\Delta_{t,i+j} X'\| > u_{t,i+j}/2\}} \right|^w \\ &\qquad \leq K \frac{\mathbb{E}\|\Delta_{t,i+j} X'\|^k}{u_{t,i+j}^{k-2w}} \leq K d_{t,i+j}^{k/2 - \varpi(k-2w)}. \end{aligned}$$

Hence, $\mathbb{E}\|\zeta_1\|^w \leq K d_t^{k/2 - \varpi(k-2w) - w}$.

By Corollary 2.1.9(a,c) in Jacod and Protter (2012), we have for any $v > 0$,

$$\mathbb{E} \left[\left(\frac{\|\Delta_{t,i+j} X''\|}{d_{t,i+j}^{\varpi}} \wedge 1 \right)^v \right] \leq K d_{t,i+j}^{(1-\varpi r) \min\{v/r, 1\}}. \tag{S.A.59}$$

Applying (S.A.59) with $v = 2w$, we have $\mathbb{E}[(\|d_{t,i+j}^{-\varpi}\Delta_{t,i+j}X''\| \wedge 1)^{2w}] \leq Kd_{t,i+j}^{1-\varpi r}$. Therefore, $\mathbb{E}\|\zeta_2\|^w \leq Kd_t^{1-\varpi r+w(2\varpi-1)}$.

We now turn to ζ_3. Let $m' = k/w$ and $m = k/(k-w)$. Observe that

$$
\begin{aligned}
&\mathbb{E}\left|\|\Delta_{t,i+j}X'\|(\|u_{t,i+j}^{-1}\Delta_{t,i+j}X''\| \wedge 1)\right|^w \\
&\leq K\left\{\mathbb{E}\|\Delta_{t,i+j}X'\|^{wm'}\right\}^{1/m'}\left\{\mathbb{E}\left[(\|u_{t,i+j}^{-1}\Delta_{t,i+j}X''\| \wedge 1)^{wm}\right]\right\}^{1/m} \\
&\leq Kd_{t,i+j}^{w/2+(1-\varpi r)\min\{w/r,(k-w)/k\}},
\end{aligned}
$$

where the first inequality is by Hölder's inequality; the second inequality is obtained by applying (S.A.59) with $v = wm$. Therefore, $\mathbb{E}\|\zeta_3\|^w \leq Kd_t^{w(\varpi-1/2)+(1-\varpi r)\min\{w/r,(k-w)/k\}}$.

Combining the above bounds for $\mathbb{E}\|\zeta_j\|^w$, $j = 1, 2$ or 3, we readily derive the assertion of the lemma. *Q.E.D.*

Appendix S.B Extensions: details on stepwise procedures

S.B.1 The StepM procedure

In this subsection, we provide details for implementing the StepM procedure of Romano and Wolf (2005) using proxies, so as to complete the discussion in Section 4.2 of the main text. Recall that we are interested in testing $\bar{k}$ pairs of hypotheses

$$
\text{Multiple SPA}\begin{cases} H_{j,0}: \mathbb{E}[f_{j,t+\tau}^{\dagger}] \leq 0 \text{ for all } t \geq 1, \\ H_{j,a}: \liminf_{T\to\infty} \mathbb{E}[\bar{f}_{j,T}^{\dagger}] > 0, \end{cases} \qquad 1 \leq j \leq \bar{k}. \tag{S.B.1}
$$

We denote the test statistic for the jth testing problem as $\varphi_{j,T} \equiv \varphi_j(a_T\bar{f}_T, a_T'S_T)$, where $\varphi_j(\cdot,\cdot)$ is a measurable function. The StepM procedure involves critical values $\hat{c}_{1,T} \geq \hat{c}_{2,T} \geq \cdots$, where $\hat{c}_{l,T}$ is the critical value in step l. Given these notations, we can describe Romano and Wolf's StepM algorithm as follows.[1]

Algorithm 1 (StepM): Step 1. Set $l = 1$ and $\mathcal{A}_{0,T} = \{1, \ldots, \bar{k}\}$.

Step 2. Compute the step-l critical value $\hat{c}_{l,T}$. Reject the null hypothesis $H_{j,0}$ if $\varphi_{j,T} > \hat{c}_{l,T}$.

