Supplemental Material for "Jump Regressions"

Jia Li,* Viktor Todorov,† and George Tauchen‡

August 30, 2016

Abstract

This document contains three supplemental appendices for the main text. Supplemental Appendix A presents additional theoretical results. Supplemental Appendix B contains some numerical analysis for the econometric procedures proposed in the main text. Supplemental Appendix C contains all proofs.

JEL classification: C51, C52, G12.

*Department of Economics, Duke University, Durham, NC 27708; e-mail: jl410@duke.edu.

†Department of Finance, Kellogg School of Management, Northwestern University, Evanston, IL 60208; e-mail: v-todorov@northwestern.edu.

‡Department of Economics, Duke University, Durham, NC 27708; e-mail: george.tauchen@duke.edu.

Supplemental Appendix A: Additional theoretical results

SA.1 Inference when some jumps arrive at deterministic times

In this subsection, we extend the results in the main text to a setting where a subset of jump times can be identified using prior information. Examples of such jump events are the ones caused by pre-scheduled macro announcements (Andersen, Bollerslev, Diebold, and Vega (2003)). In a liquid market, one may expect that price jumps "immediately" after the announcement, so that the announcement time can be used to locate the jump time. Pre-scheduled announcement times are deterministic, which, technically speaking, are excluded from model (2.1) that features random jump arrivals. That being said, inference procedures in this paper can be extended straightforwardly to accommodate fixed jump times as we now show. Here, we use the specification test as an example for a detailed illustration. Modifications to other inference procedures are essentially the same and, hence, will be omitted for brevity.

Formally, we denote with N_t the number of pre-scheduled announcements before time t. We then extend model (2.1) as follows

$$X_t = x_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s + J_t, \quad J_t = \int_0^t \int_{\mathbb{R}} \delta\left(s,u\right)\mu\left(ds,du\right) + \sum_{s \leq N_t} \tilde{\delta}_s, \tag{SA.1}$$

where the jump size process $\tilde{\delta}$ is càdlàg adapted. Economically, $\tilde{\delta}$ reflects the unanticipated information content in the announcement, aggregated among market participants, with $\tilde{\delta}_t = 0$ indicating no surprise. The jump times associated with the announcements are indexed by

$$\mathcal{A}_{\mathcal{D}} \equiv \left\{p \in \mathcal{P}_{\mathcal{D}} : \Delta N_{\tau_p} \neq 0\right\}.$$

We now describe how to modify the results in Section 3.1 so as to accommodate the presence of fixed jump times. It is instructive to start with some intuition. We note that the variable R_p (see (3.2) in the main text) represents asymptotically the (scaled) diffusive disturbance $\Delta_n^{-1/2}\left(\Delta_i^n X - \Delta X_{\tau_p}\right)$ around the jump at time τ_p. In particular, the variable κ_p represents the relative location of the jump time within the sampling interval in the limiting problem. The uniform $[0,1]$ distribution of κ_p is a consequence of the absolute continuity of the compensator of the jump measure μ, and reflects the fact that the econometrician is "locally uninformed" about the exact jump time. On the other hand, if we can use external information (such as announcement times) to pin down the jump time within the sampling interval, we can resolve this layer of statistical uncertainty. Technically, we assume the following.

Assumption S1. *The following conditions hold for each* $p \in \mathcal{A}_{\mathcal{D}}$*: (a)* τ_p *is not an integer multiple of* Δ_n*; (b)* $\tau_p/\Delta_n - \lfloor \tau_p/\Delta_n \rfloor$ *converges to some known constant* $\kappa_p \in [0,1]$.

Part (a) of Assumption S1 is a convenient device to avoid uninteresting technical issues that arise when the pre-scheduled announcement is at $i\Delta_n$ for some i. Part (b) of Assumption S1 formalizes the form of prior knowledge about jump times. For example, the assumption that the price jumps immediately after the announcement time corresponds to $\kappa_p = 0$.

For random jump times indexed by $p \in \mathcal{P}_{\mathcal{D}} \setminus \mathcal{A}_{\mathcal{D}}$, we draw κ_p from a uniform $[0,1]$ distribution like in Section 3.1 of the main text and for $p \in \mathcal{A}_{\mathcal{D}}$ we set κ_p at its known value. We then define R_p and ς_p respectively using (3.2) and (3.12) for all $p \in \mathcal{P}_{\mathcal{D}}$. The variables $(\tilde{\kappa}_i, \tilde{R}_{n,i}, \tilde{\varsigma}_{n,i})$ in Algorithm 1 are modified analogously. The formal extension of Theorem 1 is stated as follows.

Theorem S1. *Suppose (SA.1), Assumption S1 and the conditions in Theorem 1. Under aforemenioned modifications for the definitions of* $(\kappa_p, R_p, \varsigma_p)$ *and* $(\tilde{\kappa}_i, \tilde{R}_{n,i}, \tilde{\varsigma}_{n,i})$*, the statements in Theorem 1 hold true.*

SA.2 Higher-order asymptotics for the optimally weighted estimator

In this subsection, we proceed with designing refined confidence sets of the optimally weighted estimator based on a higher-order asymptotic expansion. To motivate, we observe that while the optimally weighted estimator $\hat{\beta}_n(\mathcal{D}, w^*)$ depends on the spot covariance estimates $(\hat{c}_{n,i-}, \hat{c}_{n,i+})$, the sampling variability of the latter is not reflected in the asymptotic distribution described by Theorem 2. The reason is that these estimates enter only the weights and their sampling errors are annihilated in the second-order asymptotics. In finite samples, the sampling variability of the spot covariance estimates may still have some effect, because the latter enjoy only a nonparametric convergence rate. To account for such effects, we need a refined characterization of the asymptotic behavior of the optimally weighted estimator, so we proceed to derive its higher-order expansion. Based on this expansion, we provide a refinement to the confidence interval construction described in Algorithm 2.

We first need some additional regularity on the spot volatility σ, namely that it is an Itô semimartingale.

Assumption S2. *The process* σ *is an Itô semimartingale of the form*

$$\begin{aligned} vec(\sigma_t) &= vec(\sigma_0) + \int_0^t \tilde{b}_s ds + \int_0^t \tilde{\sigma}_s d\widetilde{W}_s + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s,u) 1_{\{\|\tilde{\delta}(s,u)\|>1\}} \tilde{\mu}(ds,du) \\ &\quad + \int_0^t \int_{\mathbb{R}} \tilde{\delta}(s,u) 1_{\{\|\tilde{\delta}(s,u)\|\leq 1\}} (\tilde{\mu} - \tilde{\nu})(ds,du), \end{aligned}$$

where the processes $\tilde{b}$ *and* $\tilde{\sigma}$ *are locally bounded and take values respectively in* $\mathbb{R}^4$ *and* $\mathbb{R}^{4\times 4}$, $\widetilde{W}$ *is a 4-dimensional Brownian motion,* $\tilde{\delta} : \Omega \times \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}^4$ *is a predictable function,* $\tilde{\mu}$ *is a Poisson random measure with compensator* $\tilde{\nu}$ *of the form* $\tilde{\nu}(dt, du) = dt \otimes \tilde{\lambda}(du)$ *for some* σ*-finite measure* $\tilde{\lambda}$*. Moreover, there exists a localizing sequence of stopping times* $(T_m)_{m\geq 1}$ *and* $\tilde{\lambda}$*-integrable functions* $(\tilde{\Gamma}_m)_{m\geq 1}$*, such that* $\|\tilde{\delta}(\omega, t, u)\|^2 \wedge 1 \leq \tilde{\Gamma}(u)$ *for all* $\omega \in \Omega$*,* $t \leq T_m$ *and* $u \in \mathbb{R}$*.*

Assumption S2 is needed for characterizing the stable convergence of the spot covariance estimates. This assumption is fairly unrestrictive and is satisfied by many models in finance. In particular, it allows for "leverage effect," that is, the Brownian motions W and $\widetilde{W}$ can be correlated. Moreover, Assumption S2 allows for volatility jumps, and it does not restrict their activity and dependence with the price jumps. However, this assumption does rule out certain long-memory volatility models driven by the fractional Brownian motion (see Comte and Renault (1996)).

We now present the higher-order asymptotic expansion for the optimally weighted estimator. We need some additional notation for this. Let v denote the spot variance of the residual $Y - \beta_0 Z$. That is,

$$v_t \equiv (-\beta_0, 1)\, c_t\, (-\beta_0, 1)^\top . \tag{SA.2}$$

Further, let $\left(\xi'_{p-}, \xi'_{p+}\right)_{p\geq 1}$ be a collection of mutually independent random variables which are also independent of $\mathcal{F}$ and $(\kappa_p, \xi_{p-}, \xi_{p+})_{p\geq 1}$, such that ξ'_{p-} and ξ'_{p+} are scalar standard normal variables. We set for $p \geq 1$,

$$\phi_p \equiv \frac{(-\beta_0, 1)\left(c_{\tau_p -} + c_{\tau_p}\right)(-\beta_0, 1)^\top}{2}, \quad F_p \equiv \frac{v_{\tau_p -}\xi'_{p-} + v_{\tau_p}\xi'_{p+}}{\sqrt{2}}. \tag{SA.3}$$

To guide intuition, we note that the variable F_p captures the sampling variability for approximating the average residual volatility $(v_{\tau_p -} + v_{\tau_p})/2$. We further note that $\phi_p^{-1} = w^*(c_{\tau_p -}, c_{\tau_p}, \beta_0)$, so the limit variable $\zeta_\beta(\mathcal{D}, w^*)$ defined in (4.3) can be rewritten as

$$\zeta_\beta(\mathcal{D}, w^*) \equiv \frac{\sum_{p\in\mathcal{P}_\mathcal{D}} \Delta Z_{\tau_p} \varsigma_p / \phi_p}{\sum_{p\in\mathcal{P}_\mathcal{D}} \Delta Z_{\tau_p}^2 / \phi_p}. \tag{SA.4}$$

Theorem S2. *Let* $k_n \asymp \Delta_n^{-a}$ *for some constant* $a \in (0, 1/2)$*. Suppose Assumptions 1, 2 and S2 hold for* $\varpi \in (a/4, 1/2)$*. Then we have the following expansion for the optimally weighted estimator:*

$$\Delta_n^{-1/2}\left(\hat{\beta}_n(\mathcal{D}, w^*) - \beta_0\right) = \zeta^*_{n,\beta}(\mathcal{D}) + k_n^{-1/2} H^*_{n,\beta}(\mathcal{D}) + o_p(k_n^{-1/2}), \tag{SA.5}$$

for some sequences of variables $\zeta^*_{n,\beta}(\mathcal{D})$ *and* $H^*_{n,\beta}(\mathcal{D})$ *satisfying*

$$\left(\zeta^*_{n,\beta}(\mathcal{D}), H^*_{n,\beta}(\mathcal{D})\right) \xrightarrow{\mathcal{L}\text{-}s} \left(\zeta_\beta(\mathcal{D}, w^*), H^*_\beta(\mathcal{D})\right), \tag{SA.6}$$

where

$$H^*_\beta(\mathcal{D}) \equiv \frac{\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p} F_p}{\phi_p^2}\right)\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}\varsigma_p}{\phi_p}\right) - \left(\sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}\varsigma_p F_p}{\phi_p^2}\right)\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p}}{\phi_p}\right)}{\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p}}{\phi_p}\right)^2}. \tag{SA.7}$$

The leading term $\zeta^*_{n,\beta}(\mathcal{D})$ in (SA.5) is what drives the convergence in Theorem 2. The higher-order term $k_n^{-1/2}H^*_{n,\beta}(\mathcal{D})$ is $O_p(k_n^{-1/2})$ and hence is asymptotically dominated by $\zeta^*_{n,\beta}(\mathcal{D})$. The limiting variable $H^*_\beta(\mathcal{D})$ involves both F_p and ς_p, which capture respectively the sampling variability that arise from the estimation of the spot covariance and the estimation of jumps.

Because of the higher-order asymptotic effect played by $\hat{c}_{n,i\pm}$ in the efficient beta estimation, the user has a lot of freedom in setting the block size k_n. Indeed, as seen from Theorem S2, we need only $k_n \asymp \Delta_n^{-a}$ with a in the wide range of $(0, 1/2)$. This is unlike the block-based volatility estimators, see e.g., Jacod and Rosenbaum (2013), where one has significantly less freedom in choosing k_n. Having the refined asymptotic result in Theorem S2 helps since if k_n is relatively small, the higher-order term $k_n^{-1/2}H^*_{n,\beta}(\mathcal{D})$ might have nontrivial finite sample effect.

For concreteness, we describe in Algorithm 3 a finite-sample correction for the CIs described in Algorithm 2, based on Theorem S2, where we set for $i \in \mathcal{I}'_n(\mathcal{D})$,

$$\hat{\phi}_{n,i} \equiv \frac{\left(-\tilde{\beta}_n, 1\right)(\hat{c}_{n,i-} + \hat{c}_{n,i+})\left(-\tilde{\beta}_n, 1\right)}{2}, \quad \hat{v}_{n,i\pm} \equiv \left(-\tilde{\beta}_n, 1\right)\hat{c}_{n,i\pm}\left(-\tilde{\beta}_n, 1\right)^\top. \tag{SA.8}$$

The proof of Theorem 2(c) can be easily adapted to show that $\mathrm{CI}_n^{*\alpha}$ defined in Algorithm 3 has asymptotic level $1-\alpha$, that is, $\mathbb{P}(\beta_0 \in \mathrm{CI}_n^\alpha) \to 1-\alpha$; the details are omitted for brevity.

ALGORITHM 3. (1) Simulate $(\tilde{\varsigma}_{n,i})_{i\in\mathcal{I}'_n(\mathcal{D})}$ as in step 1 of Algorithm 1. Simulate $(\tilde{\xi}'^n_{i-}, \tilde{\xi}'^n_{i+})_{i\in\mathcal{I}'_n(\mathcal{D})}$ consisting of independent copies of $\left(\xi'_{p-}, \xi'_{p+}\right)$. Set

$$\tilde{F}_{n,i} \equiv \frac{\hat{v}_{n,i-}\tilde{\xi}'_{n,i-} + \hat{v}_{n,i+}\tilde{\xi}'_{n,i+}}{\sqrt{2}}.$$

(2) Compute

$$\begin{aligned}
&\tilde{\zeta}^*_{n,\beta}(\mathcal{D}) \\
&\equiv \left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z\,\tilde{\varsigma}_{n,i}}{\hat{\phi}_{n,i}}\right) \Big/ \left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z^2}{\hat{\phi}_{n,i}}\right) \\
&\quad + \frac{\left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z^2 \tilde{F}_{n,i}}{\hat{\phi}^2_{n,i}}\right)\left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z\,\tilde{\varsigma}_{n,i}}{\hat{\phi}_{n,i}}\right) - \left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z\,\tilde{\varsigma}_{n,i}\,\tilde{F}_{n,i}}{\hat{\phi}^2_{n,i}}\right)\left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z^2}{\hat{\phi}_{n,i}}\right)}{k_n^{1/2}\left(\sum_{i\in\mathcal{I}'_n(\mathcal{D})} \frac{\Delta_i^n Z^2}{\hat{\phi}_{n,i}}\right)^2}.
\end{aligned}$$

(3) Generate a large number of Monte Carlo simulations in the first two steps and set $cv_{n,\beta}^{\alpha/2}$ as the $(1-\alpha/2)$-quantile of $\tilde{\zeta}_{n,\beta}^{*}(\mathcal{D},w)$ in the Monte Carlo sample. Set the $1-\alpha$ level two-sided CI as $\mathrm{CI}_n^{*\alpha}=[\hat{\beta}_n(\mathcal{D},w^*)-\Delta_n^{1/2}cv_{n,\beta}^{\alpha/2},\hat{\beta}_n(\mathcal{D},w^*)+\Delta_n^{1/2}cv_{n,\beta}^{\alpha/2}]$. □

SA.3 Adaptive estimation under a common-beta restriction

In Section 4.2 of the main text, we have shown that the optimally weighted jump beta estimator attains the adaptive bound (4.21) when the diffusive beta (i.e., β_t^c) coincides with the jump beta. If this parametric restriction is known *a priori*, it can be exploited to construct a more efficient estimator of the (common) beta parameter by using both jump and diffusive returns. In this appendix, we show that the common beta can be estimated adaptively with respect to nonparametric nuisance components in the model. This result is related to the recent work of Li, Todorov, and Tauchen (2016), which considers the adaptive estimation of the constant diffusive beta. Li, Todorov, and Tauchen (2016) treat the jump component as a nonparametric nuisance. By contrast, here we also exploit the information content of the constant jump beta restriction.

Below, we maintain the same setting as in Section 4.2 of the main text with the additional restriction that the diffusive beta equals the jump beta, that is,

$$c_{ZY,t}=\beta_0\,c_{ZZ,t},\quad \text{all}\quad t\in[0,T]\,. \tag{SA.9}$$

Like in Section 4.2, we fix $\mathcal{D}=[0,T]\times\mathbb{R}_*$ and suppress it in future notation. Model (2.1) can then be represented as

$$\begin{cases} dZ_t=b_{Z,t}dt+\sqrt{c_{ZZ,t}}dW_{Z,t}+dJ_{Z,t},\\ dY_t=b_{Y,t}dt+\beta\sqrt{c_{ZZ,t}}dW_{Z,t}+\sqrt{v_t^c}dW_{Y,t}+\beta dJ_{Z,t}+d\epsilon_t,\end{cases} \tag{SA.10}$$

where W_Z and W_Y are univariate independent Brownian motions, ϵ is a pure jump process that contains the Y-specific jumps, and the spot idiosyncratic variance v_t^c is connected with c_t via $v_t^c=c_{YY,t}-c_{ZY,t}^2/c_{ZZ,t}$.

We proceed as follows. We first derive the efficiency bound for estimating β in a parametric submodel of (SA.10) where the only unknown parameter is β. The resulting efficiency bound is what one would attain if the processes $(b,c_{ZZ},v^c,J_Z,\epsilon)$ were observable. We then provide a feasible estimator that attains this efficiency bound in the original semiparametric model. The proposed estimator is therefore adaptive with respect to the nonparametric nuisance components $(b,c_{ZZ},v^c,J_Z,\epsilon)$. Henceforth, we denote the law of $(\Delta_i^nX)_{1\le i\le n}$ in this submodel by P_β^n.

