Supplemental Appendix for "Robust Jump Regressions"*

Jia Li† Viktor Todorov‡ George Tauchen§

December 28, 2015

Abstract

This document contains two appendices. Supplemental Appendix A contains the proofs for the results in the main text. Supplemental Appendix B provides an example that illustrates the adverse consequence of mistakenly ignoring measurement errors in the jump regression setting.

*Li's and Todorov's research have been partially supported by NSF grants SES-1326819 and SES-0957330 and SES-1530748, respectively. Yuan Xue provided excellent research assistance.

†Department of Economics, Duke University.

‡Department of Finance, Northwestern University.

§Department of Economics, Duke University.

Supplemental Appendix A: Proofs

We prove the results in the main text in this appendix. Below, we use K to denote a generic constant that may change from line to line. This constant does not depend on the index of a process or a series. We sometimes emphasize its dependence on some parameter v by writing K_v. We write "w.p.a.1" for "with probability approaching one." By a standard localization procedure (see, e.g., Section 4.4.1 of [5]), we can strengthen Assumptions 2 and 4 to the following versions without loss of generality.

Assumption S1. *We have Assumption 2. Moreover, the processes* $(\mathbf{b}_t)_{t\geq 0}$, $(\boldsymbol{\sigma}_t)_{t\geq 0}$ *and* $(\mathbf{X}_t)_{t\geq 0}$ *are bounded.*

Assumption S2. *We have Assumption 4. Moreover, the process* $(\mathbf{a}_t)_{t\geq 0}$ *is bounded.*

For notational simplicity, we denote the continuous part of $\mathbf{X}$ by $\mathbf{X}^c$, that is,

$$\mathbf{X}_t^c = \mathbf{X}_0 + \int_0^t \mathbf{b}_s ds + \int_0^t \boldsymbol{\sigma}_s d\mathbf{W}_s.$$

The notations Y^c and $\mathbf{Z}^c$ are defined similarly.

SA.1 Proof of Theorem 1

For each $p \geq 1$, let $i_{n,p}$ denote the unique integer i such that $\tau_p \in ((i-1)\Delta_n, i\Delta_n]$. Since the jumps of $\mathbf{X}$ are finitely active, each interval $((i-1)\Delta_n, i\Delta_n]$ contains at most one jump w.p.a.1. By Proposition 1 of [9], $\mathbb{P}(\mathcal{J}_n = \mathcal{J}_n^*) \to 1$. Therefore, w.p.a.1,

$$\begin{aligned} M_n(\mathbf{h}) &= \Delta_n^{-q/2} \sum_{p\in\mathcal{P}} \rho\left(\Delta_{i_{n,p}}^n Y - \left(\boldsymbol{\beta}^* + \Delta_n^{1/2}\mathbf{h}\right)^\top \Delta_{i_{n,p}}^n \mathbf{Z}\right) \\ &= \sum_{p\in\mathcal{P}} \rho\left(\Delta_n^{-1/2}\left(\Delta_{i_{n,p}}^n Y^c - \boldsymbol{\beta}^{*\top}\Delta_{i_{n,p}}^n \mathbf{Z}^c\right) - \mathbf{h}^\top \Delta_{i_{n,p}}^n \mathbf{Z}\right), \end{aligned}$$

where the second equality is due to Assumption 1 and (1). By Proposition 4.4.10 of [5],

$$\Delta_n^{-1/2}\left(\Delta_{i_{n,p}}^n Y^c - \boldsymbol{\beta}^{*\top}\Delta_{i_{n,p}}^n \mathbf{Z}^c\right)_{p\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\varsigma_p)_{p\geq 1}.$$

It is easy to see that $\Delta_{i_{n,p}}^n \mathbf{Z} \to \Delta \mathbf{Z}_{\tau_p}$. Since $\rho(\cdot)$ is convex, it is also continuous on $\mathbb{R}$. We then deduce (12) using the continuous mapping theorem and the properties of stable convergence (see (2.2.5) in [5]).