Step 3. If no (further) null hypotheses are rejected or all hypotheses have been rejected, stop; otherwise, let $\mathcal{A}_{l,T}$ be the index set for hypotheses that have yet been rejected, that is, $\mathcal{A}_{l,T} = \{j : 1 \leq j \leq \bar{k},\ \varphi_{j,T} \leq \hat{c}_{l,T}\}$, set $l = l + 1$ and then return to Step 2.

To specify the critical value $\hat{c}_{l,T}$, we make the following assumption. Below, $\alpha \in (0,1)$ denotes the significance level and (ξ, S) is defined in Assumption A1 in the main text.

[1] The presentation here unifies Algorithms 3.1 (non-studentized StepM) and Algorithm 4.1 (studentized StepM) in Romano and Wolf (2005).

Assumption S: For any nonempty nonrandom $\mathcal{A} \subseteq \{1, \ldots, \bar{k}\}$, the distribution function of $\max_{j\in\mathcal{A}} \varphi_j(\xi, S)$ is continuous at its $1-\alpha$ quantile $c(\mathcal{A}, 1-\alpha)$. Moreover, there exists a sequence of estimators $\hat{c}_T(\mathcal{A}, 1-\alpha)$ such that $\hat{c}_T(\mathcal{A}, 1-\alpha) \xrightarrow{\mathbb{P}} c(\mathcal{A}, 1-\alpha)$ and $\hat{c}_T(\mathcal{A}, 1-\alpha) \leq \hat{c}_T(\mathcal{A}', 1-\alpha)$ whenever $\mathcal{A} \subseteq \mathcal{A}'$.

The step-l critical value is then given by $\hat{c}_{l,T} = \hat{c}_T(\mathcal{A}_{l-1,T}, 1-\alpha)$. Notice that $\hat{c}_{1,T} \geq \hat{c}_{2,T} \geq \cdots$ in finite samples by construction. The bootstrap critical values proposed by Romano and Wolf (2005) verify Assumption S.

The following proposition describes the asymptotic properties of the StepM procedure. We remind the reader that Assumptions A1, A2, B1 and C1 are given in the main text.

Proposition S.B.1: Suppose that Assumptions A1, C1 and S hold and that Assumptions A2 and B1 hold for each $\varphi_j(\cdot)$, $1 \leq j \leq \bar{k}$. Then (a) the null hypothesis $H_{j,0}$ is rejected with probability tending to one under the alternative hypothesis $H_{j,a}$; (b) Algorithm 1 asymptotically controls the familywise error rate (FWE) at level α.

Proof. By Assumptions A1 and C1,

$$(a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^\dagger]), a_T' S_T) \xrightarrow{d} (\xi, S). \tag{S.B.2}$$

The proof is then similar to that in Romano and Wolf (2005). The details are given below.

First consider part (a), so $H_{j,a}$ is true for some j. By (S.B.2) and Assumption B1(b), $\varphi_{j,T}$ diverges to $+\infty$ in probability. By Assumption S, it is easy to see that $\hat{c}_{l,T}$ forms a tight sequence for fixed l. Hence, $\varphi_{j,T} > \hat{c}_{l,T}$ with probability tending to one. From here the assertion in part (a) follows.

Now turn to part (b). Let $I_0 = \{j : 1 \leq j \leq \bar{k},\ H_{0,j} \text{ is true}\}$ and $\text{FWE}_T = \mathbb{P}(H_{j,0} \text{ is rejected for some } j \in I_0)$. If I_0 is empty, $\text{FWE}_T = 0$ and there is nothing to prove. We can thus suppose that I_0 is nonempty without loss of generality. By part (a), all false hypotheses are rejected in the first step with probability approaching one. Since $\hat{c}_T(I_0, 1-\alpha) \leq \hat{c}_{1,T}$,