Theorem S3, below, shows that the aforementioned submodel satisfies the LAMN property and characterizes the information matrix for the estimation of β. To fit the current setting, we modify slightly Assumption 4 in the main text as follows.

Assumption S3. *We have Assumption 1 and the processes* $(b_t)_{t\geq 0}$, $(\sigma_t)_{t\geq 0}$ *and* $(J_t)_{t\geq 0}$ *are independent of* $(W_t)_{t\geq 0}$*, and the joint law of* $(b, c_{ZZ}, v^c, J_Z, \epsilon)$ *does not depend on* β.

Theorem S3. *Under Assumptions S3 and 5, the sequence* $(P_\beta^n : \beta \in \mathbb{R})$ *satisfies the LAMN property at* $\beta = \beta_0$ *with information* $\int_0^T (1/v_s^c)\, d[Z,Z]_s$.

The proofs for all results in this section are given in Supplemental Appendix C.

From Theorem S3 and the conditional convolution theorem (Jeganathan (1982, 1983)), we deduce that the efficiency bound for estimating β_0 in the adaptive case is given by the inverse of the information, that is,

$$\left(\int_0^T \frac{d\,[Z,Z]_s}{v_s^c}\right)^{-1}. \tag{SA.11}$$

Next, we construct an estimator for β_0 that attains this bound, for which we need some additional notation. For each i, we set

$$\hat{c}_{n,i} \equiv \begin{pmatrix} \hat{c}_{ZZ,n,i} & \hat{c}_{ZY,n,i} \\ \hat{c}_{ZY,n,i} & \hat{c}_{YY,n,i} \end{pmatrix} \equiv \frac{1}{k_n \Delta_n} \sum_{j=1}^{k_n} (\Delta_{i+j}^n X)(\Delta_{i+j}^n X)^\top 1_{\{-v_n' \leq \Delta_{i+j}^n X \leq v_n'\}},$$

and then

$$\hat{v}_{n,i}^c \equiv \hat{c}_{YY,n,i} - \hat{c}_{ZY,n,i}^2 / \hat{c}_{ZZ,n,i}.$$

The adaptive estimator for β_0 is given by

$$\hat{\beta}_n^\star \equiv \frac{(1-\frac{3}{k_n})\Delta_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \frac{\hat{c}_{ZY,n,i}}{\hat{v}_{n,i}^c} + \sum_{i\in\mathcal{I}_n'} \frac{\Delta_i^n Z \Delta_i^n Y}{\hat{v}_{n,i}^c}}{(1-\frac{3}{k_n})\Delta_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \frac{\hat{c}_{ZZ,n,i}}{\hat{v}_{n,i}^c} + \sum_{i\in\mathcal{I}_n'} \frac{(\Delta_i^n Z)^2}{\hat{v}_{n,i}^c}}. \tag{SA.12}$$

The factor $1-3/k_n$ in (SA.12) is used to correct a nonlinearity bias in the estimation of integrated volatility functionals as in, for example, Jacod and Rosenbaum (2013) and Li, Todorov, and Tauchen (2016).

Theorem S4, below, establishes the stable convergence in law of $\hat{\beta}_n^\star$. In particular, it shows that $\hat{\beta}_n^\star$ attains the adaptive bound given by (SA.11).

Theorem S4. *Suppose (a) Assumptions 1, 5 and S2 hold; (b)* $k_n \asymp \Delta_n^{-\gamma}$ *such that* $\gamma \in (1/3, 1/2)$ *and* $(1-\gamma)/2 \leq \varpi < 1/2$. *Then*

$$\Delta_n^{-1/2}\left(\hat{\beta}_n^\star - \beta_0\right) \xrightarrow{\mathcal{L}\text{-}s} \mathcal{MN}\left(0, \left(\int_0^T \frac{d\,[Z,Z]_s}{v_s^c}\right)^{-1}\right).$$

Comment. Theorem S4 does not require the independence conditions in Assumption S3, which are only needed for deriving the LAMN property and the efficiency bound. Assumption S2 is needed for deriving the stable convergence of the integrated volatility functionals $\Delta_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \hat{c}_{ZY,n,i}/\hat{v}_{n,i}^c$ and $\Delta_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \hat{c}_{ZZ,n,i}/\hat{v}_{n,i}^c$.

SA.4 Jump regressions with microstructure noise and irregular sampling

In this subsection, we consider jump regressions under a setting with microstructure noise and irregular asynchronous sampling. Section SA.4.1 describes the setting. In Section SA.4.2, we propose and analyze a new estimator that is robust to these complications. The proofs are in Supplemental Appendix C. To simplify the exposition, we set $\mathcal{D} = [0, T] \times \mathbb{R}_*$ and suppress it henceforth. In other words, the jump regression relationship $\Delta Y_\tau = \beta_0 \Delta Z_\tau$ is supposed to hold for all jump times τ of Z.

SA.4.1 The setting with microstructure noise and irregular sampling

We extend the jump regression setting in the main text in three directions. First, the observed prices are contaminated by measurement error. The measurement error is often referred to as "microstructure noise," which may be attributed to various trading frictions. Second, the prices are sampled at irregular times, i.e., the sampling interval is no longer constant. Third, the sampling times for Y and Z are possibly asynchronous.

To describe the formal setting, we start with the sampling scheme. We consider an array of deterministic times $t(n,i)$, $i \geq 0$, that is increasing in i. Our analysis for this deterministic sampling scheme can be equivalently considered as an analysis conditioning on a strictly exogenous random sampling design. We allow these times to be irregularly spaced, that is, $\Delta_{n,i} \equiv t(n,i) - t(n,i-1)$ may depend on i. Below, we refer to $(t(n,i))_{i\geq 0}$ as the *sampling basis* for simplicity. The processes Y and Z are sampled at two deterministic subsets of the sampling basis, indexed by $(i_{Y,n,k})_{k\geq 0}$ and $(i_{Z,n,k})_{k\geq 0}$, respectively.[1] Asynchronous sampling arises when these subsets are different. To simplify notation, we denote the sampling times of Y and Z using

$$t(Y,n,k) \equiv t(n, i_{Y,n,k}), \quad t(Z,n,k) \equiv t(n, i_{Z,n,k}).$$

As an additional generalization of our original setup in the text, we assume that only noisy observations of Y and Z are available at the sampling times

$$Y'_{t(Y,n,k)} = Y_{t(Y,n,k)} + \varepsilon'_{Y,t(Y,n,k)}, \quad Z'_{t(Z,n,k)} = Z_{t(Z,n,k)} + \varepsilon'_{Z,t(Z,n,k)},$$

where ε'_Y and ε'_Z are the noise terms. We assume the following for the noise.

Assumption S4. *We have $\varepsilon'_{Y,t} = \sqrt{A_{Y,t}}\varepsilon_{Y,t}$ and $\varepsilon'_{Z,t} = \sqrt{A_{Z,t}}\varepsilon_{Z,t}$ such that (a) the processes A_Y and A_Z are continuous, $(\mathcal{F}_t)$-adapted and locally bounded; (b) the variables $\varepsilon_{Y,t}$ and $\varepsilon_{Z,t}$, $t \geq 0$, are*

[1] Alternatively, if we start with the sampling times of Y and Z, we can set the sampling basis as the union of these sampling times (i.e., the collection of times at which at least one process is sampled).

mutually independent and independent of $\mathcal{F}$ with zero mean and unit variance, and with finite pth moment for all $p \geq 1$.

The essential part of Assumption S4 is that the noise terms $(\varepsilon'_{t(Y,n,k)}, \varepsilon'_{t(Z,n,k)})_{k\geq 0}$ are $\mathcal{F}$-conditionally independent with zero mean. For the results below, we only need $\varepsilon_{Y,t}$ and $\varepsilon_{Z,t}$ to have finite moments up to a certain order; assuming finite moments for all orders is merely for technical convenience. Finally, we note that the noise terms are allowed to be heteroskedastic and serially dependent through their volatility processes $\sqrt{A_Y}$ and $\sqrt{A_Z}$.

We now turn to the regularity conditions for the sampling scheme. Below, we denote

$$\delta_{Y,n,k} \equiv i_{Y,n,k} - i_{Y,n,k-1}, \quad \delta_{Z,n,k} \equiv i_{Z,n,k} - i_{Z,n,k-1},$$

which are the numbers of fine sampling intervals (i.e., those determined by the sampling basis) contained in the kth sampling intervals of Y and Z, respectively.

Assumption S5. *The following conditions hold for a real sequence $\Delta_n \to 0$ and functions f, ϕ_Y, $\phi_Z : [0,T] \mapsto (0,\infty)$: (a) $\Delta_{n,i}/\Delta_n$ is uniformly bounded; (b) for each $t \in (0,T)$ and any integer sequence L_n with $L_n \to \infty$ and $L_n\Delta_n \to 0$,*

$$\frac{1}{L_n} \sum_{i=\max\{i:t(n,i)<t\}-L_n}^{\min\{i:t(n,i)\geq t\}+L_n} \left| \frac{\Delta_{n,i}}{\Delta_n} - f(t) \right| = o(1); \tag{SA.13}$$

(c) $\delta_{Y,n,k}$ and $\delta_{Z,n,k}$ are uniformly bounded; (d) for each $t \in (0,T)$ and any integer sequence L_n with $L_n \to \infty$ and $L_n\Delta_n \to 0$,

$$\begin{cases} \dfrac{1}{L_n} \displaystyle\sum_{k=\max\{k:t(Y,n,k)<t\}-L_n}^{\min\{k:t(Y,n,k)\geq t\}+L_n} |\delta_{Y,n,k} - \phi_Y(t)| = o(1), \\ \dfrac{1}{L_n} \displaystyle\sum_{k=\max\{k:t(Y,n,k)<t\}-L_n}^{\min\{k:t(Y,n,k)\geq t\}+L_n} |\delta_{Z,n,k} - \phi_Z(t)| = o(1). \end{cases} \tag{SA.14}$$

Part (a) of Assumption S5 ensures that the mesh of the sampling basis shrinks to zero at least as fast as Δ_n. Part (b) formalizes the notion that the irregularity of the sampling basis is "locally moderate." This condition allows the sampling intensity (measured by $1/f(t)$) to vary across different fixed times. It requires that, within a shrinking neighborhood of each time t, "most" sampling intervals are of size close to $f(t)\Delta_n$. Note that parts (a) and (b) hold trivially when the sampling basis is regular. Part (c) requires that the sampling grids of Y and Z are not "too coarse" relative to the sampling basis, and part (d) further requires that the coarse sampling of each process is "moderately irregular." These two conditions hold trivially if the sampling of Y and Z are synchronized: in this case, we can set the sampling basis as the common sampling grid of Y and Z, so that (SA.14) hold with the $o(1)$ terms being identically zero.

SA.4.2 The estimator and its asymptotic properties

We propose a pre-averaging method to deal with the noisy data: we first smooth locally the noisy returns and then conduct the jump regression. The pre-averaging method has been introduced by Podolskij and Vetter (2009), Jacod, Li, Mykland, Podolskij, and Vetter (2009) for the estimation of integrated variance using noisy high-frequency data. The inference for jumps is limited to high-order jump power variations (Jacod, Podolskij, and Vetter (2010), Li (2013)) for univariate regularly sampled processes. The current setting is hence distinct from prior work in a non-trivial way.

In order to pre-smooth the noisy returns, we consider a weight function $g : \mathbb{R} \mapsto \mathbb{R}_+$ that is supported on $[0,1]$, continuously differentiable with Lipschitz continuous derivative and is strictly positive on $(0,1)$. We also consider an integer sequence h_n of smoothing bandwidth parameters that satisfies

$$h_n = \lfloor \theta \Delta_n^{-1/2} \rfloor, \quad \text{for some } \theta \in (0, \infty). \tag{SA.15}$$

The pre-averaged returns are defined as locally weighted averages of observed noisy returns. For each $i \geq 0$ (which is the index of the sampling basis), we set

$$\begin{cases} \bar{Z}'_{n,i} \equiv \sum_k g\left(\frac{i_{Z,n,k} - i}{h_n}\right)\left(Z'_{t(Z,n,k)} - Z'_{t(Z,n,k-1)}\right), \\ \bar{Y}'_{n,i} \equiv \sum_k g\left(\frac{i_{Y,n,k} - i}{h_n}\right)\left(Y'_{t(Y,n,k)} - Y'_{t(Y,n,k-1)}\right). \end{cases} \tag{SA.16}$$

By construction, these pre-averaged returns are synchronized with respect to the sampling basis.

We adapt the (unweighted) jump regression estimator using the pre-averaged returns as follows:

$$\hat{\beta}'_n \equiv \frac{\sum_{i\geq 0: t(n,i+h_n)\leq T} \bar{Z}'_{n,i}\bar{Y}'_{n,i} 1_{\{|\bar{Z}'_{n,i}|>\bar{v}_n\}}}{\sum_{i\geq 0: t(n,i+h_n)\leq T} \bar{Z}'^2_{n,i} 1_{\{|\bar{Z}'_{n,i}|>\bar{v}_n\}}}. \tag{SA.17}$$

The truncation threshold $\bar{v}_n$ plays the same role as v_n in (2.12). In the current context, we suppose that

$$\bar{v}_n \asymp \Delta_n^{\varpi'}, \quad \text{for some } \varpi' \in (0, 1/4).$$

The intuition for the condition $\varpi' < 1/4$ is that, in the absence of jumps, the order of magnitude of each pre-averaged return is $O_p(\Delta_n^{1/4})$. Hence, the truncation threshold $\bar{v}_n$ can be used to separate consistently the smoothed jump returns from the diffusive ones.

Theorem S5, below, describes the asymptotic distribution of $\hat{\beta}'_n$, which requires some additional notation. Let $(T_m)_{m\geq 1}$ be the successive jump times of the Poisson process $t \mapsto \mu([0,t] \times \mathbb{R}_*)$. We consider variables $(\zeta_{m-}, \zeta_{m+}, \zeta'_{Y,m}, \zeta'_{Z,m})_{m\geq 1}$ defined on an extension of the space $(\Omega, \mathcal{F}, \mathbb{P})$ which, conditionally on $\mathcal{F}$, are mutually independent centered Gaussian with variances given by

$$\begin{cases} \mathbb{E}\left[\zeta_{m-}^{2}|\mathcal{F}\right] = \theta v_{T_m-} f\left(T_m\right) \int_{-1}^{0} \left(\int g\left(s-u\right) g\left(s\right) ds\right)^{2} du, \\ \mathbb{E}\left[\zeta_{m+}^{2}|\mathcal{F}\right] = \theta v_{T_m} f\left(T_m\right) \int_{0}^{1} \left(\int g\left(s-u\right) g\left(s\right) ds\right)^{2} du, \\ \mathbb{E}\left[\zeta_{Y,m}^{\prime 2}|\mathcal{F}\right] = \theta^{-1} A_{Y,T_m} \phi_Y(T_m) \int \left(\int_{0}^{1} g\left(s\right) g'\left(s+u\right) ds\right)^{2} du, \\ \mathbb{E}\left[\zeta_{Z,m}^{\prime 2}|\mathcal{F}\right] = \theta^{-1} A_{Z,T_m} \phi_Z(T_m) \int \left(\int_{0}^{1} g\left(s\right) g'\left(s+u\right) ds\right)^{2} du. \end{cases} \tag{SA.18}$$

Theorem S5. *Under Assumptions 1, S4 and S5,*

$$\Delta_n^{-1/4}\left(\hat{\beta}_n' - \beta_0\right) \xrightarrow{\mathcal{L}\text{-}s} \frac{\sum_{m\geq 1: T_m \leq T} \Delta Z_{T_m}\left(\zeta_{m-} + \zeta_{m+} + \zeta_{Y,m}' - \beta_0 \zeta_{Z,m}'\right)}{\left(\int_0^1 g\left(s\right)^2 ds\right) \sum_{m\geq 1: T_m \leq T} \Delta Z_{T_m}^2}.$$

Theorem S5 shows that $\hat{\beta}_n'$ is asymptotically centered at the true value β_0 with convergence rate $\Delta_n^{-1/4}$. The limiting distribution is $\mathcal{F}$-conditionally centered Gaussian and its asymptotic variance depends on the spot variances of the efficient prices, the spot variances of the noise terms, as well as the local sampling intensities, around jump times.

Supplemental Appendix B: Numerical experiments

We now assess the efficiency gain provided by our efficient estimation procedure and we further examine the finite-sample performance of the asymptotic theory developed in the paper in realistically calibrated simulations.

SB.1 Relative efficiency of beta estimation

We start with gauging the efficiency gains of our efficient estimation procedure in empirically relevant scenarios. As seen from the asymptotic theory in Section 4, the sampling variability in the estimation of beta depends on the volatility processes of Y and Z, as well as the number and sizes of jumps. Therefore, to make the efficiency comparisons practically relevant, we calibrate the numerical environment using estimates of these quantities from our empirical application in Section 5. In particular, in the calculation of the asymptotic variances we will use the detected jumps in our empirical data sets and we will further set $c_t = \frac{1}{2}\left(\hat{c}_{n,i-} + \hat{c}_{n,i+}\right)$ for $t \in ((i-1)\Delta_n, i\Delta_n]$.

We conduct two efficiency comparisons. First, in order to have a general sense about how accurate the jump beta can be estimated, we compare the efficiency bound for estimating the jump beta, which is attained by our optimally weighted estimator, with that for estimating the continuous beta. Under the assumption that β_t^c (recall (4.15)) is a constant, Li, Todorov, and Tauchen (2016) show that the sharp lower efficiency bound for estimating the continuous beta is $(\int_0^T c_{ZZ,s}/v_s^c ds)^{-1}$. For the 9 assets studied in our empirical application, we find that estimating the continuous beta is 6 to 7 times more accurate, measured by the $\mathcal{F}$-conditional asymptotic standard deviation, than estimating the jump beta from the same data set. We note that this is in spite of the fact that the jump beta estimation is (effectively) based on 74 jump returns detected in the sample while the continuous beta is based on the remaining of the total 56,886 high-frequency increments. The intuition is that, although the number of jump returns is small, these returns have much higher signal-to-noise ratio than their diffusive counterparts for the estimation of betas.[2]

Our second efficiency comparison concerns the role of the optimal weighting in the efficient estimation of jump beta. That is, we are interested in the efficiency gains from using the optimal weight function $w^*(\cdot)$ over the case of no weighting, corresponding to $w(\cdot) = 1$, which has been done in prior work such as Gobbi and Mancini (2012) and Todorov and Bollerslev (2010). The comparison is, again, implemented using estimates of jumps and volatility paths as explained above.