Next, we show that $\hat{\mathbf{h}}_n \xrightarrow{\mathcal{L}\text{-}s} \hat{\mathbf{h}}$ in restriction to Ω_0 using a convexity argument similar to Lemma A of [7]. We need to adapt this argument to the case of stable convergence. Fix any bounded $\mathcal{F}$-measurable random variable ξ. By Proposition VIII.5.33 in [6], it suffices to show that $(\hat{\mathbf{h}}_n, \xi)$ converges in law to $(\hat{\mathbf{h}}, \xi)$ in restriction to Ω_0. Let D be a countable dense subset of $\mathbb{R}^{d-1}$ and $\xi_0 \equiv 1_{\Omega_0}$. We consider $(M_n(\mathbf{h})_{\mathbf{h}\in D}, \xi, \xi_0)$ as a $\mathbb{R}^\infty$-valued random variable, where $\mathbb{R}^\infty$ is equipped with the product Euclidean topology. By (12), $(M_n(\mathbf{h})_{\mathbf{h}\in D_0}, \xi, \xi_0) \xrightarrow{\mathcal{L}} (M(\mathbf{h})_{\mathbf{h}\in D_0}, \xi, \xi_0)$ for any finite subset $D_0 \subset D$. By Skorokhod's representation, there exists $(M_n^*(\mathbf{h})_{\mathbf{h}\in D}, M^*(\mathbf{h})_{\mathbf{h}\in D}, \xi^*, \xi_0^*)$ that has the same finite-dimensional distributions as $(M_n(\cdot), M(\cdot), \xi, \xi_0)$, and $M_n^*(\cdot) \to M^*(\cdot)$ in finite dimensions almost surely. Note that in restriction to $\{\xi_0^* = 1\}$, $M^*(\cdot)$ is uniquely minimized at a random variable $\hat{\mathbf{h}}^*$ that has the same distribution as $\hat{\mathbf{h}}$. We can then use the pathwise argument in the proof of Lemma A of [7] to deduce that $\hat{\mathbf{h}}_n^* \to \hat{\mathbf{h}}^*$ almost surely in restriction to the set $\{\xi_0^* = 1\}$, where $\hat{\mathbf{h}}_n^*$ minimizes $M_n^*(\cdot)$. This further implies that $(\hat{\mathbf{h}}_n^*, \xi^*) \to (\hat{\mathbf{h}}^*, \xi^*)$ almost surely in restriction to $\{\xi_0^* = 1\}$. By a reverse use of Skorokhod's representation, we deduce $(\hat{\mathbf{h}}_n, \xi) \xrightarrow{\mathcal{L}} (\hat{\mathbf{h}}, \xi)$ in restriction to Ω_0 as wanted. $\square$

SA.2 Proof of Theorem 2

Denote $\tilde{M}_n(\mathbf{h}) \equiv \sum_{i\in\mathcal{J}_n} \rho\left(\tilde{\varsigma}_{n,i} - \mathbf{h}^\top \Delta_i^n \mathbf{Z}\right)$. Let $i_{n,p}$ be defined as in the proof of Theorem 1. By Proposition 1 of [9], $\mathbb{P}(\mathcal{J}_n = \mathcal{J}_n^*) \to 1$. Hence, w.p.a.1,

$$\tilde{M}_n(\mathbf{h}) = \sum_{p\in\mathcal{P}} \rho\left(\tilde{\varsigma}_{n,i_{n,p}} - \mathbf{h}^\top \Delta_{i_{n,p}}^n \mathbf{Z}\right).$$

Since the set $\mathcal{P}$ is finite almost surely, the probability of $\{|\mathcal{P}| > \bar{p}\}$ can be made arbitrarily small by setting the constant $\bar{p}$ sufficiently large. Therefore, in order to prove the asserted convergence in probability, we can restrict attention to the set $\{|\mathcal{P}| \le \bar{p}\}$.