$$\begin{aligned}
\limsup_{T\to\infty} \text{FWE}_T &= \limsup_{T\to\infty} \mathbb{P}\left(\varphi_j(a_T\bar{f}_T, a_T'S_T) > \hat{c}_T(I_0, 1-\alpha) \text{ for some } j \in I_0\right) \\
&\leq \limsup_{T\to\infty} \mathbb{P}\left(\varphi_j(a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^\dagger]), a_T'S_T) > \hat{c}_T(I_0, 1-\alpha) \text{ for some } j \in I_0\right) \\
&= \limsup_{T\to\infty} \mathbb{P}\left(\max_{j\in I_0} \varphi_j(a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^\dagger]), a_T'S_T) > \hat{c}_T(I_0, 1-\alpha)\right) \\
&= \mathbb{P}\left(\max_{j\in I_0} \varphi_j(\xi, S) > c(I_0, 1-\alpha)\right) \\
&= \alpha.
\end{aligned}$$

This is the assertion of part (b). *Q.E.D.*

S.B.2 Model confidence sets

In this subsection, we provide details for constructing the model confidence set (MCS) using proxies. In so doing, we complete the discussion in Section 4.3 of the main text. Below, we denote the paper of Hansen, Lunde, and Nason (2011) by HLN.

Recall that the set of superior forecasts is defined as

$$\overline{\mathcal{M}}^{\dagger} \equiv \left\{ j \in \{1, \ldots, \bar{k}\} : \mathbb{E}[f^{\dagger}_{j,t+\tau}] \geq \mathbb{E}[f^{\dagger}_{l,t+\tau}] \text{ for all } 1 \leq l \leq \bar{k} \text{ and } t \geq 1 \right\},$$

and the set of asymptotically inferior forecasts is given by

$$\underline{\mathcal{M}}^{\dagger} \equiv \left\{ j \in \{1, \ldots, \bar{k}\} : \begin{array}{l} \liminf_{T\to\infty} \left(\mathbb{E}[\bar{f}^{\dagger}_{l,T}] - \mathbb{E}[\bar{f}^{\dagger}_{j,T}] \right) > 0 \\ \text{for some (and hence any) } l \in \overline{\mathcal{M}}^{\dagger} \end{array} \right\}.$$

The formulation above slightly generalizes HLN's setting by allowing data heterogeneity. Under (mean) stationarity, $\overline{\mathcal{M}}^{\dagger}$ coincides with HLN's definition of MCS; in particular, it is nonempty and complemental to $\underline{\mathcal{M}}^{\dagger}$. In the heterogeneous setting, $\overline{\mathcal{M}}^{\dagger}$ may be empty and the union of $\overline{\mathcal{M}}^{\dagger}$ and $\underline{\mathcal{M}}^{\dagger}$ may be inexhaustive. We avoid these scenarios by imposing

Assumption M1: $\overline{\mathcal{M}}^{\dagger}$ is nonempty and $\overline{\mathcal{M}}^{\dagger} \cup \underline{\mathcal{M}}^{\dagger} = \{1, \ldots, \bar{k}\}$.

We now describe the MCS algorithm. We first need to specify some test statistics. Below, for any subset $\mathcal{M} \subseteq \{1, \ldots, \bar{k}\}$, we denote its cardinality by $|\mathcal{M}|$. We consider the test statistic

$$\varphi_{\mathcal{M},T} = \varphi_{\mathcal{M}}(a_T \bar{f}_T, a'_T S_T), \quad \text{where} \quad \varphi_{\mathcal{M}}(\cdot,\cdot) = \max_{j\in\mathcal{M}} \varphi_{j,\mathcal{M}}(\cdot,\cdot),$$

and, as in HLN (see Section 3.1.2 there), $\varphi_{j,\mathcal{M}}(\cdot,\cdot)$ may take either of the following two forms: for $u \in \mathbb{R}^{\bar{k}}$ and $1 \leq j \leq \bar{k}$,

$$\varphi_{j,\mathcal{M}}(u,s) = \begin{cases} \max_{i\in\mathcal{M}} \dfrac{u_i - u_j}{\sqrt{s_{ij}}}, & \text{where } s_{ij} = s_{ji} \in (0,\infty) \text{ for all } 1 \leq i \leq \bar{k}, \\ \dfrac{|\mathcal{M}|^{-1} \sum_{i\in\mathcal{M}} u_i - u_j}{\sqrt{s_j}}, & \text{where } s_j \in (0,\infty). \end{cases}$$

We also need to specify critical values, for which we need Assumption M2 below. We remind the reader that the variables (ξ, S) are defined in Assumption A1 in the main text.