[2]Of course the above efficiency comparison is based on the premise that the continuous beta and the jump beta remain constant over the same time interval. Results from tests for temporal stability of continuous betas in Reiß, Todorov, and Tauchen (2015) indicate that continuous betas vary even over short periods such as months. This is unlike our empirical findings for the jump betas, reported in Section 5, for which we find temporal stability over time spans of years.

Table S1: Relative Efficiency of Jump Beta Estimation

Asset	$\frac{a.s.e.\hat{\beta}_n}{a.s.e.\tilde{\beta}_n}$	**Asset**	$\frac{a.s.e.\hat{\beta}_n}{a.s.e.\tilde{\beta}_n}$	**Asset**	$\frac{a.s.e.\hat{\beta}_n}{a.s.e.\tilde{\beta}_n}$
XLB	0.68	XLI	0.69	XLU	0.55
XLE	0.61	XLK	0.61	XLV	0.61
XLF	0.44	XLP	0.66	XLY	0.61

Note: Calculations of the asymptotic standard error (a.s.e.) are based on detected jumps and volatility paths extracted from the empirical data set discussed in Section 5. The efficient estimator $\hat{\beta}_n$ and the unweighted estimator $\tilde{\beta}_n$ correspond to $\hat{\beta}_n([0,T]\times\mathbb{R}_*, w)$ with $w(\cdot) = w^*(\cdot)$ and $w(\cdot) = 1$, respectively.

Table S1 reports the relative efficiency of the unweighted estimator versus the efficient estimator. We see that the optimal weighting indeed provides nontrivial efficiency gains, with the minimal gain being 31% among all assets in our sample. Not surprisingly, these gains vary across assets and are bigger for those with more volatility variations over the time period. We note that optimal weighting is of particular relevance for the jump beta estimation: by their very nature, jumps are rare events and, hence, for estimating the jump beta we naturally pool information from distinct time periods which typically have very different volatility levels.

SB.2 Monte Carlo

We proceed with assessing the performance of our inference techniques on simulated data from the following model

$$\begin{aligned} dZ_t &= \sigma_t dL_t, \quad dY_t = \beta_t dZ_t + \sigma_t d\widetilde{L}_t, \quad \sigma_t^2 = V_{1,t} + V_{2,t}, \\ dV_{1,t} &= 0.0105(0.5 - V_{1,t})dt + 0.0717\sqrt{V_{1,t}}dB_{1,t}, \\ dV_{2,t} &= 0.6931(0.5 - V_{2,t})dt + 0.5828\sqrt{V_{2,t}}dB_{2,t}, \end{aligned} \tag{SB.1}$$

where L and $\widetilde{L}$ are two independent Lévy processes with characteristic triplets $\left(0, 1, \frac{e^{-|x|}}{24}\right)$ and $\left(0, \frac{1}{\sqrt{2}}, \frac{e^{-|x|}}{96}\right)$ with respect to the zero truncation function; (B_1, B_2) is a two-dimensional standard Brownian motion independent of L and $\widetilde{L}$. This means that L and $\widetilde{L}$ are Brownian motions plus compound Poisson jumps, with jumps having double-exponential distribution. The frequency and jump size distributions are calibrated to mimic those in the real data that we are going to use. The stochastic volatility has a two-factor affine volatility structure, with the first factor being slow mean-reverting (with half-life of sixty-six days) and the second factor being fast mean-reverting (with half-life of one day). Finally, for the beta process we consider

$$\begin{cases} \beta_t = 1, \quad \text{for } t \in [0,T], & \text{under } H_0 \text{ (null hypothesis)}, \\ d\beta_t = 0.005(1-\beta_t)dt + 0.005\sqrt{\beta_t}d\widetilde{B}_t, & \text{under } H_a \text{ (alternative hypothesis)}, \end{cases} \tag{SB.2}$$

where $\widetilde{B}$ is a Brownian motion independent of B, L and $\widetilde{L}$. The unconditional mean of β_t under the alternative hypothesis is 1 and the expected range of the process $(\beta_t)_{t\geq 0}$ over the interval of estimation is approximately 0.2.

We set $T = 1,500$ days (our unit of time is a trading day), and consider two sampling frequencies: $\Delta_n = 1/38$ which corresponds to sampling every 10 minutes in a 6.5 hours trading day, and $\Delta_n = 1/81$ which corresponds to sampling every 5 minutes. We experiment with two values of k_n for each of the sampling frequency in order to check the sensitivity of the inference techniques with respect to this tuning parameter. Finally, as is typical in truncation-based methods, we select the truncation threshold in the following data-driven way. For the increment $\Delta_i^n Z$ with $i = \lfloor (t-1)/\Delta_n \rfloor + 1, ..., \lfloor t/\Delta_n \rfloor$, we set

$$v_n = 4 \times \sqrt{BV_t} \times \Delta_n^{0.49}, \quad BV_t = \frac{\lfloor 1/\Delta_n \rfloor}{\lfloor 1/\Delta_n \rfloor - 1} \frac{\pi}{2} \sum_{\lfloor (t-1)/\Delta_n \rfloor + 2}^{\lfloor t/\Delta_n \rfloor} |\Delta_{i-1}^n Z||\Delta_i^n Z|. \tag{SB.3}$$

Here, BV_t is the Bipower Variation of Barndorff-Nielsen and Shephard (2004, 2006) which is a jump-robust estimator of volatility and importantly free of tuning parameters. For the construction of $\hat{c}_{n,i\pm}$ we include all increments for which both components are below a threshold set similarly as above but with 4 replaced by 3. There are 10,000 Monte Carlo trials.

In Table S2 we report the results from the Monte Carlo for the test of constant jump beta. As seen from the table, the test in general has good size properties. In all cases, we notice some overrejections which are higher for the larger choices of the block size k_n. These overrejections decrease when the sampling frequency increases from 10 to 5 minutes. The test also has a reasonable power against the considered alternative which increases with the sampling frequency. In Table S3 we report the coverage probability for the refined CI of jump beta that is based on the efficient estimator and is described in Algorithm 3. The coverage probabilities are in general quite close to the nominal levels of the CIs. Not surprisingly, we see again improved performance at the higher sampling frequency. We also note that the coverage probability of the CIs is not very sensitive to the choice of the block size k_n. Overall, we find quite satisfactory finite-sample performance of our inference techniques for the jump betas, even for relatively sparse sampling of $1/\Delta_n = 38$.

Table S2: Monte Carlo Rejection Rates (%) of Tests for Constant Jump Beta

Case	Under H_0			Under H_a		
	Nominal Level			Nominal Level		
	10%	5%	1%	10%	5%	1%
$1/\Delta_n = 38$, $k_n = 19$	13.42	7.72	2.99	69.76	59.39	39.28
$1/\Delta_n = 38$, $k_n = 25$	14.09	8.42	3.20	70.19	60.28	41.12
$1/\Delta_n = 81$, $k_n = 27$	12.90	7.40	2.62	90.37	85.19	71.69
$1/\Delta_n = 81$, $k_n = 35$	13.01	7.29	2.40	90.13	85.22	72.45

Table S3: Monte Carlo Coverage Probability (%) of Confidence Intervals

Case	Nominal Level		
	90%	95%	99%
$1/\Delta_n = 38$, $k_n = 19$	88.43	93.60	98.46
$1/\Delta_n = 38$, $k_n = 25$	88.52	94.01	98.43
$1/\Delta_n = 81$, $k_n = 27$	88.68	94.04	98.50
$1/\Delta_n = 81$, $k_n = 35$	88.93	94.32	98.59

Supplemental Appendix C: Proofs

Throughout this appendix, we use K to denote a generic constant that may change from line to line; we sometimes emphasize the dependence of this constant on some parameter q by writing K_q. We use $0_{k\times q}$ to denote a $k\times q$ matrix of zeros and when $q=1$, we write 0_k for notational simplicity; 0_k is understood to be empty when $k=0$. For any sequence of variables $(\xi_{n,p})_{p\geq 1}$, the convergence $(\xi_{n,p})_{p\geq 1}\to(\xi_p)_{p\geq 1}$ is understood as $n\to\infty$ under the product topology. We write w.p.a.1 for "with probability approaching 1."

By a standard localization procedure (see Section 4.4.1 of Jacod and Protter (2012)), we can strengthen Assumption 1 to the following stronger version without loss of generality.

Assumption S6. *We have Assumption 1. Moreover, the processes X_t, b_t and σ_t are bounded.*

SC.1 Proofs of results in the main text

Proof of Proposition 1. (a) Since the jumps of Z have finite activity, we can assume without loss of generality that each interval $((i-1)\Delta_n, i\Delta_n]$ contains at most one jump; otherwise we can restrict our calculation to the w.p.a.1 set of sample paths on which this condition holds. We denote the continuous part of Z by Z^c, that is,

$$Z_t^c = Z_t - \sum_{s\leq t}\Delta Z_s,\quad t\geq 0. \tag{SC.1}$$

Note that $\mathcal{I}_n(\mathcal{D})$ is the union of two disjoint sets $\mathcal{I}_{1n}(\mathcal{D})$ and $\mathcal{I}_{2n}(\mathcal{D})$ that are defined as

$$\mathcal{I}_{1n}(\mathcal{D}) = \mathcal{I}_n(\mathcal{D})\cap\{i(p):p\in\mathcal{P}\},\quad \mathcal{I}_{2n}(\mathcal{D}) = \mathcal{I}_n(\mathcal{D})\backslash\mathcal{I}_{1n}(\mathcal{D}). \tag{SC.2}$$

It suffices to show that, w.p.a.1,

$$\mathcal{I}_{1n}(\mathcal{D}) = \mathcal{I}(\mathcal{D}),\quad \mathcal{I}_{2n}(\mathcal{D}) = \emptyset. \tag{SC.3}$$

First consider $\mathcal{I}_{1n}(\mathcal{D})$. Since $v_n\to 0$, we have $|\Delta_{i(p)}^n Z|>v_n$ for all $p\in\mathcal{P}$, when n is large enough. Therefore,

$$\mathcal{I}_{1n}(\mathcal{D}) = \left\{i(p):p\in\mathcal{P},((i(p)-1)\Delta_n,\Delta_{i(p)}^n Z)\in\mathcal{D}\right\}\text{ w.p.a.1.} \tag{SC.4}$$

Now, observe that

$$\sup_{p\in\mathcal{P}}\left\|\left((i(p)-1)\Delta_n,\Delta_{i(p)}^n Z\right)-\left(\tau_p,\Delta Z_{\tau_p}\right)\right\|\to 0\quad\text{a.s.} \tag{SC.5}$$

Indeed, almost surely,

$$\begin{aligned}\sup_{p\in\mathcal{P}}\left\|\left((i(p)-1)\Delta_n,\Delta_{i(p)}^n Z\right)-\left(\tau_p,\Delta Z_{\tau_p}\right)\right\| &= \sup_{p\in\mathcal{P}}\left\|\left((i(p)-1)\Delta_n-\tau_p,\Delta_{i(p)}^n Z^c\right)\right\| \\ &\leq \Delta_n+\sup_{s,t\leq T,|s-t|\leq\Delta_n}|Z_t^c-Z_s^c|\to 0. \quad \text{(SC.6)}\end{aligned}$$

By Assumption 2, the marks $(\tau_p,\Delta Z_{\tau_p})_{p\in\mathcal{P}_{\mathcal{D}}}$ are contained in the interior of $\mathcal{D}$ a.s. Then, by (SC.5), $((i(p)-1)\Delta_n,\Delta_{i(p)}^n Z)_{p\in\mathcal{P}_{\mathcal{D}}}\subseteq\mathcal{D}$ w.p.a.1. With the same argument but with $\mathcal{D}^c$ (i.e. the complement of $\mathcal{D}$) replacing $\mathcal{D}$, we deduce $((i(p)-1)\Delta_n,\Delta_{i(p)}^n Z)_{p\in\mathcal{P}\setminus\mathcal{P}_{\mathcal{D}}}\subseteq\mathcal{D}^c$ w.p.a.1. Therefore, the set on the right-hand side of (SC.4) coincides with $\mathcal{I}(\mathcal{D})$ w.p.a.1. From here, the first claim of (SC.3) readily follows.

It remains to show that $\mathcal{I}_{2n}(\mathcal{D})$ is empty w.p.a.1. Note that for $i\in\mathcal{I}_{2n}(\mathcal{D})$, $\Delta_i^n Z=\Delta_i^n Z^c$. Hence, for any $q>2/(1-2\varpi)$,

$$\mathbb{P}(\mathcal{I}_{2n}(\mathcal{D})\neq\emptyset)\leq\sum_{i=1}^{\lfloor T/\Delta_n\rfloor}\mathbb{P}(|\Delta_i^n Z^c|>v_n)\leq K_q\Delta_n^{-1}\frac{\Delta_n^{q/2}}{v_n^q}\to 0, \quad \text{(SC.7)}$$

where the second inequality is by Markov's inequality and $\mathbb{E}|\Delta_i^n Z^c|^q\leq K_q\Delta_n^{q/2}$; the convergence is due to (2.12) and our choice of q. The proof of part (a) is now complete.

(b) By part (a), it suffices to show that

$$((i-1)\Delta_n,\Delta_i^n X)_{i\in\mathcal{I}(\mathcal{D})}-\left(\tau_p,\Delta X_{\tau_p}\right)_{p\in\mathcal{P}_{\mathcal{D}}}=o_p(1). \quad \text{(SC.8)}$$

Observe that $((i-1)\Delta_n,\Delta_i^n X)_{i\in\mathcal{I}(\mathcal{D})}$ is simply $((i(p)-1)\Delta_n,\Delta_{i(p)}^n X)_{p\in\mathcal{P}_{\mathcal{D}}}$. We deduce the desired convergence via the same argument as that for (SC.5). *Q.E.D.*

Proof of Theorem 1. (a) Let

$$\bar{\beta}(\mathcal{D})\equiv\frac{Q_{ZY}(\mathcal{D})}{Q_{ZZ}(\mathcal{D})}. \quad \text{(SC.9)}$$

For each $p\geq 1$, we set

$$R_{n,p}=\Delta_n^{-1/2}(\Delta_{i(p)}^n X-\Delta X_{\tau_p}) \quad \text{and} \quad \varsigma_{n,p}=(-\bar{\beta}(\mathcal{D}),1)R_{n,p}. \quad \text{(SC.10)}$$

With these notations, we have in restriction to $\Omega_0(\mathcal{D})$,

$$\Delta_{i(p)}^n Y=\beta_0\Delta_{i(p)}^n Z+\Delta_n^{1/2}\varsigma_{n,p}. \quad \text{(SC.11)}$$

By Proposition 4.4.10 in Jacod and Protter (2012), $(R_{n,p})_{p\geq 1}\xrightarrow{\mathcal{L}\text{-}s}(R_p)_{p\geq 1}$, where R_p is defined in (3.2). Consequently (recall the notation (3.12)),

$$(\varsigma_{n,p})_{p\geq 1}\xrightarrow{\mathcal{L}\text{-}s}(\varsigma_p)_{p\geq 1}. \quad \text{(SC.12)}$$

By Proposition 1(a), w.p.a.1.,

$$\det\left[Q_n\left(\mathcal{D}\right)\right]=\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nZ^2\right)\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nY^2\right)-\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nZ\Delta_{i(p)}^nY\right)^2. \tag{SC.13}$$

Plug (SC.11) into (SC.13). After some algebra, we deduce

$$\Delta_n^{-1}\det[Q_n\left(\mathcal{D}\right)]=\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nZ^2\right)\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\varsigma_{n,p}^2\right)-\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nZ\,\varsigma_{n,p}\right)^2. \tag{SC.14}$$

Note that for each $p\geq 1$, $\Delta_{i(p)}^nZ\to\Delta Z_{\tau_p}$. Combining this convergence with (SC.12), we use the property of stable convergence to derive the joint convergence

$$\left(\varsigma_{n,p},\Delta_{i(p)}^nZ\right)_{p\geq 1}\xrightarrow{\mathcal{L}\text{-}s}\left(\varsigma_p,\Delta Z_{\tau_p}\right)_{p\geq 1}. \tag{SC.15}$$

Since the set $\mathcal{P}_\mathcal{D}$ is a.s. finite, the assertion of part (a) follows from (SC.14), (SC.15) and the continuous mapping theorem.

(b) By a standard localization argument (see Section 4.4.1 of Jacod and Protter (2012)), we assume that Assumption S6 holds without loss of generality. Since $\mathcal{P}_\mathcal{D}$ is a.s. finite, we can also assume that $|\mathcal{P}_\mathcal{D}|\leq M$ for some constant $M>0$ for the purpose of proving convergence in probability; otherwise, we can fix some large M to make $\mathbb{P}\left(|\mathcal{P}_\mathcal{D}|>M\right)$ arbitrarily small and restrict the calculation below on the set $\{|\mathcal{P}_\mathcal{D}|\leq M\}$.