Assumption M2: For any nonempty nonrandom $\mathcal{M} \subseteq \{1, \ldots, \bar{k}\}$, the distribution of $\varphi_{\mathcal{M}}(\xi, S)$

is continuous at its $1-\alpha$ quantile $c(\mathcal{M}, 1-\alpha)$. Moreover, there exists a sequence of estimators $\hat{c}_T(\mathcal{M}, 1-\alpha)$ such that $\hat{c}_T(\mathcal{M}, 1-\alpha) \xrightarrow{\mathbb{P}} c(\mathcal{M}, 1-\alpha)$.

With $\hat{c}_T(\mathcal{M}, 1-\alpha)$ given in Assumption M2, we define a test $\phi_{\mathcal{M},T} = \mathbf{1}\{\varphi_{\mathcal{M},T} > \hat{c}_T(\mathcal{M}, 1-\alpha)\}$ and an elimination rule $e_{\mathcal{M}} = \arg\max_{j\in\mathcal{M}} \varphi_{j,\mathcal{M},T}$, where $\varphi_{j,\mathcal{M},T} \equiv \varphi_{j,\mathcal{M}}(a_T \bar{f}_T, a_T' S_T)$. The MCS algorithm, when applied with the proxy as the evaluation benchmark, is given as follows.

Algorithm 2 (MCS): Step 1: Set $\mathcal{M} = \{1, \ldots, \bar{k}\}$.
Step 2: if $|\mathcal{M}| = 1$ or $\phi_{\mathcal{M},T} = 0$, then stop and set $\widehat{\mathcal{M}}_{T,1-\alpha} = \mathcal{M}$; otherwise continue.
Step 3. Set $\mathcal{M} = \mathcal{M} \setminus e_{\mathcal{M}}$ and return to Step 2.

The following proposition summarizes the asymptotic property of $\widehat{\mathcal{M}}_{T,1-\alpha}$. In particular, it shows that the MCS algorithm is asymptotically valid even though it is applied to the proxy instead of the true target.

Proposition S.B.2: Suppose Assumptions A1, C1, M1 and M2. Then (4.5) in the main text holds, that is,

$$\liminf_{T\to\infty} \left(\overline{\mathcal{M}}^\dagger \subseteq \widehat{\mathcal{M}}_{T,1-\alpha}\right) \geq 1-\alpha, \quad \mathbb{P}\left(\widehat{\mathcal{M}}_{T,1-\alpha} \cap \underline{\mathcal{M}}^\dagger = \emptyset\right) \to 1.$$

Proof. Under Assumptions A1 and C1, we have $(a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^\dagger]), a_T' S_T) \xrightarrow{d} (\xi, S)$. For each $\mathcal{M} \subseteq \{1, \ldots, \bar{k}\}$, we consider the null hypothesis $H_{0,\mathcal{M}} : \mathcal{M} \subseteq \overline{\mathcal{M}}^\dagger$ and the alternative hypothesis $H_{a,\mathcal{M}} : \mathcal{M} \cap \underline{\mathcal{M}}^\dagger \neq \emptyset$. Under $H_{0,\mathcal{M}}$, $\varphi_{\mathcal{M},T} = \varphi_{\mathcal{M}}(a_T \bar{f}_T, a_T' S_T) = \varphi_{\mathcal{M}}(a_T(\bar{f}_T - \mathbb{E}[\bar{f}_T^\dagger]), a_T' S_T)$, and, thus, by the continuous mapping theorem, $\varphi_{\mathcal{M},T} \xrightarrow{d} \varphi_{\mathcal{M}}(\xi, S)$. Therefore, by Assumption M2, $\mathbb{E}\phi_{\mathcal{M},T} \to \alpha$ under $H_{0,\mathcal{M}}$. On the other hand, under $H_{a,\mathcal{M}}$, $\varphi_{\mathcal{M},T}$ diverges in probability to $+\infty$ and thus $\mathbb{E}\phi_{\mathcal{M},T} \to 1$. Moreover, under $H_{a,\mathcal{M}}$, $\mathbb{P}(e_{\mathcal{M}} \in \overline{\mathcal{M}}^\dagger) \to 0$; this is because $\sup_{j\in\overline{\mathcal{M}}^\dagger\cap\mathcal{M}} \varphi_{j,\mathcal{M},T}$ is either tight or diverges in probability to $-\infty$, but $\varphi_{\mathcal{M},T}$ diverges to $+\infty$ in probability. The assertions then follow the same argument as in the proof of Theorem 1 in HLN. *Q.E.D.*

Appendix S.C Additional simulation results

S.C.1 Sensitivity to the choice of truncation lag in long-run variance estimation

In Tables S.I–S.VI, we present results on the finite-sample rejection frequencies of the Giacomini and White (2006) tests (GW) using the approaches of Newey and West (1987) and Kiefer and Vogelsang (2005) to conduct inference; we denote this two approaches respectively by NW and KV. In the main text, we use a truncation lag of $3P^{1/3}$ for NW and $0.5P$ for KV when computing the long-run variance. Below we further consider using $P^{1/3}$ and 5 (for all P) for NW, and $0.25P$ and P for KV.

Proxy RV_{t+1}^{Δ}	GW–NW ($m = 5$) $P = 500$	$P = 1000$	$P = 2000$	GW–NW ($m = P^{1/3}$) $P = 500$	$P = 1000$	$P = 2000$
			$R = 500$			
True $Y_{t+1}^{\dagger}$	0.10	0.22	0.18	0.09	0.18	0.11
$\Delta = 5$ sec	0.10	0.23	0.18	0.09	0.18	0.11
$\Delta = 1$ min	0.09	0.23	0.18	0.09	0.17	0.11
$\Delta = 5$ min	0.10	0.23	0.18	0.09	0.18	0.12
$\Delta = 30$ min	0.10	0.27	0.22	0.08	0.22	0.16
			$R = 1000$			
True $Y_{t+1}^{\dagger}$	0.28	0.22	0.19	0.24	0.15	0.12
$\Delta = 5$ sec	0.29	0.22	0.18	0.24	0.15	0.12
$\Delta = 1$ min	0.29	0.22	0.19	0.24	0.15	0.12
$\Delta = 5$ min	0.30	0.21	0.19	0.26	0.17	0.12
$\Delta = 30$ min	0.35	0.26	0.25	0.31	0.20	0.18

Table S.I: Giacomini–White test rejection frequencies for Simulation A. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy, and m is the truncation lag in the long-run variance estimation.

Overall, we confirm that feasible tests using proxies have finite-sample rejection rates similar to those of the infeasible test using the true target. That is, the negligibility result is likely in force. More specifically, we find that the GW–KV approach has good size control across various settings provided that the sample size is sufficiently large ($P = 1000$ or 2000), although the test is somewhat conservative in Simulation A. In contrast, the performance of the GW–NW test is less robust. The GW–NW test has good size control in Simulation B, but has substantial size distortion in Simulations A and C. This finding is not surprising, and it confirms the insight from the literature on inconsistent long-run variance estimation; see Kiefer and Vogelsang (2005), Müller (2012) and references therein.

S.C.2 Disagreement between feasible and infeasible test indicators

In Tables S.VII–S.IX, we report the disagreement on test decisions (i.e., rejection or non-rejection) between infeasible tests based on the true target variable and feasible tests based on proxies. In view of the size distortion of the GW–NW test, we only consider the GW–KV test for brevity. The setting is the same as that in Section 5 of the main text. In the columns headed “Weak” we report the finite-sample rejection frequency of the feasible test minus that for the infeasible test. Under the theory developed in Section 2, which ensures “weak negligibility,” the differences should be