By Theorem 9.3.2 in Jacod and Protter (2012), we have,

$$\hat{c}_{n,i(p)-}\xrightarrow{\mathbb{P}}c_{\tau_p-},\quad \hat{c}_{n,i(p)+}\xrightarrow{\mathbb{P}}c_{\tau_p},\quad \text{all}\quad 1\leq p\leq M. \tag{SC.16}$$

By Proposition 1(b),

$$Q_n\left(\mathcal{D}\right)\xrightarrow{\mathbb{P}}Q\left(\mathcal{D}\right), \tag{SC.17}$$

which further implies (with $\tilde{\beta}_n\equiv Q_{ZY,n}\left(\mathcal{D}\right)/Q_{ZZ,n}\left(\mathcal{D}\right)$)

$$\tilde{\beta}_n\xrightarrow{\mathbb{P}}\bar{\beta}\left(\mathcal{D}\right). \tag{SC.18}$$

Furthermore, by essentially the same argument as in the proof of Proposition 1(a), we deduce

$$\mathcal{I}_n'\left(\mathcal{D}\right)=\mathcal{I}\left(\mathcal{D}\right)\quad\text{w.p.a.1.} \tag{SC.19}$$

Therefore,

$$\tilde{\zeta}_n\left(\mathcal{D}\right)=\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nZ^2\right)\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\tilde{\varsigma}_{n,i(p)}^2\right)-\left(\sum_{p\in\mathcal{P}_\mathcal{D}}\Delta_{i(p)}^nZ\,\varsigma_{n,i(p)}\right)^2\quad\text{w.p.a.1.} \tag{SC.20}$$

Fix any subsequence $\mathbb{N}_1 \subseteq \mathbb{N}$. By (SC.16) and (SC.18), we can extract a further subsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$, such that along $\mathbb{N}_2$,

$$\left(\left(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}\right)_{1\leq p\leq M}, \tilde{\beta}_n\right) \to \left(\left(c_{\tau_p-}, c_{\tau_p}\right)_{1\leq p\leq M}, \bar{\beta}\left(\mathcal{D}\right)\right) \tag{SC.21}$$

on some set $\tilde{\Omega}$ with $\mathbb{P}(\tilde{\Omega}) = 1$. Then, for each $\omega \in \tilde{\Omega}$ fixed, the transition kernel of $\tilde{\zeta}_n(\mathcal{D})$ given $\mathcal{F}$ converges weakly to the $\mathcal{F}$-conditional law of $\zeta(\mathcal{D})$. Moreover, observe that the $\mathcal{F}$-conditional law of the variables $(\varsigma_p)_{1\leq p\leq M}$ does not have atoms and has full support on $\mathbb{R}^M$. Therefore, the $\mathcal{F}$-conditional distribution function of $\zeta(\mathcal{D})$ is continuous and strictly increasing. By Lemma 21.2 in van der Vaart (1998), we deduce that on each path $\omega \in \tilde{\Omega}$, along the subsequence $\mathbb{N}_2$, $cv_n^\alpha \to cv^\alpha$, where cv^α is the $\mathcal{F}$-conditional $(1-\alpha)$-quantile of $\zeta(\mathcal{D})$. Since the subsequence $\mathbb{N}_1$ is arbitrarily chosen, we further deduce that $cv_n^\alpha \xrightarrow{\mathbb{P}} cv^\alpha$ by the subsequence characterization of convergence in probability. The proof for part (b) is now complete.

(c) By part (a) and part (b), as well as the property of stable convergence, we have

$$(\Delta_n^{-1}\det\left[Q_n(\mathcal{D})\right], cv_n^\alpha, 1_{\Omega_0(\mathcal{D})}) \xrightarrow{\mathcal{L}\text{-}s} (\zeta(\mathcal{D}), cv^\alpha, 1_{\Omega_0(\mathcal{D})}). \tag{SC.22}$$

In particular,

$$\mathbb{P}\left(\{\Delta_n^{-1}\det\left[Q_n(\mathcal{D})\right] > cv_n^\alpha\} \cap \Omega_0(\mathcal{D})\right) \to \mathbb{P}\left(\{\zeta(\mathcal{D}) > cv^\alpha\} \cap \Omega_0(\mathcal{D})\right). \tag{SC.23}$$

Since $\mathbb{P}(\zeta(\mathcal{D}) > cv^\alpha|\mathcal{F}) = \alpha$ and $\Omega_0(\mathcal{D}) \in \mathcal{F}$, the right-hand side of (SC.23) equals to $\alpha\mathbb{P}(\Omega_0(\mathcal{D}))$. The first assertion of part (c) then follows from (SC.23). To show the second assertion of part (c), we first observe that (SC.17) implies $\det\left[Q_n(\mathcal{D})\right] \xrightarrow{\mathbb{P}} \det\left[Q(\mathcal{D})\right]$. In restriction to $\Omega_a(\mathcal{D})$, $\det\left[Q(\mathcal{D})\right] > 0$ and, hence, $\Delta_n^{-1}\det\left[Q_n(\mathcal{D})\right]$ diverges to $+\infty$ in probability. Part (b) implies that cv_n^α is tight in restriction to $\Omega_a(\mathcal{D})$. Consequently, $\mathbb{P}\left(\Delta_n^{-1}\det\left[Q_n(\mathcal{D})\right] > cv_n^\alpha|\Omega_a(\mathcal{D})\right) \to 1$ as asserted. *Q.E.D.*

Proof of Theorem 2. (a) Observe that

$$Q_{ZY,n}(\mathcal{D}, w) - \beta_0 Q_{ZZ,n}(\mathcal{D}, w) = \sum_{i\in\mathcal{I}_n'(\mathcal{D})} w\left(\hat{c}_{i-}^n, \hat{c}_{i+}^n, \tilde{\beta}_n\right)\Delta_i^n Z\left(\Delta_i^n Y - \beta_0\Delta_i^n Z\right). \tag{SC.24}$$

Recall the notation $\varsigma_{n,p}$ from (SC.10). By (SC.19), we further deduce that, w.p.a.1,

$$\Delta_n^{-1/2}\left(Q_{ZY,n}(\mathcal{D}, w) - \beta_0 Q_{ZZ,n}(\mathcal{D}, w)\right) = \sum_{p\in\mathcal{P}_\mathcal{D}} w\left(\hat{c}_{n,i(p)-}, \hat{c}_{n,i(p)+}, \tilde{\beta}_n\right)\Delta_{i(p)}^n Z\,\varsigma_{n,p}. \tag{SC.25}$$

By (SC.16), (SC.18) and Assumption 3,

$$w\left(\hat{c}_{i(p)-}^n, \hat{c}_{i(p)+}^n, \tilde{\beta}_n\right) \xrightarrow{\mathbb{P}} w\left(c_{\tau_p-}, c_{\tau_p}, \beta_0\right), \quad p \geq 1. \tag{SC.26}$$

Since $\mathcal{P}_{\mathcal{D}}$ is a.s. finite, we use properties of stable convergence to deduce from (SC.12) and (SC.26) that

$$\Delta_n^{-1/2}\left(Q_{ZY,n}\left(\mathcal{D},w\right)-\beta_0 Q_{ZZ,n}\left(\mathcal{D},w\right)\right)\xrightarrow{\mathcal{L}\text{-}s}\sum_{p\in\mathcal{P}_{\mathcal{D}}}w\left(c_{\tau_p-},c_{\tau_p},\beta_0\right)\Delta Z_{\tau_p}\varsigma_p. \tag{SC.27}$$

Note that

$$\Delta_n^{-1/2}(\hat{\beta}_n\left(\mathcal{D},w\right)-\beta_0)=\frac{\Delta_n^{-1/2}\left(Q_{ZY,n}\left(\mathcal{D},w\right)-\beta_0 Q_{ZZ,n}\left(\mathcal{D},w\right)\right)}{Q_{ZZ,n}\left(\mathcal{D},w\right)}. \tag{SC.28}$$

By (SC.19),

$$Q_n\left(\mathcal{D},w\right)=\sum_{p\in\mathcal{P}_{\mathcal{D}}}w\left(\hat{c}^n_{i(p)-},\hat{c}^n_{i(p)+},\tilde{\beta}_n\right)\Delta^n_{i(p)}X\Delta^n_{i(p)}X^{\top}. \tag{SC.29}$$

By $\Delta^n_{i(p)}X\to\Delta X_{\tau_p}$ and (SC.26), we deduce

$$Q_n\left(\mathcal{D},w\right)\xrightarrow{\mathbb{P}}\sum_{p\in\mathcal{P}_{\mathcal{D}}}w\left(c_{\tau_p-},c_{\tau_p},\beta_0\right)\Delta X_{\tau_p}\Delta X_{\tau_p}^{\top}. \tag{SC.30}$$

The first assertion of part (a), that is, $\Delta_n^{-1/2}(\hat{\beta}_n\left(\mathcal{D},w\right)-\beta_0)\xrightarrow{\mathcal{L}\text{-}s}\zeta_\beta\left(\mathcal{D},w\right)$ readily follows from (SC.27), (SC.28) and (SC.30).

Turning to the second assertion of part (a), we first observe that when c_t does not jump at the same time as Z_t, each ς_p is $\mathcal{F}$-conditionally centered Gaussian; moreover, the variables $(\varsigma_p)_{p\geq 1}$ are $\mathcal{F}$-conditionally independent. Therefore, the limiting variable $\zeta_\beta\left(\mathcal{D}\right)$ is centered Gaussian conditional on $\mathcal{F}$, with conditional variance given by $\Sigma\left(\mathcal{D},w\right)$. This finishes the proof of the second assertion.

(b) For notational simplicity, we denote

$$A_p=\frac{\left(-\beta_0,1\right)\left(c_{\tau_p-}+c_{\tau_p}\right)\left(-\beta_0,1\right)^{\top}}{2\Delta Z^2_{\tau_p}},\quad B_p=w\left(c_{\tau_p-},c_{\tau_p},\beta_0\right)\Delta Z^2_{\tau_p}.$$

Then we can rewrite $\Sigma\left(\mathcal{D},w\right)$ and $\Sigma\left(\mathcal{D},w^*\right)$ as

$$\Sigma\left(\mathcal{D},w\right)=\frac{\sum_{p\in\mathcal{P}_{\mathcal{D}}}B_p^2 A_p}{\left(\sum_{p\in\mathcal{P}_{\mathcal{D}}}B_p\right)^2},\quad \Sigma\left(\mathcal{D},w^*\right)=\left(\sum_{p\in\mathcal{P}_{\mathcal{D}}}A_p^{-1}\right)^{-1}.$$

The assertion of part (b) is then proved by observing

$$\sqrt{\frac{\Sigma\left(\mathcal{D},w\right)}{\Sigma\left(\mathcal{D},w^*\right)}}=\frac{\sqrt{\sum_{p\in\mathcal{P}_{\mathcal{D}}}B_p^2A_p}\sqrt{\sum_{p\in\mathcal{P}_{\mathcal{D}}}A_p^{-1}}}{\sum_{p\in\mathcal{P}_{\mathcal{D}}}B_p}\geq 1,$$

where the inequality is by the Cauchy-Schwarz inequality.

(c) By (SC.19) and (SC.26), as well as $\Delta^n_{i(p)}Z\to\Delta Z_{\tau_p}$, we deduce that the $\mathcal{F}$-conditional law of $\tilde{\zeta}_{n,\beta}(\mathcal{D},w)$ converges in probability to that of $\zeta_\beta(\mathcal{D},w)$ under any metric for weak convergence.

From here, by using an argument similar to that in the proof of Theorem 1(b), we further deduce that

$$cv_{n,\beta}^{\alpha/2} \xrightarrow{\mathbb{P}} cv_{\beta}^{\alpha/2}, \tag{SC.31}$$

where $cv_{\beta}^{\alpha/2}$ denotes the $(1-\alpha/2)$-quantile of the $\mathcal{F}$-conditional law of $\zeta_\beta(\mathcal{D}, w)$. It is easy to see that the $\mathcal{F}$-conditional law of $\zeta_\beta(\mathcal{D}, w)$ is symmetric. The assertion of part (c) then follows from part (a) and (SC.31). *Q.E.D.*

Proof of Theorem 3. (a) Fix $S \in \mathbf{S}$ and let $m = \dim(S) - 1$. We consider a sequence of subsets Ω_n defined by

$$\Omega_n = \left\{ \begin{array}{c} \text{For every } 1 \leq i \leq \lfloor T/\Delta_n \rfloor \text{, if } ((i-1)\Delta_n, i\Delta_n] \text{ contains} \\ \text{some jump of } Z\text{, then this interval is contained in } (S_{j-1}, S_j] \\ \text{for some } 1 \leq j \leq m \text{ and it contains exactly one jump of } Z. \end{array} \right\}$$

Under Assumption 1, the process Z has finitely active jumps without any fixed time of discontinuity. Hence, $\mathbb{P}(\Omega_n) \to 1$, so we can restrict our calculation below on Ω_n without loss of generality.

Below, we write $h = (h_0, \ldots h_m)^\top$ and denote the log likelihood ratio by

$$L_n(h) = \log \frac{dP^n_{\theta_0 + \Delta_n^{1/2} h}}{dP^n_{\theta_0}}.$$

For each $i \geq 1$, we set $h(n,i) = h_j$, where j is the unique integer in $\{1, \ldots, m\}$ such that $i\Delta_n \in (S_{j-1}, S_j]$. On the set Ω_n, with $\theta = \theta_0 + \Delta_n^{1/2} h$, we have

$$\Delta_i^n X = \int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds + \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s + \left(\begin{array}{c} (1 + \Delta_n^{1/2} h(n,i)) \Delta_i^n J_Z \\ \left(\beta_0 + \Delta_n^{1/2} h_0\right)(1 + \Delta_n^{1/2} h(n,i)) \Delta_i^n J_Z + \Delta_i^n \epsilon \end{array} \right).$$

To simplify notations, we denote for each $i \geq 1$,

$$\begin{aligned} x_{n,i} &\equiv \Delta_n^{-1/2} \int_{(i-1)\Delta_n}^{i\Delta_n} \sigma_s dW_s, \\ \bar{b}_{n,i} &\equiv \int_{(i-1)\Delta_n}^{i\Delta_n} b_s ds, \quad \bar{c}_{n,i} \equiv \Delta_n^{-1} \int_{(i-1)\Delta_n}^{i\Delta_n} c_s ds, \\ J_{n,i} &\equiv \left(\begin{array}{c} \Delta_i^n J_Z \\ \beta_0 \Delta_i^n J_Z + \Delta_i^n \epsilon \end{array} \right), \quad d_{n,i} \equiv \left(\begin{array}{c} h(n,i) \\ h_0 + \beta_0 h(n,i) + \Delta_n^{1/2} h_0 h(n,i) \end{array} \right). \end{aligned}$$

Note that under Assumption 4, $(x_{n,i})_{i \geq 1}$ are independent conditional on $(b_t, \sigma_t, J_{Z,t}, \epsilon_t)_{t \geq 0}$ and each $x_{n,i}$ is distributed as $\mathcal{N}(0, \bar{c}_{n,i})$. With these notations, we can write the log likelihood ratio explicitly as

$$L_n(h) = \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \Delta_i^n J_Z d_{n,i}^\top \bar{c}_{n,i}^{-1} x_{n,i} - \frac{1}{2} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \Delta_i^n J_Z^2 d_{n,i}^\top \bar{c}_{n,i}^{-1} d_{n,i}. \tag{SC.32}$$

Note that on Ω_n, $\Delta_i^n J_Z \neq 0$ only if $((i-1)\Delta_n, i\Delta_n]$ contains one (and only one) jump of Z. Therefore,

$$L_n(h) = \sum_{p \in \mathcal{P}} \Delta Z_{\tau_p} d_{n,i(p)}^\top \bar{c}_{n,i(p)}^{-1} x_{n,i(p)} - \frac{1}{2} \sum_{p \in \mathcal{P}} \Delta Z_{\tau_p}^2 d_{n,i(p)}^\top \bar{c}_{n,i(p)}^{-1} d_{n,i(p)}. \tag{SC.33}$$

By Proposition 4.4.10 in Jacod and Protter (2012), $\left(x_{n,i(p)}\right)_{p\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (R_p)_{p\geq 1}$. Under Assumption 5, the variables $(R_p)_{p\geq 1}$ are $\mathcal{F}$-conditionally independent, where the $\mathcal{F}$-conditional law of R_p is $\mathcal{N}\left(0, c_{\tau_p}\right)$; moreover, $\bar{c}_{n,i(p)} \to c_{\tau_p}$ a.s. for each $p \geq 1$. Further note that for each $p \geq 1$,

$$d_{n,i(p)} \longrightarrow D_p h. \tag{SC.34}$$

where the matrix D_p is defined as

$$D_p \equiv \begin{pmatrix} 0 & 0_{j-1}^\intercal & 1 & 0_{m-j}^\intercal \\ 1 & 0_{j-1}^\intercal & \beta_0 & 0_{m-j}^\intercal \end{pmatrix} \quad \text{for } j \text{ such that } \tau_p \in (S_{j-1}, S_j]. \tag{SC.35}$$

Since $\mathcal{P}$ is a.s. finite, we deduce (4.9) from (SC.33) and (SC.34), that is,

$$L_n(h) = h^\top \Gamma_n^{1/2} \psi_n - \frac{1}{2} h^\top \Gamma_n h + o_p(1), \tag{SC.36}$$

where

$$\Gamma_n \equiv \sum_{p \in \mathcal{P}} \Delta Z_{\tau_p}^2 D_p^\top \bar{c}_{n,i(p)}^{-1} D_p, \quad \psi_n = \Gamma_n^{-1/2} \sum_{p \in \mathcal{P}} \Delta Z_{\tau_p} D_p^\top \bar{c}_{n,i(p)}^{-1} x_{n,i(p)}. \tag{SC.37}$$

In addition, (4.10) follows with

$$\Gamma \equiv \sum_{p \in \mathcal{P}} \Delta Z_{\tau_p}^2 D_p^\top c_{\tau_p}^{-1} D_p, \quad \psi \equiv \Gamma^{-1/2} \sum_{p \in \mathcal{P}} \Delta Z_{\tau_p} D_p^\top c_{\tau_p}^{-1} R_p. \tag{SC.38}$$

It is easy to verify that Γ defined in (SC.38) equals to $\Gamma(S)$ defined by (4.17). To see, we make the following explicit calculation using (SC.35),

$$D_p^\top c_{\tau_p}^{-1} D_p = \begin{pmatrix} \frac{1}{v_{\tau_p}^c} & 0_{j-1}^\top & \frac{\beta_0 - \beta_{\tau_p}^c}{v_{\tau_p}^c} & 0_{m-j}^\top \\ 0_{j-1} & 0_{(j-1)\times(j-1)} & 0_{j-1} & 0_{(j-1)\times(m-j)} \\ \frac{\beta_0 - \beta_{\tau_p}^c}{v_{\tau_p}^c} & 0_{j-1}^\top & \frac{\left(\beta_0 - \beta_{\tau_p}^c\right)^2}{v_{\tau_p}^c} + \frac{1}{c_{ZZ,\tau_p}} & 0_{m-j}^\top \\ 0_{m-j} & 0_{(m-j)\times(j-1)} & 0_{m-j} & 0_{(m-j)\times(m-j)} \end{pmatrix}. \tag{SC.39}$$

Finally, we note that conditional on $\mathcal{F}$, ψ has a standard normal distribution and, hence, is independent of $\mathcal{F}$. The proof for the LAMN property is now complete.

From the proof of Theorem 3 of Jeganathan (1982), we see that the convolution theorem can be applied in restriction to the set $\Omega(S) \equiv \{\Gamma(S) \text{ is nonsingular}\}$. The information bound for estimating β, that is, the first diagonal element of $\Gamma(S)^{-1}$, can then be easily computed by using the inversion formula for partitioned matrices.