Proxy BV_{t+1}^{Δ}	GW–NW ($m = 5$) $P = 500$	$P = 1000$	$P = 2000$	GW–NW ($m = P^{1/3}$) $P = 500$	$P = 1000$	$P = 2000$
			$R = 500$			
True $Y_{t+1}^{\dagger}$	0.05	0.06	0.06	0.05	0.06	0.06
$\Delta = 5$ sec	0.06	0.06	0.06	0.06	0.06	0.06
$\Delta = 1$ min	0.07	0.08	0.07	0.07	0.08	0.07
$\Delta = 5$ min	0.03	0.05	0.04	0.03	0.05	0.04
$\Delta = 30$ min	0.03	0.02	0.00	0.03	0.02	0.00
			$R = 1000$			
True $Y_{t+1}^{\dagger}$	0.03	0.04	0.04	0.03	0.04	0.04
$\Delta = 5$ sec	0.03	0.04	0.04	0.03	0.04	0.04
$\Delta = 1$ min	0.04	0.05	0.06	0.04	0.05	0.06
$\Delta = 5$ min	0.03	0.04	0.05	0.03	0.04	0.05
$\Delta = 30$ min	0.02	0.01	0.01	0.02	0.01	0.01

Table S.II: Giacomini–White test rejection frequencies for Simulation B. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy, and m is the truncation lag in the long-run variance estimation.

zero asymptotically.[2] In the columns headed "Strong" we report the proportion of times in which the feasible and infeasible rejection indicators disagreed. If "strong negligibility," in the sense of comment (ii) to Theorem 2.1, holds, then this proportion should be zero asymptotically.

As noted in the main text, the weak negligibility result holds well across all three simulation designs, with the differences reported in these columns generally being close to zero, except for the lowest frequency proxy. The results for strong negligibility are more mixed: in Simulations A and C we see evidence in support of strong negligibility, while for Simulation B we observe a large proportion of disagreement. Indeed, as the nominal level of each test is 0.05, the probability of disagreement should be bounded by 0.1 asymptotically, so a disagreement proportion between 0.03 to 0.07 should be considered sizable.

[2]Positive (negative) values indicate that the feasible test based on a proxy rejects more (less) often than the corresponding infeasible test based on the true target variable.

Proxy RC^{Δ}_{t+1}	GW–NW ($m = 5$)			GW–NW ($m = P^{1/3}$)		
	$P = 500$	$P = 1000$	$P = 2000$	$P = 500$	$P = 1000$	$P = 2000$
			$R = 500$			
True $Y^{\dagger}_{t+1}$	0.35	0.33	0.29	0.33	0.30	0.25
$\Delta = 5$ sec	0.35	0.32	0.28	0.33	0.30	0.25
$\Delta = 1$ min	0.35	0.32	0.29	0.33	0.30	0.25
$\Delta = 5$ min	0.34	0.34	0.30	0.33	0.30	0.24
$\Delta = 30$ min	0.34	0.34	0.28	0.32	0.32	0.25
			$R = 1000$			
True $Y^{\dagger}_{t+1}$	0.36	0.26	0.28	0.31	0.22	0.20
$\Delta = 5$ sec	0.36	0.26	0.28	0.31	0.22	0.21
$\Delta = 1$ min	0.36	0.26	0.28	0.32	0.22	0.21
$\Delta = 5$ min	0.34	0.25	0.26	0.31	0.23	0.22
$\Delta = 30$ min	0.31	0.24	0.26	0.30	0.22	0.22

Table S.III: Giacomini–White test rejection frequencies for Simulation C. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy, and m is the truncation lag in the long-run variance estimation.

Proxy RV^{Δ}_{t+1}	GW–KV ($m = 0.25P$)			GW–KV ($m = P$)		
	$P = 500$	$P = 1000$	$P = 2000$	$P = 500$	$P = 1000$	$P = 2000$
			$R = 500$			
True $Y^{\dagger}_{t+1}$	0.00	0.02	0.01	0.01	0.03	0.02
$\Delta = 5$ sec	0.00	0.02	0.01	0.01	0.02	0.02
$\Delta = 1$ min	0.01	0.02	0.01	0.01	0.02	0.02
$\Delta = 5$ min	0.00	0.03	0.02	0.01	0.03	0.02
$\Delta = 30$ min	0.00	0.04	0.03	0.01	0.04	0.05
			$R = 1000$			
True $Y^{\dagger}_{t+1}$	0.06	0.01	0.02	0.06	0.00	0.02
$\Delta = 5$ sec	0.06	0.01	0.02	0.06	0.00	0.02
$\Delta = 1$ min	0.06	0.01	0.02	0.06	0.00	0.02
$\Delta = 5$ min	0.06	0.01	0.01	0.08	0.01	0.02
$\Delta = 30$ min	0.10	0.02	0.03	0.08	0.01	0.03

Table S.IV: Giacomini–White test rejection frequencies for Simulation A. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy, and m is the truncation lag in the long-run variance estimation.

Proxy BV_{t+1}^{Δ}	GW–KV ($m = 0.25P$) $P = 500$	$P = 1000$	$P = 2000$	GW–KV ($m = P$) $P = 500$	$P = 1000$	$P = 2000$
			$R = 500$			
True $Y_{t+1}^{\dagger}$	0.03	0.03	0.05	0.03	0.04	0.04
$\Delta = 5$ sec	0.05	0.04	0.05	0.03	0.05	0.05
$\Delta = 1$ min	0.04	0.06	0.05	0.05	0.05	0.07
$\Delta = 5$ min	0.02	0.05	0.04	0.03	0.06	0.05
$\Delta = 30$ min	0.03	0.03	0.01	0.03	0.03	0.01
			$R = 1000$			
True $Y_{t+1}^{\dagger}$	0.02	0.04	0.05	0.02	0.03	0.05
$\Delta = 5$ sec	0.02	0.04	0.05	0.04	0.04	0.05
$\Delta = 1$ min	0.03	0.04	0.07	0.03	0.04	0.06
$\Delta = 5$ min	0.03	0.03	0.05	0.04	0.02	0.05
$\Delta = 30$ min	0.02	0.01	0.02	0.02	0.02	0.01

Table S.V: Giacomini–White test rejection frequencies for Simulation B. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy, and m is the truncation lag in the long-run variance estimation.

Proxy RC_{t+1}^{Δ}	GW–KV ($m = 0.25P$) $P = 500$	$P = 1000$	$P = 2000$	GW–KV ($m = P$) $P = 500$	$P = 1000$	$P = 2000$
			$R = 500$			
True $Y_{t+1}^{\dagger}$	0.13	0.09	0.05	0.11	0.06	0.05
$\Delta = 5$ sec	0.13	0.09	0.05	0.11	0.06	0.05
$\Delta = 1$ min	0.13	0.09	0.05	0.11	0.06	0.05
$\Delta = 5$ min	0.12	0.09	0.04	0.11	0.07	0.05
$\Delta = 30$ min	0.13	0.08	0.05	0.12	0.07	0.05
			$R = 1000$			
True $Y_{t+1}^{\dagger}$	0.14	0.08	0.03	0.11	0.08	0.03
$\Delta = 5$ sec	0.15	0.08	0.03	0.11	0.08	0.03
$\Delta = 1$ min	0.14	0.08	0.04	0.11	0.08	0.03
$\Delta = 5$ min	0.14	0.08	0.04	0.10	0.08	0.03
$\Delta = 30$ min	0.14	0.07	0.03	0.10	0.06	0.02

Table S.VI: Giacomini–White test rejection frequencies for Simulation C. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy, and m is the truncation lag in the long-run variance estimation.

Proxy RV_{t+1}^{Δ}	$P = 500$ Weak	$P = 500$ Strong	$P = 1000$ Weak	$P = 1000$ Strong	$P = 2000$ Weak	$P = 2000$ Strong
			$R = 500$			
$\Delta = 5$ sec	0.00	0.00	0.00	0.00	0.00	0.00
$\Delta = 1$ min	0.00	0.00	0.00	0.01	0.00	0.01
$\Delta = 5$ min	0.00	0.00	0.00	0.01	0.01	0.01
$\Delta = 30$ min	0.00	0.02	0.00	0.03	0.02	0.02
			$R = 1000$			
$\Delta = 5$ sec	0.00	0.00	0.00	0.00	0.00	0.00
$\Delta = 1$ min	0.01	0.02	0.00	0.00	0.00	0.00
$\Delta = 5$ min	0.00	0.03	0.00	0.00	0.00	0.00
$\Delta = 30$ min	0.04	0.05	0.01	0.01	0.01	0.02

Table S.VII: Giacomini–White test rejection indicator disagreement frequencies for Simulation A. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy. Columns headed "Weak" report the difference between the feasible and infeasible tests' rejection frequencies. Columns headed "Strong" report the proportion of simulations in which the feasible and infeasible tests disagree.