(b) Since the jumps of Z have finite activity, on each sample path $\omega \in \Omega$ there exists some $S^*(\omega) \in \mathbf{S}$ that shatters its jumps. That is, each interval $(S^*_{j-1}(\omega), S^*_j(\omega)]$ contains at exactly one jump time of Z. We can then evaluate $\bar{\Sigma}_\beta(\cdot)$ at S^* on each sample path and obtain

$$\bar{\Sigma}_\beta(S^*) = \left(\sum_{s \le T}\left(\frac{\Delta Z_s^2}{v_s^c} - \frac{\gamma_{1s}^2}{\gamma_{2s}}\right)\right)^{-1}. \tag{SC.40}$$

Plugging the definitions of γ_{1s} and γ_{2s} (see (4.16)) into (SC.40), we can verify that

$$\bar{\Sigma}_\beta(S^*) = \left(\sum_{s \le T}\frac{\Delta Z_s^2}{c_{YY,s} - 2\beta_0 c_{ZY,s} + \beta_0^2 c_{ZZ,s}}\right)^{-1}. \tag{SC.41}$$

Recall that we fix $\mathcal{D} = [0,T] \times \mathbb{R}_*$ and $\Sigma^* \equiv \Sigma(\mathcal{D}, w^*)$, with the latter given by (4.8). Under Assumption 5, we see $\bar{\Sigma}_\beta(S^*) = \Sigma^*$.

It remains to verify that $\bar{\Sigma}_\beta(S^*) \ge \bar{\Sigma}_\beta(S)$ for all $S \in \mathbf{S}$. By the Cauchy–Schwarz inequality,

$$\frac{\left(\sum_{S_{j-1} < s \le S_j}\gamma_{1s}\right)^2}{\sum_{S_{j-1} < s \le S_j}\gamma_{2s}} \le \sum_{S_{j-1} < s \le S_j}\frac{\gamma_{1s}^2}{\gamma_{2s}}. \tag{SC.42}$$

From (4.19), (SC.40) and (SC.42), $\bar{\Sigma}_\beta(S^*) \ge \bar{\Sigma}_\beta(S)$ readily follows. *Q.E.D.*

SC.2 Proofs of results in Appendices SA.1, SA.2 and SA.3

Proof of Theorem S1. The proof is essentially the same as that of Theorem 1 except that we derive the convergence $R_{n,p} \xrightarrow{\mathcal{L}\text{-}s} R_p$ for $p \in \mathcal{A}_\mathcal{D}$ using Assumption S1. To do so, we denote $\kappa_{n,p} = \tau_p/\Delta_n - \lfloor \tau_p/\Delta_n \rfloor$ and observe

$$\begin{aligned} R_{n,p} &= \Delta_n^{-1/2}\int_{(i(p)-1)\Delta_n}^{i(p)\Delta_n}\sigma_s dW_s + o_p(1) \\ &= \sigma_{(i(p)-1)\Delta_n}\sqrt{\kappa_{n,p}}\frac{W_{\tau_p} - W_{(i(p)-1)\Delta_n}}{\sqrt{\tau_p - (i(p)-1)\Delta_n}} + \sigma_{\tau_p}\sqrt{1-\kappa_{n,p}}\frac{W_{i(p)\Delta_n} - W_{\tau_p}}{\sqrt{i(p)\Delta_n - \tau_p}} + o_p(1), \end{aligned}$$

where the $o_p(1)$ term in the first equality is due to the drift and, in the second equality, we use the standard local Gaussian approximation (using Itô's isometry and the càdlàg property of σ) to the continuous martingale component before and after the jump. From here, the claimed convergence of $R_{n,p}$ follows from the càdlàg property of σ and Assumption S1. *Q.E.D.*

Proof of Theorem S2. We complement the notations in (SA.8) with

$$\tilde{v}_{n,i\pm} \equiv (-\beta_0, 1)\, \hat{c}_{n,i\pm} (-\beta_0, 1)^\top . \tag{SC.43}$$

Observe that v_t (recall (SA.2)) is the spot covariance matrix of the process $Y - \beta_0 Z$. Then, by applying Theorem 13.3.3(c) of Jacod and Protter (2012) to the process $Y - \beta_0 Z$, we deduce that

$$k_n^{1/2} \left(\tilde{v}_{n,i(p)-} - v_{\tau_p -}, \tilde{v}_{n,i(p)+} - v_{\tau_p}\right)_{p\geq 1} \xrightarrow{\mathcal{L}\text{-}s} \left(\sqrt{2} v_{\tau_p -} \xi'_{p-}, \sqrt{2} v_{\tau_p} \xi'_{p+}\right)_{p\geq 1}. \tag{SC.44}$$

Recall the notations in (SA.3) and (SA.8). For each $p \geq 1$, we can decompose

$$\hat{\phi}_{n,i(p)} = \phi_p + k_n^{-1/2} F_{n,p} + G_{n,p}, \tag{SC.45}$$

where

$$\begin{cases} F_{n,p} \equiv k_n^{1/2} \left(\left(\tilde{v}_{n,i(p)-} + \tilde{v}_{n,i(p)+}\right)/2 - \phi_p\right), \\ G_{n,p} \equiv \hat{\phi}_{n,p} - \left(\tilde{v}_{n,i(p)-} + \tilde{v}_{n,i(p)+}\right)/2. \end{cases} \tag{SC.46}$$

From (SC.44), it follows that

$$\left(F_{n,p}\right)_{p\geq 1} \xrightarrow{\mathcal{L}\text{-}s} \left(F_p\right)_{p\geq 1}. \tag{SC.47}$$

We also see from Theorem 2(a) that $\tilde{\beta}_n - \beta_0 = O_p(\Delta_n^{1/2}) = o_p(k_n^{-1/2})$, so we further deduce

$$G_{n,p} = o_p(k_n^{-1/2}). \tag{SC.48}$$

We now turn to the estimator $\hat{\beta}_n(\mathcal{D}, w^*)$. By (SC.19), we have

$$\Delta_n^{-1/2}\left(\hat{\beta}_n(\mathcal{D}, w^*) - \beta_0\right) = \frac{\Delta_n^{-1/2} \sum_{p\in\mathcal{P}} \Delta_{i(p)}^n Z \left(\Delta_{i(p)}^n Y - \beta_0 \Delta_{i(p)}^n Z\right)/\hat{\phi}_{n,i(p)}}{\sum_{p\in\mathcal{P}} \Delta_{i(p)}^n Z^2 / \hat{\phi}_{n,i(p)}} \quad \text{w.p.a.1.} \tag{SC.49}$$

Recall the notations $R_{n,p}$ and $\varsigma_{n,p}$ from (SC.10) and write $R_{n,p} = (R_{Z,n,p}, R_{Y,n,p})^\top$. We can rewrite (SC.49) as

$$\Delta_n^{-1/2}\left(\hat{\beta}_n(\mathcal{D}, w^*) - \beta_0\right) = \frac{\sum_{p\in\mathcal{P}} \left(\Delta Z_{\tau_p} + \Delta_n^{1/2} R_{Z,n,p}\right) \varsigma_{n,p}/\hat{\phi}_{n,i(p)}}{\sum_{p\in\mathcal{P}} \left(\Delta Z_{\tau_p} + \Delta_n^{1/2} R_{Z,n,p}\right)^2 / \hat{\phi}_{n,i(p)}} \quad \text{w.p.a.1.} \tag{SC.50}$$

Next, we derive expansions for the numerator and the denominator of the right-hand side of (SC.50) separately. Observe that the numerator satisfies

$$
\begin{aligned}
&\sum_{p\in\mathcal{P}} \frac{\left(\Delta Z_{\tau_p} + \Delta_n^{1/2} R_{Z,n,p}\right) \varsigma_{n,p}}{\hat{\phi}_{n,i(p)}} - \sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}\, \varsigma_{n,p}}{\phi_p} \\
&\quad = \sum_{p\in\mathcal{P}} \frac{\left(\Delta Z_{\tau_p} + \Delta_n^{1/2} R_{Z,n,p}\right) \varsigma_{n,p}\, \phi_p - \Delta Z_{\tau_p}\, \varsigma_{n,p} \left(\phi_p + k_n^{-1/2} F_{n,p} + G_{n,p}\right)}{\hat{\phi}_{n,i(p)} \phi_p} \\
&\quad = -k_n^{-1/2} \sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}\, \varsigma_{n,p}\, F_{n,p}}{\hat{\phi}_{n,i(p)} \phi_p} + \sum_{p\in\mathcal{P}} \frac{\Delta_n^{1/2} R_{Z,n,p} \varsigma_{n,p}\, \phi_p - \Delta Z_{\tau_p}\, \varsigma_{n,p} G_{n,p}}{\hat{\phi}_{n,i(p)} \phi_p} \\
&\quad = -k_n^{-1/2} \sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}\, \varsigma_{n,p}\, F_{n,p}}{\phi_p^2} + o_p(k_n^{-1/2}),
\end{aligned}
\tag{SC.51}
$$

where the first equality is obtained by using (SC.45); the second equality is obvious; the third equality follows from $R_{Z,n,p} = O_p(1)$, $\varsigma_{n,p} = O_p(1)$, $\hat{\phi}_{n,i(p)} - \phi_p = o_p(1)$ and (SC.48). Similarly, the denominator of the right-hand side of (SC.50) satisfies

$$
\begin{aligned}
&\sum_{p\in\mathcal{P}} \frac{\left(\Delta Z_{\tau_p} + \Delta_n^{1/2} R_{Z,n,p}\right)^2}{\hat{\phi}_{n,i(p)}} - \sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}^2}{\phi_p} \\
&\quad = \sum_{p\in\mathcal{P}} \frac{\left(\Delta Z_{\tau_p} + \Delta_n^{1/2} R_{Z,n,p}\right)^2 \phi_p - \Delta Z_{\tau_p}^2 \left(\phi_p + k_n^{-1/2} F_{n,p} + G_{n,p}\right)}{\hat{\phi}_{n,i(p)} \phi_p} \\
&\quad = -k_n^{-1/2} \sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}^2 F_{n,p}}{\hat{\phi}_{n,i(p)} \phi_p} \\
&\qquad + \sum_{p\in\mathcal{P}} \frac{\left(2\Delta_n^{1/2} \Delta Z_{\tau_p} R_{Z,n,p} + \Delta_n R_{Z,n,p}^2\right) \phi_p - \Delta Z_{\tau_p}^2 G_{n,p}}{\hat{\phi}_{n,i(p)} \phi_p} \\
&\quad = -k_n^{-1/2} \sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p}^2 F_{n,p}}{\phi_p^2} + o_p(k_n^{-1/2}).
\end{aligned}
\tag{SC.52}
$$

Finally, we plug the expansions (SC.51) and (SC.52) into (SC.50) and deduce, w.p.a.1,

$$
\begin{aligned}
&\Delta_n^{-1/2} \left(\hat{\beta}_n\left(\mathcal{D}, w^*\right) - \beta_0\right) \\
&\quad = \frac{\sum_{p\in\mathcal{P}} \Delta Z_{\tau_p}\, \varsigma_{n,p}/\phi_p - k_n^{-1/2} \sum_{p\in\mathcal{P}} \Delta Z_{\tau_p}\, \varsigma_{n,p}\, F_{n,p}/\phi_p^2 + o_p(k_n^{-1/2})}{\sum_{p\in\mathcal{P}} \Delta Z_{\tau_p}^2/\phi_p - k_n^{-1/2} \sum_{p\in\mathcal{P}} \Delta Z_{\tau_p}^2 F_{n,p}/\phi_p^2 + o_p(k_n^{-1/2})} \\
&\quad = \zeta_{n,\beta}^*\left(\mathcal{D}\right) + k_n^{-1/2} H_{n,\beta}^*\left(\mathcal{D}\right) + o_p(k_n^{-1/2}),
\end{aligned}
\tag{SC.53}
$$

where

$$\begin{cases} \zeta^*_{n,\beta}(\mathcal{D}) \equiv \left(\sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p} \varsigma_{n,p}}{\phi_p}\right) / \left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p}}{\phi_p}\right), \\ H^*_{n,\beta}(\mathcal{D}) \equiv \frac{\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p} F_{n,p}}{\phi^2_p}\right)\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p} \varsigma_{n,p}}{\phi_p}\right) - \left(\sum_{p\in\mathcal{P}} \frac{\Delta Z_{\tau_p} \varsigma_{n,p} F_{n,p}}{\phi^2_p}\right)\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p}}{\phi_p}\right)}{\left(\sum_{p\in\mathcal{P}} \frac{\Delta Z^2_{\tau_p}}{\phi_p}\right)^2}. \end{cases} \tag{SC.54}$$

We now observe that the estimators $\hat{c}_{n,i-}$ and $\hat{c}_{n,i+}$ do not involve the increment $\Delta^n_i X$. From here, it is easy to see that the convergence in (SC.12) and (SC.47) hold jointly with $\mathcal{F}$-conditionally independent limits, that is,

$$(\varsigma_{n,p}, F_{n,p})_{p\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\varsigma_p, F_p)_{p\geq 1}. \tag{SC.55}$$

By properties of stable convergence, we deduce

$$\left(\zeta^*_{n,\beta}(\mathcal{D}), H^*_{n,\beta}(\mathcal{D})\right) \xrightarrow{\mathcal{L}\text{-}s} \left(\zeta^*_{\beta}(\mathcal{D}), H^*_{\beta}(\mathcal{D})\right). \tag{SC.56}$$

This finishes the proof. *Q.E.D.*

Proof of Theorem S3. We consider a sequence Ω_n of subsets defined by

$$\Omega_n = \left\{ \begin{array}{c} \text{For every } 1 \leq i \leq \lfloor T/\Delta_n \rfloor, \ ((i-1)\Delta_n, i\Delta_n] \\ \text{contains at most one jump of } Z. \end{array} \right\}.$$

Under the maintained assumptions, the process Z has finitely active jumps. Hence, $\mathbb{P}(\Omega_n) \to 1$, so we can restrict our calculation below on Ω_n without loss of generality.

We denote the log likelihood ratio by

$$L_n(h) \equiv \log \frac{dP^n_{\beta_0+\Delta_n^{1/2}h}}{dP^n_{\beta_0}}, \quad h \in \mathbb{R}.$$

Let $\mathcal{G}$ denote the σ-field generated by the processes $(b, c_{ZZ}, v^c, J_Z, \epsilon)$. Given the maintained assumptions, we see that, under the law P^n_β, the observed returns $(\Delta^n_i X)_{i\geq 0}$ are independently normally distributed conditional on $\mathcal{G}$. Using this fact, we can obtain an explicit expression for $L_n(h)$. For notational simplicity, we denote

$$\begin{aligned} z_{n,i} &\equiv \Delta_n^{-1/2} \int_{(i-1)\Delta_n}^{i\Delta_n} \sqrt{c_{ZZ,s}} dW_{Z,s}, \quad y_{n,i} \equiv \Delta_n^{-1/2} \int_{(i-1)\Delta_n}^{i\Delta_n} \sqrt{v^c_s} dW_{Y,s}, \\ \bar{c}_{n,i} &\equiv \Delta_n^{-1} \int_{(i-1)\Delta_n}^{i\Delta_n} c_{ZZ,s} ds, \quad \bar{v}_{n,i} \equiv \Delta_n^{-1} \int_{(i-1)\Delta_n}^{i\Delta_n} v^c_s ds. \end{aligned}$$

Some straightforward algebra yields

$$L_n(h) = h\tilde{\psi}_n - \frac{h^2}{2}\Gamma_n, \tag{SC.57}$$

where

$$\tilde{\psi}_n \equiv \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \frac{y_{n,i}\left(\Delta_i^n J_Z + \Delta_n^{1/2} z_{n,i}\right)}{\bar{v}_{n,i}}, \quad \Gamma_n \equiv \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \frac{\left(\Delta_i^n J_Z + \Delta_n^{1/2} z_{n,i}\right)^2}{\bar{v}_{n,i}}.$$

It remains to analyze the asymptotic properties of $\tilde{\psi}_n$ and Γ_n. We decompose $\tilde{\psi}_n = \tilde{\psi}'_n + \tilde{\psi}''_n$, where $\tilde{\psi}'_n$ and $\tilde{\psi}''_n$ are sums over the subset $\{i : \Delta_i^n J_Z \neq 0\}$ and its complement, respectively. Similarly, we decompose $\Gamma_n = \Gamma'_n + \Gamma''_n$.

We now proceed to deriving the joint convergence in law of $(\tilde{\psi}'_n, \tilde{\psi}''_n)$ under the $\mathcal{G}$-conditional probability. We note that $\Delta_i^n J_Z$ and $\bar{v}_{n,i}$ are $\mathcal{G}$-measurable, and $(z_{n,i}, y_{n,i})$ are $\mathcal{G}$-conditionally independent with conditional distributions given by

$$z_{n,i}|\mathcal{G} \sim \mathcal{MN}(0, \bar{c}_{n,i}), \quad y_{n,i}|\mathcal{G} \sim \mathcal{MN}(0, \bar{v}_{n,i}).$$

In particular, $\tilde{\psi}'_n$ and $\tilde{\psi}''_n$ are $\mathcal{G}$-conditionally independent, so it is enough to derive the marginal convergence of each sequence. Since the jumps are finitely active, it is easy to see that

$$\tilde{\psi}'_n = \sum_{i:\Delta_i^n J_Z \neq 0} \frac{y_{n,i}\Delta_i^n J_Z}{\bar{v}_{n,i}} + o_p(1), \quad \tilde{\psi}''_n = \Delta_n^{1/2} \sum_{i=1}^{\lfloor T/\Delta_n \rfloor} \frac{y_{n,i} z_{n,i}}{\bar{v}_{n,i}} + o_p(1).$$

By applying the Lindeberg–Lévy central limit theorem under the $\mathcal{G}$-conditional probability, we deduce the following $\mathcal{G}$-conditional convergence in law:

$$\tilde{\psi}'_n \xrightarrow{\mathcal{L}} \mathcal{MN}\left(0, \sum_{\tau \in \mathcal{T}} \frac{\Delta Z_\tau^2}{v_\tau^c}\right), \quad \tilde{\psi}''_n \xrightarrow{\mathcal{L}} \mathcal{MN}\left(0, \int_0^T \frac{c_{ZZ,s}}{v_s^c} ds\right).$$

From here, we deduce the following convergence under the $\mathcal{G}$-conditional probability,

$$\tilde{\psi}_n \xrightarrow{\mathcal{L}} \tilde{\psi} \sim \mathcal{MN}\left(0, \int_0^T \frac{d\,[Z,Z]_s}{v_s^c}\right). \tag{SC.58}$$

Similarly, we can derive the convergence in probability for Γ_n:

$$\Gamma_n \xrightarrow{\mathbb{P}} \Gamma \equiv \int_0^T \frac{d\,[Z,Z]_s}{v_s^c}. \tag{SC.59}$$

Since Γ_n is $\mathcal{G}$-measurable, (SC.58) and (SC.59) imply that $(\tilde{\psi}_n, \Gamma_n)$ converges in law to $(\tilde{\psi}, \Gamma)$. From here, the assertion of the theorem readily follows (recall (SC.57)). *Q.E.D.*

Proof of Theorem S4. We first define some functions. For any 2×2 symmetric matrix $c = [c_{jk}]_{1 \leq j,k \leq 2}$, we define

$$g_1(c) \equiv \frac{c_{12}}{c_{22} - c_{12}^2/c_{11}}, \quad g_2(c) \equiv \frac{c_{11}}{c_{22} - c_{12}^2/c_{11}}, \quad g_3(c) \equiv g_1(c) - \beta_0 g_2(c).$$

With any twice continuously differentiable function g, we associate a function

$$\mathbb{B}g(c) \equiv \frac{1}{2} \sum_{j,k,l,m=1}^{2} \partial_{jk,lm}^2 g(c)(c_{jl}c_{km} + c_{jm}c_{kl}),$$

and a statistic

$$\hat{S}_n(g) \equiv \Delta_n \sum_{i=0}^{\lfloor T/\Delta_n \rfloor - k_n} \left(g(\hat{c}_{n,i}) - \frac{1}{k_n} \mathbb{B}g\left(\hat{c}_{n,i}\right) \right).$$

We also set

$$Q_{n,ZY}^{\star} \equiv \sum_{i \in \mathcal{I}_n'} \frac{\Delta_i^n Z \Delta_i^n Y}{\hat{v}_{n,i}^c}, \quad Q_{n,ZZ}^{\star} \equiv \sum_{i \in \mathcal{I}_n'} \frac{\left(\Delta_i^n Z\right)^2}{\hat{v}_{n,i}^c}.$$

By direct calculation, we see that $\mathbb{B}g_1(c) = 3g_1(c)$ and $\mathbb{B}g_2(c) = 3g_2(c)$. With this notation, we can rewrite $\hat{\beta}_n^{\star}$ as

$$\hat{\beta}_n^{\star} = \frac{\hat{S}_n(g_1) + Q_{n,ZY}^{\star}}{\hat{S}_n(g_2) + Q_{n,ZZ}^{\star}}.$$

Hence,

$$\Delta_n^{-1/2}\left(\hat{\beta}_n^{\star} - \beta_0\right) = \frac{\Delta_n^{-1/2}\hat{S}_n(g_3) + \Delta_n^{-1/2}\left(Q_{n,ZY}^{\star} - \beta_0 Q_{n,ZZ}^{\star}\right)}{\hat{S}_n(g_2) + Q_{n,ZZ}^{\star}}. \tag{SC.60}$$

As a special case of (SC.27), we deduce

$$\Delta_n^{-1/2}\left(Q_{n,ZY}^{\star} - \beta_0 Q_{n,ZZ}^{\star}\right) \xrightarrow{\mathcal{L}\text{-}s} \sum_{p \in \mathcal{P}} \frac{\Delta Z_{\tau_p} \varsigma_p}{v_{\tau_p}^c} \sim \mathcal{MN}\left(0, \sum_{\tau \in \mathcal{T}} \frac{\Delta Z_{\tau}^2}{v_{\tau}^c}\right). \tag{SC.61}$$

In addition, we note that

$$\int_0^T g_3\left(c_s\right) ds = \int_0^T \frac{c_{ZY,s} - \beta_0 c_{ZZ,s}}{v_s^c} ds = 0.$$

Therefore, by Theorem 4 of Li, Todorov, and Tauchen (2016),

$$\Delta_n^{-1/2}\hat{S}_n(g_3) \xrightarrow{\mathcal{L}\text{-}s} \mathcal{MN}\left(0, \int_0^T \frac{c_{ZZ,s}}{v_s^c} ds\right). \tag{SC.62}$$

Note that the convergence (SC.61) is driven by a fixed number of Brownian increments around the jump times. By a routine argument, we can show that (SC.61) and (SC.62) hold jointly with $\mathcal{F}$-conditionally independent limits. Hence,

$$\Delta_n^{-1/2}\hat{S}_n(g_3) + \Delta_n^{-1/2}\left(Q_{n,ZY}^{\star} - \beta_0 Q_{n,ZZ}^{\star}\right) \xrightarrow{\mathcal{L}\text{-}s} \mathcal{MN}\left(0, \int_0^T \frac{d\left[Z,Z\right]_s}{v_s^c}\right). \tag{SC.63}$$

By Theorem 3 of Li, Todorov, and Tauchen (2016), $\hat{S}_n(g_2) \xrightarrow{\mathbb{P}} \int_0^T c_{ZZ,s}/v_s^c ds$. As a special case of (SC.30), we also have $Q^{\star}_{n,ZZ} \xrightarrow{\mathbb{P}} \sum_{\tau \in \mathcal{T}} \Delta Z_\tau^2 / v_\tau^c$. Hence,

$$\hat{S}_n(g_2) + Q^{\star}_{n,ZZ} \xrightarrow{\mathbb{P}} \int_0^T \frac{d\,[Z,Z]_s}{v_s^c}. \tag{SC.64}$$

The assertion of the theorem readily follows from (SC.60), (SC.63) and (SC.64). *Q.E.D.*

SC.3 Proofs for Theorem S5

In this subsection, we prove Theorem S5. Section SC.3.1 collects some notation. Section SC.3.2 contains the main proof. Technical lemmas are proved in Section SC.3.3.

SC.3.1 Notations and preliminary results

By a standard localization argument (see Section 4.4.1 of Jacod and Protter (2012)), we can, without loss of generality, strengthen Assumption S4 to the following version.

Assumption S7. *We have Assumption S4. Moreover, the processes A_Y and A_Z are bounded.*

We now introduce some additional notation. For notational simplicity, we set

$$g_n(j) \equiv g_n(j/h_n).$$

Since the continuous function $g(\cdot)$ is supported on $[0,1]$, $g_n(j)$ is non-zero only when $1 \leq j \leq h_n - 1$; we shall use this simple fact implicitly below without further mention. We denote

$$\tilde{X}_{n,i} \equiv \begin{pmatrix} \tilde{Z}_{n,i} \\ \tilde{Y}_{n,i} \end{pmatrix} \equiv \sum_j g_n(j)\, \Delta_{i+j}^n X, \tag{SC.65}$$

where the first-difference operator Δ_i^n is now defined with respect to the sampling basis, that is,

$$\Delta_i^n X \equiv X_{t(n,i)} - X_{t(n,i-1)}, \quad i \geq 1.$$

We denote the continuous component of X by

$$X_t^c \equiv \begin{pmatrix} Z_t^c \\ Y_t^c \end{pmatrix} \equiv X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s.$$

We also denote by $U_t = Y_t^c - \beta_0 Z_t^c$ the diffusive residual process. The instantaneous drift and the diffusion coefficient of U are given by, respectively,

$$b_{U,t} \equiv (-\beta_0, 1) b_t, \quad \sigma_{U,t} \equiv (-\beta_0, 1)\sigma_t, \quad t \geq 0. \tag{SC.66}$$

Recall from (SA.2) in the main text that $v_t \equiv (-\beta_0, 1)\sigma_t\sigma_t^\top(-\beta_0, 1)^\top$. Hence,

$$v_t = \sigma_{U,t}\sigma_{U,t}^\top. \tag{SC.67}$$

We consider a filtration given by $\mathcal{H}_t \equiv \mathcal{F}_t \vee \sigma(T_m : m \geq 1)$. Note that the random times $(T_m)_{m\geq 1}$ are independent of the Brownian motion W. Hence, W remains a Brownian motion with respect to the filtration $(\mathcal{H}_t)_{t\geq 0}$. For each $m \geq 1$, we denote by $i(n,m)$ the unique integer i such that $T_m \in (t(n,i-1), t(n,i)]$. The random integer $i(n,m)$ is $\mathcal{H}_0$-measurable.

We consider a sequence Ω_n of events defined by

$$\Omega_n \equiv \left\{ \begin{array}{c} \text{For each } i \text{ with } t(n,i) \leq T,\ (t(n,i-h_n), t(n,i+2h_n)] \\ \text{contains at most one jump time in } (T_m)_{m\geq 1}. \end{array} \right\}.$$

Under the maintained assumptions, the length of the interval $(t(n,i-h_n), t(n,i+2h_n)]$ shrinks to zero uniformly. Since the jumps have finite activity, $\mathbb{P}(\Omega_n) \to 1$. Therefore, we can restrict the calculations below to the set Ω_n without loss of generality. By doing so, we can suppose that each pre-averaged return contains at most one jump return.

Finally, we recall the concept of convergence in $\mathcal{F}$-conditional law. We write $\xi_n \xrightarrow{\mathcal{L}|\mathcal{F}} \xi$ if the $\mathcal{F}$-conditional law of ξ_n converges in probability to the $\mathcal{F}$-conditional law of ξ under any metric for the weak convergence of probability measures. See Appendix A of Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008) for more details and many useful results.

SC.3.2 Proof of Theorem S5

We now prove Theorem S5. The proof relies on several technical lemmas. For readability, we defer the proof of these lemmas to Section SC.3.3.

Lemma S1, below, collects some estimates that are repeatedly used in the sequel. Notations such as $\bar{J}_{Z,n,i}$ and $\tilde{J}_{Z,n,i}$ are interpreted as in (SA.16) and (SC.65), respectively.

Lemma S1. *The following statements hold under the assumptions of Theorem S5: for $p \geq 2$,*

$$\left|\bar{J}_{Z,n,i} - \tilde{J}_{Z,n,i}\right| + \left|\bar{J}_{Y,n,i} - \tilde{J}_{Y,n,i}\right| \leq K h_n^{-1}, \tag{SC.68}$$

$$\left|\bar{J}_{Y,n,i} - \beta_0 \bar{J}_{Z,n,i}\right| 1_{\Omega_n \cap \{\tilde{J}_{Z,n,i} \neq 0\}} \leq K h_n^{-1}, \tag{SC.69}$$

$$\mathbb{E}\left[\left|\bar{Y}^c_{n,i}\right|^p + \left|\bar{Z}^c_{n,i}\right|^p \middle| \mathcal{H}_0\right] \leq K_p (h_n \Delta_n)^{p/2}, \tag{SC.70}$$

$$\mathbb{E}\left[\left|\bar{\varepsilon}'_{Y,n,i}\right|^p + \left|\bar{\varepsilon}'_{Z,n,i}\right|^p \middle| \mathcal{H}_0\right] \leq K_p h_n^{-p/2}, \tag{SC.71}$$

$$\mathbb{E}\left[\left|\bar{Y}^c_{n,i} - \tilde{Y}^c_{n,i}\right|^2 + \left|\bar{Z}^c_{n,i} - \tilde{Z}^c_{n,i}\right|^2 \middle| \mathcal{H}_0\right] \leq K \Delta_n^{3/2}. \tag{SC.72}$$

As we shall show later, the asymptotics of $\hat{\beta}'_n$ is driven by the following variables:

$$\begin{cases} \zeta_{n,m-} \equiv \Delta_n^{-1/4} h_n^{-1} \sum_{k<0} \left(\sum_j g_n(j-k)\, g_n(j) \right) \Delta_{i(n,m)+k}^n U, \\ \zeta_{n,m+} \equiv \Delta_n^{-1/4} h_n^{-1} \sum_{k>0} \left(\sum_j g_n(j-k)\, g_n(j) \right) \Delta_{i(n,m)+k}^n U, \\ \zeta'_{Y,n,m} \equiv \Delta_n^{-1/4} h_n^{-1} \sum_{l=1}^{h_n} g_n(l)\, \bar{\varepsilon}'_{Y,n,i(n,m)-l}, \\ \zeta'_{Z,n,m} \equiv \Delta_n^{-1/4} h_n^{-1} \sum_{l=1}^{h_n} g_n(l)\, \bar{\varepsilon}'_{Z,n,i(n,m)-l}. \end{cases} \tag{SC.73}$$

Their asymptotic distributions are characterized by Lemma S2 and Lemma S3 below.

Lemma S2. $(\zeta_{n,m-}, \zeta_{n,m+})_{m\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\zeta_{m-}, \zeta_{m+})_{m\geq 1}$ *under the product topology.*

Lemma S3. $(\zeta'_{Y,n,m}, \zeta'_{Z,n,m})_{m\geq 1} \xrightarrow{\mathcal{L}|\mathcal{F}} (\zeta'_{Y,m}, \zeta'_{Z,m})_{m\geq 1}$ *under the product topology.*

We set

$$\mathcal{J}_n^* \equiv \{ i : (t(n,i), t(n,i+h_n)] \text{ contains at least one jump of } Z \}, \tag{SC.74}$$

which collects local windows on the sampling basis in which Z jumps. The proof of Theorem S5 relies on approximating $\hat{\beta}'_n$ via

$$\hat{\beta}'^{*}_n \equiv \frac{\sum_{i\in\mathcal{J}_n^*} \bar{Z}'_{n,i} \bar{Y}'_{n,i}}{\sum_{i\in\mathcal{J}_n^*} \bar{Z}'^{2}_{n,i}}. \tag{SC.75}$$

Theorem S5 is evidently implied by Proposition S1 below, which is followed by its proof.

Proposition S1. *Under the assumptions of Theorem S5,*

(a) the sequence $\Delta_n^{-1/4}(\hat{\beta}'^{*}_n - \beta_0)$ *converges stably in law to*

$$\frac{\sum_{m\geq 1: T_m \leq T} \Delta Z_{T_m} \left(\zeta_{m-} + \zeta_{m+} + \zeta'_{Y,m} - \beta_0 \zeta'_{Z,m} \right)}{\left(\int_0^1 g(s)^2\, ds \right) \sum_{m\geq 1: T_m\leq T} \Delta Z_{T_m}^2};$$

(b) $\hat{\beta}'_n - \hat{\beta}'^{*}_n = o_p(\Delta_n^{1/4})$.

Proof of Proposition S1(a). Step 1. In this step, we outline the proof. We note that

$$\Delta_n^{-1/4} \left(\hat{\beta}'^{*}_n - \beta_0 \right) = \frac{\Delta_n^{-1/4} h_n^{-1} \sum_{i\in\mathcal{J}_n^*} \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right)}{h_n^{-1} \sum_{i\in\mathcal{J}_n^*} \bar{Z}'^{2}_{n,i}}. \tag{SC.76}$$

In step 2, below, we show that

$$\frac{1}{h_n} \sum_{i\in\mathcal{J}_n^*} \bar{Z}'^{2}_{n,i} \xrightarrow{\mathbb{P}} \int_0^1 g(s)^2\, ds \sum_{m\geq 1: T_m\leq T} \Delta Z_{T_m}^2. \tag{SC.77}$$

The numerator of the right-hand side of (SC.76) can be rewritten as

$$\frac{1}{h_n \Delta_n^{1/4}} \sum_{i \in \mathcal{J}_n^*} \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right) = \sum_{m \geq 1: T_m \leq T} 1_{\{\Delta Z_{T_m} \neq 0\}} N_{n,m}, \tag{SC.78}$$

where

$$N_{n,m} \equiv \frac{1}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} \bar{Z}'_{n,i(n,m)-l} \left(\bar{Y}'_{n,i(n,m)-l} - \beta_0 \bar{Z}'_{n,i(n,m)-l} \right).$$

In step 3, below, we show that for each $m \geq 1$,

$$N_{n,m} = \Delta Z_{T_m} \left(\zeta_{n,m-} + \zeta_{n,m+} + \zeta'_{Y,n,m} - \beta_0 \zeta'_{Z,n,m} \right) + o_p(1). \tag{SC.79}$$

By Proposition 5 in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), we can combine Lemma S2 and Lemma S3 to deduce that

$$(\zeta_{n,m-}, \zeta_{n,m+}, \zeta'_{Y,n,m}, \zeta'_{Z,n,m})_{m \geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\zeta_{m-}, \zeta_{m+}, \zeta'_{Y,m}, \zeta'_{Z,m})_{m \geq 1}. \tag{SC.80}$$

From (SC.79) and (SC.80), we deduce

$$(N_{n,m})_{m \geq 1} \xrightarrow{\mathcal{L}\text{-}s} \Delta Z_{T_m} \left(\zeta_{m-} + \zeta_{m+} + \zeta'_{Y,m} - \beta_0 \zeta'_{Z,m} \right). \tag{SC.81}$$

The assertion of part (a) of Proposition S1 then follows from (SC.76), (SC.77), (SC.78) and (SC.81).

Step 2. In this step, we show (SC.77). We can rewrite

$$\frac{1}{h_n} \sum_{i \in \mathcal{J}_n^*} \bar{Z}'^2_{n,i} = \sum_{m \geq 1: T_m \leq T} 1_{\{\Delta Z_{T_m} \neq 0\}} \frac{1}{h_n} \sum_{l=1}^{h_n} \bar{Z}'^2_{n,i(n,m)-l}. \tag{SC.82}$$

We note that (recall the notation (SC.65))

$$\begin{aligned} \sum_{m \geq 1: T_m \leq T} 1_{\{\Delta Z_{T_m} \neq 0\}} \frac{1}{h_n} \sum_{l=1}^{h_n} \tilde{J}^2_{Z,n,i(n,m)-l} &= \sum_{m \geq 1: T_m \leq T} 1_{\{\Delta Z_{T_m} \neq 0\}} \frac{1}{h_n} \sum_{l=1}^{h_n} g_n(l)^2 \, \Delta Z^2_{T_m} \\ &\to \left(\int_0^1 g(s)^2 \, ds \right) \sum_{m \geq 1: T_m \leq T} \Delta Z^2_{T_m}. \end{aligned} \tag{SC.83}$$

Next, we note that

$$\bar{Z}'_{n,i} - \tilde{J}_{Z,n,i} = \bar{J}_{Z,n,i} - \tilde{J}_{Z,n,i} + \bar{Z}^c_{n,i} + \bar{\varepsilon}'_{Z,n,i}.$$

Since the random integer $i(n,m)$ is $\mathcal{H}_0$-measurable, Lemma S1 implies

$$\mathbb{E}\left[\left| \bar{Z}'_{n,i(n,m)-l} - \tilde{J}_{Z,n,i(n,m)-l} \right|^2 \middle| \mathcal{H}_0 \right] \leq K \left(h_n^{-2} + h_n \Delta_n + h_n^{-1} \right) \leq K \Delta_n^{1/2}. \tag{SC.84}$$

Therefore,

$$\sum_{m\geq 1:T_m\leq T} 1_{\{\Delta Z_{T_m}\neq 0\}} \frac{1}{h_n} \sum_{l=1}^{h_n} \left| \bar{Z}'_{n,i(n,m)-l} - \tilde{J}_{Z,n,i(n,m)-l} \right|^2 = O_p(\Delta_n^{1/2}) = o_p(1). \tag{SC.85}$$

With an appeal to the Cauchy–Schwarz inequality, we deduce (SC.77) from (SC.82), (SC.83) and (SC.85).