Proxy BV_{t+1}^{Δ}	$P = 500$ Weak	$P = 500$ Strong	$P = 1000$ Weak	$P = 1000$ Strong	$P = 2000$ Weak	$P = 2000$ Strong
			$R = 500$			
$\Delta = 5$ sec	0.01	0.01	0.00	0.00	0.01	0.01
$\Delta = 1$ min	0.01	0.04	0.01	0.01	0.01	0.03
$\Delta = 5$ min	-0.01	0.04	0.01	0.05	0.00	0.06
$\Delta = 30$ min	-0.01	0.06	-0.02	0.06	-0.03	0.05
			$R = 1000$			
$\Delta = 5$ sec	0.01	0.01	0.01	0.01	0.00	0.00
$\Delta = 1$ min	0.00	0.04	0.00	0.04	0.02	0.03
$\Delta = 5$ min	0.01	0.04	0.00	0.04	0.01	0.07
$\Delta = 30$ min	-0.01	0.04	-0.02	0.04	-0.04	0.05

Table S.VIII: Giacomini–White test rejection indicator disagreement frequencies for Simulation B. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy. Columns headed "Weak" report the difference between the feasible and infeasible tests' rejection frequencies. Columns headed "Strong" report the proportion of simulations in which the feasible and infeasible tests disagree.

Proxy RC_{t+1}^{Δ}	$P = 500$ Weak	$P = 500$ Strong	$P = 1000$ Weak	$P = 1000$ Strong	$P = 2000$ Weak	$P = 2000$ Strong
			$R = 500$			
$\Delta = 5$ sec	0.00	0.00	0.00	0.00	0.00	0.00
$\Delta = 1$ min	0.00	0.01	-0.01	0.01	0.00	0.00
$\Delta = 5$ min	-0.01	0.02	0.01	0.01	0.01	0.01
$\Delta = 30$ min	0.00	0.03	0.01	0.01	0.02	0.02
			$R = 1000$			
$\Delta = 5$ sec	0.00	0.00	0.00	0.00	0.00	0.00
$\Delta = 1$ min	0.00	0.00	0.00	0.00	0.00	0.00
$\Delta = 5$ min	-0.02	0.02	0.00	0.01	0.00	0.01
$\Delta = 30$ min	0.00	0.03	-0.01	0.02	0.00	0.00

Table S.IX: Giacomini–White test rejection indicator disagreement frequencies for Simulation C. The nominal level is 0.05, R is the length of the estimation sample, P is the length of the prediction sample, Δ is the sampling frequency for the proxy. Columns headed “Weak” report the difference between the feasible and infeasible tests’ rejection frequencies. Columns headed “Strong” report the proportion of simulations in which the feasible and infeasible tests disagree.

References

GIACOMINI, R., AND H. WHITE (2006): "Tests of conditional predictive ability," *Econometrica*, 74(6), 1545–1578.

HANSEN, P. R., A. LUNDE, AND J. M. NASON (2011): "The Model Confidence Set," *Econometrica*, 79(2), pp. 453–497.

JACOD, J., AND P. PROTTER (2012): *Discretization of Processes*. Springer.

KIEFER, N. M., AND T. J. VOGELSANG (2005): "A New Asymptotic Theory for Heteroskedasticity-Autocorrelation Robust Tests," *Econometric Theory*, 21, pp. 1130–1164.

MÜLLER, U. (2012): "HAC Corrections for Strongly Autocorrelated Time Series," Discussion paper, Princeton University.

NEWEY, W. K., AND K. D. WEST (1987): "A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix," *Econometrica*, 55, 703–708.

PROTTER, P. (2004): *Stochastic Integration and Differential Equations*. Springer-Verlag, Berlin, 2nd edn.

ROMANO, J. P., AND M. WOLF (2005): "Stepwise multiple testing as formalized data snooping," *Econometrica*, 73(4), 1237–1282.