Step 3. We show (SC.79) in this step. We first approximate $N_{n,m}$ by

$$N'_{n,m} \equiv \frac{1}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} \tilde{J}_{Z,n,i(n,m)-l} \left(\bar{Y}'_{n,i(n,m)-l} - \beta_0 \bar{Z}'_{n,i(n,m)-l} \right).$$

Note that (SC.69), (SC.70) and (SC.71) imply

$$\mathbb{E}\left[\left| \bar{Y}'_{n,i(n,m)-l} - \beta_0 \bar{Z}'_{n,i(n,m)-l} \right|^2 \middle| \mathcal{H}_0 \right] \leq K\Delta_n^{1/2}. \tag{SC.86}$$

By (SC.84), (SC.86) and the Cauchy–Schwarz inequality, $\mathbb{E}\left| N_{n,m} - N'_{n,m} \right| \leq K\Delta_n^{1/4}$. Hence,

$$N_{n,m} = N'_{n,m} + o_p(1). \tag{SC.87}$$

Next, we observe that $N'_{n,m}$ can be rewritten as

$$N'_{n,m} = \frac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} g_n(l) \left(\bar{Y}'_{n,i(n,m)-l} - \beta_0 \bar{Z}'_{n,i(n,m)-l} \right).$$

We further approximate $N'_{n,m}$ by $N''_{n,m} \equiv N''_{n,m,1} + N''_{n,m,2}$, where

$$\begin{cases} N''_{n,m,1} \equiv \dfrac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \displaystyle\sum_{l=1}^{h_n} g_n(l) \left(\bar{Y}^c_{n,i(n,m)-l} - \beta_0 \bar{Z}^c_{n,i(n,m)-l} \right), \\ N''_{n,m,2} \equiv \dfrac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \displaystyle\sum_{l=1}^{h_n} g_n(l) \left(\bar{\varepsilon}'_{Y,n,i(n,m)-l} - \beta_0 \bar{\varepsilon}'_{Z,n,i(n,m)-l} \right). \end{cases}$$

Note that

$$N'_{n,m} - N''_{n,m} = \frac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} g_n(l) \left(\bar{J}_{Y,n,i(n,m)-l} - \beta_0 \bar{J}_{Z,n,i(n,m)-l} \right).$$

By (SC.69),

$$N'_{n,m} - N''_{n,m} = O_p(\Delta_n^{1/4}) = o_p(1). \tag{SC.88}$$

Let

$$\tilde{N}''_{n,m,1} \equiv \frac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} g_n(l) \left(\tilde{Y}^c_{n,i(n,m)-l} - \beta_0 \tilde{Z}^c_{n,i(n,m)-l} \right).$$

By (SC.72),

$$N''_{n,m,1} - \tilde{N}''_{n,m,1} = O_p(\Delta_n^{1/2}) = o_p(1). \tag{SC.89}$$

Recall that $U = Y^c - \beta_0 Z^c$. Hence, we can rewrite

$$\begin{aligned}
\tilde{N}''_{n,m,1} &= \frac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \sum_l g_n(l) \tilde{U}_{n,i(n,m)-l} \\
&= \frac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \sum_l g_n(l) \sum_j g_n(j) \Delta^n_{i(n,m)-l+j} U \\
&= \frac{\Delta Z_{T_m}}{h_n \Delta_n^{1/4}} \sum_k \sum_j g_n(j-k) g_n(j) \Delta^n_{i(n,m)+k} U.
\end{aligned}$$

From here, it is easy to see that $\tilde{N}''_{n,m,1} = \Delta Z_{T_m} (\zeta_{n,m-} + \zeta_{n,m+}) + o_p(1)$. By (SC.89), we further derive

$$N''_{n,m,1} = \Delta Z_{T_m} (\zeta_{n,m-} + \zeta_{n,m+}) + o_p(1). \tag{SC.90}$$

Finally, we observe that, by definition,

$$N''_{n,m,2} = \Delta Z_{T_m} (\zeta'_{Y,n,m} - \beta_0 \zeta'_{Z,n,m}). \tag{SC.91}$$

From (SC.87), (SC.88), (SC.90) and (SC.91), the claim (SC.79) readily follows. *Q.E.D.*

Proof of Proposition S1(b). Step 1. For notational simplicity, we denote $\mathcal{J}_n \equiv \{i : |\bar{Z}'_{n,i}| > \bar{v}_n\}$. In this step, we show that

$$h_n^{-1} \sum_{i \in \mathcal{J}_n \setminus \mathcal{J}_n^*} \bar{Z}'^2_{n,i} = o_p(1), \quad h_n^{-1} \sum_{i \in \mathcal{J}_n \setminus \mathcal{J}_n^*} \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right) = o_p(\Delta_n^{1/4}). \tag{SC.92}$$

We start with some estimates for each $i \notin \mathcal{J}_n^*$. For such i, Z does not jump in the interval $(t(n,i), t(n,i+h_n)]$. Therefore, only the first non-zero summand in (SA.16) may contain a jump. Since $g(\cdot)$ is Lipschitz continuous and $g(0) = 0$, the weight on this term is bounded by Kh_n^{-1}. Therefore, $|\bar{Z}'_{n,i} - \bar{Z}'^c_{n,i}| \leq Kh_n^{-1}$, where $Z'^c \equiv Z^c + \varepsilon'_Z$. By Lemma S1, for each $p \geq 2$, $\mathbb{E}[|\bar{Z}'^c_{n,i}|^p | \mathcal{H}_0] \leq K_p \Delta_n^{p/4}$. Hence, $\mathbb{E}[|\bar{Z}'_{n,i}|^p | \mathcal{H}_0] \leq K_p \Delta_n^{p/4}$. For each $q > 0$,

$$\mathbb{E}\left[\bar{Z}'^2_{n,i} 1_{\{|\bar{Z}'_{n,i}| > \bar{v}_n\}} \middle| \mathcal{H}_0 \right] \leq \mathbb{E}\left[\left| \bar{Z}'_{n,i} \right|^{2+q} \middle| \mathcal{H}_0 \right] / \bar{v}_n^q \leq K_q \Delta_n^{1/2 + q(1/4 - \varpi')}.$$

From here, we deduce that for any $q > 0$,

$$h_n^{-1} \sum_{i \in \mathcal{J}_n \setminus \mathcal{J}_n^*} \bar{Z}'^2_{n,i} = O_p\left(\Delta_n^{q(1/4 - \varpi')} \right). \tag{SC.93}$$

Since $\varpi' \in (0, 1/4)$, by taking q sufficiently large in (SC.93), we deduce

$$h_n^{-1} \sum_{i \in \mathcal{J}_n \setminus \mathcal{J}_n^*} \bar{Z}_{n,i}'^2 = o_p(\Delta_n^{1/2}), \tag{SC.94}$$

which implies the first part of (SC.92).

From Lemma S1, it is easy to see that

$$h_n^{-1} \sum_{i \in \mathcal{J}_n \setminus \mathcal{J}_n^*} \left(\bar{Y}_{n,i}' - \beta_0 \bar{Z}_{n,i}'\right)^2 \leq h_n^{-1} \sum_i \left(\bar{Y}_{n,i}' - \beta_0 \bar{Z}_{n,i}'\right)^2 = O_p(1). \tag{SC.95}$$

By the Cauchy–Schwarz inequality, (SC.94) and (SC.95), we further deduce the second part of (SC.92).

Step 2. In this step, we show that

$$h_n^{-1} \sum_{i \in \mathcal{J}_n^* \setminus \mathcal{J}_n} \bar{Z}_{n,i}'^2 = o_p(1). \tag{SC.96}$$

We consider a positive process $f_{n,m}(\cdot)$ given by

$$f_{n,m}(s) \equiv \bar{Z}_{n,i(n,m)-1-\lfloor h_n s \rfloor}'^2 1_{\left\{\left|\bar{Z}_{n,i(n,m)-1-\lfloor h_n s \rfloor}'\right| \leq \bar{v}_n\right\} \cap \{\Delta Z_{T_m} \neq 0\}}. \tag{SC.97}$$

We can then rewrite

$$h_n^{-1} \sum_{i \in \mathcal{J}_n^* \setminus \mathcal{J}_n} \bar{Z}_{n,i}'^2 = \sum_{m \geq 1: T_m \leq T} \int_0^1 f_{n,m}(s) ds. \tag{SC.98}$$

From Lemma S1, it is easy to see that

$$\mathbb{E}\left[f_{n,m}(s)^2 | \mathcal{H}_0\right] \leq K. \tag{SC.99}$$

We now consider the behavior of $f_{n,m}(s)$ for each $s \in (0, 1)$. From Lemma S1, we see that

$$\bar{Z}_{n,i(n,m)-1-\lfloor h_n s \rfloor}' = \tilde{J}_{Z,n,i(n,m)-1-\lfloor h_n s \rfloor} + o_p(1).$$

We further note that $\tilde{J}_{Z,n,i(n,m)-1-\lfloor h_n s \rfloor} = g(s) \Delta Z_{T_m} + o_p(1)$. Hence,

$$\bar{Z}_{n,i(n,m)-1-\lfloor h_n s \rfloor}' = g\left(s\right) \Delta Z_{T_m} + o_p(1).$$

By assumption, $g\left(s\right) > 0$ for $s \in (0, 1)$. Since $\bar{v}_n \to 0$, we deduce

$$\mathbb{P}\left(\left\{\left|\bar{Z}_{n,i(n,m)-1-\lfloor h_n s \rfloor}'\right| \leq \bar{v}_n\right\} \cap \{\Delta Z_{T_m} \neq 0\} \middle| \mathcal{H}_0\right) = o_p(1). \tag{SC.100}$$

Since $\bar{Z}_{n,i(n,m)-1-\lfloor h_n s \rfloor}'^2 = O_p(1)$, (SC.100) implies that for each $s \in (0, 1)$,

$$f_{n,m}\left(s\right) = o_p(1). \tag{SC.101}$$

Note that (SC.99) implies that the sequence $f_{n,m}(s)$, $n \geq 1$, is uniformly integrable under the $\mathcal{H}_0$-conditional probability. This condition allows us to deduce $\mathbb{E}[f_{n,m}(s)|\mathcal{H}_0] = o_p(1)$ from (SC.101). Note that $\mathbb{E}[f_{n,m}(s)|\mathcal{H}_0]$ is bounded because of (SC.99). By Fubini's theorem and the bounded convergence theorem, we deduce that the right-hand side of (SC.98) is $o_p(1)$. From here, (SC.96) readily follows.

Step 3. In this step, we show that

$$\Delta_n^{-1/4} h_n^{-1} \sum_{i \in \mathcal{J}_n^* \backslash \mathcal{J}_n} \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right) = o_p(1). \tag{SC.102}$$

We use the same argument as in step 2, except that we replace the term $\bar{Z}'^2_{n,i(n,m)-1-\lfloor h_n s \rfloor}$ in (SC.97) by

$$\Delta_n^{-1/4} \left| \bar{Z}'_{n,i(n,m)-1-\lfloor h_n s \rfloor} \right| \left| \bar{Y}'_{n,i(n,m)-1-\lfloor h_n s \rfloor} - \beta_0 \bar{Z}'_{n,i(n,m)-1-\lfloor h_n s \rfloor} \right|.$$

By Lemma S1, we can verify that the term in the above display is $O_p(1)$ and (SC.99) still holds. With these modifications, the same argument in step 2 yields (SC.102).

Step 4. Combining (SC.92), (SC.96) and (SC.102), we deduce that

$$\begin{cases} h_n^{-1} \sum_i \bar{Z}'^2_{n,i} 1_{\{|\bar{Z}'_{n,i}| > \bar{v}_n\}} - h_n^{-1} \sum_{i \in \mathcal{J}_n^*} \bar{Z}'^2_{n,i} = o_p(1), \\ h_n^{-1} \sum_i \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right) 1_{\{|\bar{Z}'_{n,i}| > \bar{v}_n\}} - h_n^{-1} \sum_{i \in \mathcal{J}_n^*} \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right) = o_p(\Delta_n^{1/4}). \end{cases} \tag{SC.103}$$

From the proof of Proposition S1(a), we also have

$$\begin{cases} h_n^{-1} \sum_{i \in \mathcal{J}_n^*} \bar{Z}'^2_{n,i} = O_p(1), \\ h_n^{-1} \sum_{i \in \mathcal{J}_n^*} \bar{Z}'_{n,i} \left(\bar{Y}'_{n,i} - \beta_0 \bar{Z}'_{n,i} \right) = O_p(\Delta_n^{1/4}). \end{cases} \tag{SC.104}$$

Recall the definitions of $\hat{\beta}'_n$ and $\hat{\beta}'^*_n$ from (SA.17) and (SC.75). From (SC.103) and (SC.104), we readily deduce the assertion of Proposition S1(b). *Q.E.D.*

SC.3.3 Proofs of technical lemmas

Proof of Lemma S1. We start with (SC.68). From (SA.16) and (SC.65), we observe that $\bar{J}_{Z,n,i}$ and $\tilde{J}_{Z,n,i}$ can be rewritten as

$$\begin{aligned} \bar{J}_{Z,n,i} &= \sum_k \sum_{j=i_{Z,n,k-1}+1}^{i_{Z,n,k}} g_n \left(i_{Z,n,k} - i \right) \Delta_j^n J_Z, \\ \tilde{J}_{Z,n,i} &= \sum_k \sum_{j=i_{Z,n,k-1}+1}^{i_{Z,n,k}} g_n \left(j - i \right) \Delta_j^n J_Z. \end{aligned}$$

We can then rewrite $\bar{J}_{Z,n,i} - \tilde{J}_{Z,n,i} = \sum_j \tilde{g}_{Z,n,i}(j) \Delta_j^n J_Z$ where

$$\tilde{g}_{Z,n,i}(j) \equiv g_n(i_{Z,n,k} - i) - g_n(j - i), \quad \text{if} \quad i_{Z,n,k-1} < j \leq i_{Z,n,k}. \tag{SC.105}$$

Since g is Lipschitz continuous and $\delta_{Z,n,k}$ is bounded,

$$|\tilde{g}_{Z,n,i}(j)| \leq K h_n^{-1}. \tag{SC.106}$$

Since the jumps of Z are finitely active, we readily deduce $|\bar{J}_{Z,n,i} - \tilde{J}_{Z,n,i}| \leq K h_n^{-1}$. The part of (SC.68) concerning Y can be shown similarly.

Turning to (SC.69), we first observe that in restriction to $\Omega_n \cap \{\tilde{J}_{Z,n,i} \neq 0\}$, the interval $(t(n,i), t(n, i+h_n)]$ contains exactly one jump. Hence, $\tilde{J}_{Y,n,i} - \beta_0 \tilde{J}_{Z,n,i} = 0$. From here and (SC.68), (SC.69) readily follows.

We now show (SC.70). By parts (a,c) of Assumption S5, $|t(Z,n,k) - t(Z,n,k-1)| \leq K\Delta_n$. By using a standard estimate for continuous Itô semimartingales, we derive

$$\mathbb{E}\left[\left|\bar{Z}_{n,i}^c\right|^p \middle| \mathcal{H}_0\right] \leq K_p \left(\sum_k g_n(i_{Z,n,k} - i)\Delta_n\right)^p + K_p \left(\sum_k g_n(i_{Z,n,k} - i)^2 \Delta_n\right)^{p/2}.$$

We further note that the sequence $g_n(i_{Z,n,k} - i)$, $k \geq 0$, is bounded and contains at most h_n non-zero terms. Hence, $\mathbb{E}[|\bar{Z}_{n,i}^c|^p|\mathcal{H}_0] \leq K_p(h_n\Delta_n)^{p/2}$. Similarly, $\mathbb{E}[|\bar{Y}_{n,i}^c|^p|\mathcal{H}_0] \leq K_p(h_n\Delta_n)^{p/2}$. The inequality (SC.70) readily follows.

To show (SC.71), we first observe

$$\bar{\varepsilon}'_{Z,n,i} = \sum_k \left[g_n(i_{Z,n,k} - i) - g_n(i_{Z,n,k+1} - i)\right] \varepsilon'_{Z,t(Z,n,k)}.$$

Since $g(\cdot)$ is Lipschitz continuous and $\delta_{Z,n,k}$ is bounded,

$$|g_n(i_{Z,n,k} - i) - g_n(i_{Z,n,k+1} - i)| \leq K h_n^{-1}.$$

We further note that the noise terms are $\mathcal{F}$-conditionally independent with bounded moments. By the Burkholder–Davis–Gundy inequality, we see that $\mathbb{E}[|\bar{\varepsilon}'_{Z,n,i}|^p|\mathcal{H}_0] \leq K_p h_n^{-p/2}$. Similarly, we can show $\mathbb{E}[|\bar{\varepsilon}'_{Y,n,i}|^p|\mathcal{H}_0] \leq K_p h_n^{-p/2}$, which further implies (SC.71).

We now show (SC.72). Recall (SC.105). We can rewrite $\bar{Z}_{n,i}^c - \tilde{Z}_{n,i}^c = \sum_j \tilde{g}_{Z,n,i}(j) \Delta_j^n Z^c$. Note that $\tilde{g}_{Z,n,i}(j)$ is non-zero for at most $2h_n$ terms. By using a standard estimate for continuous Itô semimartingales and then using (SC.106), we deduce

$$\begin{aligned}\mathbb{E}\left[\left|\bar{Z}_{n,i}^c - \tilde{Z}_{n,i}^c\right|^2 \middle| \mathcal{H}_0\right] &\leq K \left(\sum_j \tilde{g}_{Z,n,i}(j)\Delta_n\right)^2 + K \sum_j \tilde{g}_{Z,n,i}^2(j)\Delta_n \\ &\leq K\Delta_n^2 + K h_n^{-1}\Delta_n.\end{aligned}$$

Hence, $\mathbb{E}[|\bar{Z}^c_{n,i} - \tilde{Z}^c_{n,i}|^2|\mathcal{H}_0] \leq K\Delta_n^{3/2}$. Similarly, $\mathbb{E}[|\bar{Y}^c_{n,i} - \tilde{Y}^c_{n,i}|^2|\mathcal{H}_0] \leq K\Delta_n^{3/2}$. From here, (SC.72) readily follows. *Q.E.D.*

Proof of Lemma S2. Recall the definitions in (SC.66). For each $k > 0$, we can decompose

$$\Delta^n_{i(n,m)+k}U = A_{n,m,k} + B_{n,m,k} + \sigma_{U,T_m}\Delta^n_{i(n,m)+k}W,$$

where

$$A_{n,m,k} \equiv \int_{t(n,i(n,m)+k-1)}^{t(n,i(n,m)+k)} b_{U,s}ds, \quad B_{n,m,k} \equiv \int_{t(n,i(n,m)+k-1)}^{t(n,i(n,m)+k)} \left(\sigma_{U,s} - \sigma_{U,T_m}\right) dW_s.$$

We note that

$$|A_{n,m,k}| \leq K\Delta_n, \quad \mathbb{E}\left[|B_{n,m,k}|^2 \middle| \mathcal{H}_0\right] \leq K\int_{t(n,i(n,m)+k-1)}^{t(n,i(n,m)+k)} \mathbb{E}\left[\|\sigma_{U,s} - \sigma_{U,T_m}\|^2 \middle| \mathcal{H}_0\right] ds. \quad \text{(SC.107)}$$

We approximate $\zeta_{n,m+}$ using

$$\zeta^*_{n,m+} \equiv \Delta_n^{-1/4}h_n^{-1}\sum_{k=1}^{h_n-1}\left(\sum_j g_n\left(j-k\right)g_n\left(j\right)\right)\sigma_{U,T_m}\Delta^n_{i(n,m)+k}W.$$

Observe that

$$\begin{aligned}
&\mathbb{E}\left[\left|\zeta_{n,m+} - \zeta^*_{n,m+}\right|^2 \middle| \mathcal{H}_0\right] \\
&\leq K\left(\Delta_n^{-1/4}h_n^{-1}\sum_{k>0}\left(\sum_j g_n\left(j-k\right)g_n\left(j\right)\right)A_{n,m,k}\right)^2 \\
&\qquad +K\Delta_n^{-1/2}h_n^{-2}\sum_{k>0}\left(\sum_j g_n\left(j-k\right)g_n\left(j\right)\right)^2 \mathbb{E}\left[B^2_{n,m,k}\middle| \mathcal{H}_0\right] \\
&\leq K\Delta_n^{1/2} + K\Delta_n^{-1/2}\int_{t(n,i(n,m))}^{t(n,i(n,m)+h_n)} \mathbb{E}\left[\|\sigma_{U,s} - \sigma_{U,T_m}\|^2 \middle| \mathcal{H}_0\right] ds \\
&\leq K\Delta_n^{1/2} + K\mathbb{E}\left[\sup_{s\in[T_m,T_m+K\Delta_n^{1/2}]} \|\sigma_{U,s} - \sigma_{U,T_m}\|^2 \middle| \mathcal{H}_0\right], \qquad \text{(SC.108)}
\end{aligned}$$

where the first inequality is derived by observing that $B_{n,m,k}$, $k \geq 1$, are martingale differences; the second inequality holds because of (SC.107); the third inequality follows from Assumption S5. By the bounded convergence theorem and the right-continuity of the process σ_U, we deduce that the majorant side of (SC.108) converges to zero in expectation. Therefore,

$$\zeta_{n,m+} = \zeta^*_{n,m+} + o_p(1). \qquad \text{(SC.109)}$$

Similarly, we can approximate $\zeta_{n,m-}$ using

$$\zeta^*_{n,m-} \equiv \Delta_n^{-1/4} h_n^{-1} \sum_{k=-(h_n-1)}^{-1} \left(\sum_j g_n\left(j-k\right) g_n\left(j\right) \right) \sigma_{U,t(n,i(n,m)-h_n)} \Delta^n_{i(n,m)+k} W,$$

and show that

$$\zeta_{n,m-} = \zeta^*_{n,m-} + o_p(1). \tag{SC.110}$$

In view of (SC.109) and (SC.110), to prove Lemma S2, it remains to derive the stable convergence in law of $(\zeta^*_{n,m-}, \zeta^*_{n,m+})_{m\geq 1}$ towards $(\zeta_{m-}, \zeta_{m+})_{m\geq 1}$. We observe that $\zeta^*_{n,m-}$ and $\zeta^*_{n,m+}$ are sums of martingale difference arrays. We then use Theorem IX.7.28 of Jacod and Shiryaev (2003) to deduce the stable convergence in law of these sequences (arguing similar to the proof of Theorem 4.3.1 of Jacod and Protter (2012)). The limiting distributions of $\zeta^*_{n,m-}$ and $\zeta^*_{n,m+}$ are, conditionally on $\mathcal{F}$, mutually independent and centered Gaussian with their asymptotic variances respectively given by the limits (in probability) of

$$\begin{cases} \Sigma_{n,m-} \equiv \Delta_n^{-1/2} \sum_{k<0} \left(\frac{1}{h_n} \sum_j g_n\left(j-k\right) g_n\left(j\right) \right)^2 v_{t(n,i(n,m)-h_n)} \Delta_{n,i(n,m)+k}, \\ \Sigma_{n,m+} \equiv \Delta_n^{-1/2} \sum_{k>0} \left(\frac{1}{h_n} \sum_j g_n\left(j-k\right) g_n\left(j\right) \right)^2 v_{T_m} \Delta_{n,i(n,m)+k}. \end{cases} \tag{SC.111}$$

It remains to verify that (recall (SA.18))

$$\Sigma_{n,m-} \xrightarrow{\mathbb{P}} \mathbb{E}\left[\zeta^2_{m-} | \mathcal{F}\right], \quad \Sigma_{n,m+} \xrightarrow{\mathbb{P}} \mathbb{E}\left[\zeta^2_{m+} | \mathcal{F}\right]. \tag{SC.112}$$

To this end, we observe that

$$\begin{aligned} \Sigma_{n,m-} &= \Delta_n^{1/2} f\left(T_m\right) \sum_{k<0} \left(\frac{1}{h_n} \sum_j g_n\left(j-k\right) g_n\left(j\right) \right)^2 v_{t(n,i(n,m)-h_n)} \\ &\quad + \Delta_n^{1/2} \sum_{k<0} \left(\frac{1}{h_n} \sum_j g_n\left(j-k\right) g_n\left(j\right) \right)^2 v_{t(n,i(n,m)-h_n)} \left(\frac{\Delta_{n,i(n,m)+k}}{\Delta_n} - f\left(T_m\right) \right). \end{aligned} \tag{SC.113}$$

The second term on the right-hand side of (SC.113) is bounded in absolute value by

$$K h_n^{-1} \sum_{k=-(h_n-1)}^{-1} \left| \frac{\Delta_{n,i(n,m)+k}}{\Delta_n} - f\left(T_m\right) \right|,$$

which converges to zero pointwise under Assumption S5. Since $\Delta_n^{1/2} \asymp \theta h_n^{-1}$, the first term on the right-hand side of (SC.113) converges to

$$\theta v_{T_m -} f\left(T_m\right) \int_{-1}^{0} \left(\int g\left(s-u\right) g\left(s\right) ds \right)^2 du.$$

From here, the first convergence in (SC.112) follows. The second convergence of (SC.112) can be proved similarly. This finishes the proof for Lemma S2. *Q.E.D.*

Proof of Lemma S3. Step 1. In this step, we show that for each $m \geq 1$,

$$\zeta'_{Y,n,m} \xrightarrow{\mathcal{L}|\mathcal{F}} \zeta'_{Y,m}. \tag{SC.114}$$

To simplify notations, we denote $i^* = i(n,m)$. Observe that

$$\begin{aligned}
\zeta'_{Y,n,m} &= \frac{1}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} g_n(l)\, \bar{\varepsilon}'_{Y,n,i^*-l} \\
&= \frac{1}{h_n \Delta_n^{1/4}} \sum_{l=1}^{h_n} g_n(l) \sum_k g_n(i_{Y,n,k} - i^* + l) \left(\varepsilon'_{Y,t(Y,n,k)} - \varepsilon'_{Y,t(Y,n,k-1)}\right) \\
&= -\frac{1}{h_n \Delta_n^{1/4}} \sum_k \left\{ \sum_{l=1}^{h_n} g_n(l) \left[g_n(i_{Y,n,k+1} - i^* + l) - g_n(i_{Y,n,k} - i^* + l) \right] \right\} \varepsilon'_{Y,t(Y,n,k)}.
\end{aligned}$$

We further decompose

$$\zeta'_{Y,n,m} = \zeta'^{*}_{Y,n,m} + R_n,$$

where

$$\begin{aligned}
\zeta'^{*}_{Y,n,m} &\equiv \sum_k \zeta'^{*}_{Y,n,m}(k), \\
\zeta'^{*}_{Y,n,m}(k) &\equiv -\frac{1}{h_n \Delta_n^{1/4}} \left(\sum_{l=1}^{h_n} g_n(l)\, g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n} \right) \frac{\delta_{Y,n,k+1}}{h_n} \right) \varepsilon'_{Y,t(Y,n,k)}, \\
R_n &\equiv -\frac{1}{h_n \Delta_n^{1/4}} \sum_k \left\{ \sum_{l=1}^{h_n} g_n(l) \left[g_n(i_{Y,n,k+1} - i^* + l) \right.\right. \\
&\qquad \left.\left. - g_n(i_{Y,n,k} - i^* + l) - g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n} \right) \frac{\delta_{Y,n,k+1}}{h_n} \right] \right\} \varepsilon'_{Y,t(Y,n,k)}.
\end{aligned}$$

Since g' is Lipschitz continuous,

$$\left| g_n(i_{Y,n,k+1} - i^* + l) - g_n(i_{Y,n,k} - i^* + l) - g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n} \right) \frac{\delta_{Y,n,k+1}}{h_n} \right| \leq K h_n^{-2}. \tag{SC.115}$$

Since the noise terms $\varepsilon'_{Y,t(Y,n,k)}$, $k \geq 0$, are $\mathcal{F}$-conditionally independent with zero mean and bounded variance, we have

$$\begin{aligned}
\mathbb{E}\left[R_n^2 | \mathcal{F}\right] &\leq \frac{K}{h_n^2 \Delta_n^{1/2}} \sum_k \left| \sum_{l=1}^{h_n} g_n(l) \left[g_n(i_{Y,n,k+1} - i^* + l) \right.\right. \\
&\qquad \left.\left. - g_n(i_{Y,n,k} - i^* + l) - g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n} \right) \frac{\delta_{Y,n,k+1}}{h_n} \right] \right|^2.
\end{aligned}$$

Here, the summation over k involves at most $2h_n$ nonzero terms. By (SC.115), we further deduce

$$\mathbb{E}\left[R_n^2|\mathcal{F}\right] \leq K h_n^{-2} \to 0. \tag{SC.116}$$

Therefore, to show (SC.114), it suffices to show that

$$\zeta_{Y,n,m}^{\prime *} \xrightarrow{\mathcal{L}|\mathcal{F}} \zeta_{Y,m}^{\prime}. \tag{SC.117}$$

We note that for each fixed m, the array $\zeta_{Y,n,m}^{\prime *}(k)$ is $\mathcal{F}$-conditionally independent with zero mean. We shall use the Lindeberg–Lévy central limit theorem under the $\mathcal{F}$-conditional probability to deduce (SC.117). It is easy to verify Lyaponov's condition. Indeed,

$$\sum_k \mathbb{E}\left[\left|\zeta_{Y,n,m}^{\prime *}(k)\right|^3 \middle| \mathcal{F}\right] \leq K \Delta_n^{1/4} \to 0.$$

It remains to compute the ($\mathcal{F}$-conditional) asymptotic variance of $\zeta_{Y,n,m}^{\prime *}$, which is given by the limit of

$$\begin{aligned} \Sigma_{Y,n,m}^{\prime} &\equiv \frac{1}{h_n^2 \Delta_n^{1/2}} \sum_k \left(\sum_{l=1}^{h_n} g_n(l)\, g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n}\right) \frac{\delta_{Y,n,k+1}}{h_n} \right)^2 A_{Y,t(Y,n,k)} \\ &= \frac{1}{h_n \Delta_n^{1/2}} \sum_k \frac{\delta_{Y,n,k+1}}{h_n} \left(\frac{1}{h_n} \sum_{l=1}^{h_n} g_n(l)\, g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n}\right) \right)^2 \delta_{Y,n,k+1} A_{Y,t(Y,n,k)}. \end{aligned}$$

Observe that, in the above display, the summand indexed by k is non-zero only when $t(Y,n,k)$ falls in a local window around T_m with length (in calendar time) shrinking to zero. Since the process A_Y is continuous and (SA.14) holds,

$$\Sigma_{Y,n,m}^{\prime} = \frac{A_{Y,T_m} \phi_Y(T_m)}{h_n \Delta_n^{1/2}} \sum_k \frac{\delta_{Y,n,k+1}}{h_n} \left(\frac{1}{h_n} \sum_{l=1}^{h_n} g_n(l)\, g'\left(\frac{i_{Y,n,k} - i^* + l}{h_n}\right) \right)^2 + o_p(1).$$

By applying a Riemann approximation, we deduce

$$\Sigma_{Y,n,m}^{\prime} \xrightarrow{\mathbb{P}} \theta^{-1} A_{Y,T_m} \phi_Y(T_m) \int \left(\int_0^1 g(s)\, g'(s+u)\, ds \right)^2 du \equiv \mathbb{E}\left[\zeta_{Y,m}^{\prime 2}|\mathcal{F}\right].$$

This finishes the proof of (SC.117) and, hence, that of (SC.114).

Step 2. In this step, we prove the assertion of the lemma. We observe that for $m \neq m'$, the noise terms in $\zeta_{Y,n,m}^{\prime}$ do not overlap with those in $\zeta_{Y,n,m'}^{\prime}$ when n is large. Since the noise terms are $\mathcal{F}$-conditionally independent, for each $\bar{m} \geq 1$, the finite collection $(\zeta_{Y,n,m}^{\prime})_{1 \leq m \leq \bar{m}}$ forms an $\mathcal{F}$-conditional independency for large n. Therefore, $(\zeta_{Y,n,m}^{\prime})_{m \geq 1} \xrightarrow{\mathcal{L}|\mathcal{F}} (\zeta_{Y,m}^{\prime})_{m \geq 1}$ under the product topology. Similarly, $(\zeta_{Z,n,m}^{\prime})_{m \geq 1} \xrightarrow{\mathcal{L}|\mathcal{F}} (\zeta_{Z,m}^{\prime})_{m \geq 1}$. By Assumption S4, $\zeta_{Y,n,m}^{\prime}$ and $\zeta_{Z,n,m}^{\prime}$ are $\mathcal{F}$-conditionally independent. The assertion of the lemma readily follows. *Q.E.D.*

References

Andersen, T. G., T. Bollerslev, F. X. Diebold, and C. Vega (2003): "Micro Effects of Macro Announcements: Real-Time Price Discovery in Foreign Exchange," *American Economic Review*, 93, 38–62.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008): "Designing Realized Kernels to Measure the ex post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76, 1481–1536.

Barndorff-Nielsen, O. E., and N. Shephard (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2, 1–37.

Barndorff-Nielsen, O. E., and N. Shephard (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4, 1–30.

Comte, F., and E. Renault (1996): "Long Memory Continuous Time Models," *Journal of Econometrics*, 73, 101–149.

Gobbi, F., and C. Mancini (2012): "Identifying the Brownian Covariation from the Co-jumps given Discrete Observations," *Econometric Theory*, 28, 249–273.

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.

Jacod, J., M. Podolskij, and M. Vetter (2010): "Limit Theorems for Moving Averages of Discretized Processes Plus Noise," *Annals of Statistics*, 38, 1478–1545.

Jacod, J., and P. Protter (2012): *Discretization of Processes*. Springer.

Jacod, J., and M. Rosenbaum (2013): "Quarticity and Other Functionals of Volatility: Efficient Estimation," *Annals of Statistics*, 41, 1462–1484.

Jacod, J., and A. N. Shiryaev (2003): *Limit Theorems for Stochastic Processes*. Springer-Verlag, New York, second edn.

Jeganathan, P. (1982): "On the Asymptotic Theory of Estimation when the Limit of the Log-likelihood is Mixed Normal," *Sankhya, Ser. A*, 44, 173–212.

——— (1983): "Some Asymptotic Properties of Risk Functions When the Limit of the Experiment Is Mixed Normal," *Sankhya: The Indian Journal of Statistics, Series A (1961-2002)*, 45(1), 66–87.

Li, J. (2013): "Robust Estimation and Inference for Jumps in Noisy High Frequency Data: A Local-to-Continuity Theory for the Pre-averaging Method," *Econometrica*, 81, 1673–1693.

Li, J., V. Todorov, and G. Tauchen (2016): "Adaptive Estimation of Continuous-Time Regression Models using High-Frequency Data," Discussion paper, Duke Univeristy.

Podolskij, M., and M. Vetter (2009): "Estimation of Volatility Functionals in the Simultaneous Presence of Microstructure Noise and Jumps," *Bernoulli*, 15, 634–658.

Reiss, M., V. Todorov, and G. Tauchen (2015): "Nonparametric Test for a Constant Beta between Ito Semimartingales based on High-Frequency Data," *Stochastic Processes and Their Applications*, 125, 2955–2988.

Todorov, V., and T. Bollerslev (2010): "Jumps and Betas: A New Framework for Disentangling and Estimating Systematic Risks," *Journal of Econometrics*, 157, 220–235.

van der Vaart, A. W. (1998): *Asymptotic Statistics*. Cambridge University Press.