Fix an arbitrary subsequence $\mathbb{N}_1 \subseteq \mathbb{N}$. By Proposition 9.3.2 in [5], $\hat{\Sigma}_{n,i_{n,p}-} \xrightarrow{\mathbb{P}} \Sigma_{\tau_p-}$ and $\hat{\Sigma}_{n,i_{n,p}+} \xrightarrow{\mathbb{P}} \Sigma_{\tau_p}$ for each $p \ge 1$. Theorem 1 implies that $\hat{\boldsymbol{\beta}}_n \xrightarrow{\mathbb{P}} \boldsymbol{\beta}^*$. Therefore, we can select a further subsequence $\mathbb{N}_2 \subseteq \mathbb{N}_1$ such that $((\hat{\Sigma}_{n,i_{n,p}-}, \hat{\Sigma}_{n,i_{n,p}+})_{1\le p\le\bar{p}}, \hat{\boldsymbol{\beta}}_n)$ converges almost surely to $((\Sigma_{\tau_p-}, \Sigma_{\tau_p})_{1\le p\le\bar{p}}, \boldsymbol{\beta}^*)$ along $\mathbb{N}_2$. By the construction of $\tilde{\varsigma}_{n,i}$, it is then easy to see that the $\mathcal{F}$-conditional law of $(\tilde{\varsigma}_{n,i_{n,p}})_{1\le p\le\bar{p}}$ converges (under any metric for weak convergence) almost surely to that of $(\varsigma_p)_{1\le p\le\bar{p}}$ along $\mathbb{N}_2$. Note that $\Delta_{i_{n,p}}^n \mathbf{Z} \to \Delta \mathbf{Z}_{\tau_p}$ pathwise for each $p \ge 1$. Therefore, by the continuous mapping theorem, we further deduce that, along $\mathbb{N}_2$, the $\mathcal{F}$-conditional law of $(\tilde{M}_n(\mathbf{h}))_{\mathbf{h}\in D}$ converge almost surely to that of $(M(\mathbf{h}))_{\mathbf{h}\in D}$ for any countable dense subset $D \subseteq \mathbb{R}^{d-1}$. By using a convexity argument as in Lemma A of [7], we deduce that, along $\mathbb{N}_2$, the $\mathcal{F}$-conditional law of $\tilde{\mathbf{h}}_n$ converges almost surely to that of $\hat{\mathbf{h}}$ in restriction to Ω_0. By the subsequence characterization of convergence in probability, we deduce the assertion of the theorem. □

SA.3 Proof of Theorem 3

Step 1. We outline the proof in this step. We shall use two technical results that are proved in steps 2 and 3. Below, we use $o_{pu}(1)$ to denote a uniformly $o_p(1)$ term.

We shall use an alternative representation for $M'(\cdot)$. Let $(T_m)_{m\geq1}$ be the successive jump times of the Poisson process $t \mapsto \mu([0,t]\times\mathbb{R})$. We consider $\mathbb{R}$-valued processes $(\tilde{\zeta}_m(\cdot), \tilde{\zeta}'_m(\cdot))_{m\geq1}$ which, conditional on $\mathcal{F}$, are mutually independent centered Gaussian processes with covariance functions given by

$$\begin{cases} \mathbb{E}\left[\tilde{\zeta}_m(s)\tilde{\zeta}_m(t)|\mathcal{F}\right] = \theta\Sigma_{T_m-}\int_{-1}^{0} g(s+u)\, g(t+u)\, du + \theta\Sigma_{T_m}\int_0^1 g(s+u)\, g(t+u)\, du, \\ \mathbb{E}\left[\tilde{\zeta}'_m(s)\tilde{\zeta}'_m(t)|\mathcal{F}\right] = \frac{A_{T_m-}}{\theta}\int_{-1}^{0} g'(s+u)\, g'(t+u)\, du + \frac{A_{T_m}}{\theta}\int_0^1 g'(s+u)\, g'(t+u)\, du. \end{cases}$$

We then observe that $M'(\cdot)$ can be represented as

$$M'(\mathbf{h}) = \sum_{m\geq1: T_m\leq T} 1_{\{\Delta\mathbf{Z}_{T_m}\neq0\}}\int_0^1 \rho\left(\tilde{\zeta}_m(s) + \tilde{\zeta}'_m(s) - \mathbf{h}^\top\Delta\mathbf{Z}_{T_m} g(s)\right) ds.$$

Note that the stopping times $(T_m)_{m\geq1}$ are independent of the Brownian motion $\mathbf{W}$. Hence, with $\mathcal{H}_t \equiv \mathcal{F}_t \vee \sigma(T_m : m\geq1)$, the process $\mathbf{W}$ is still a Brownian motion with respect to the filtration $(\mathcal{H}_t)_{t\geq0}$. We consider a sequence Ω_n of events on which the stopping times $(T_m)_{m\geq1}$ do not occur on the sampling grid $\{i\Delta_n : i\geq0\}$ and $|T_m - T_{m'}| > 3k_n\Delta_n$ whenever $m\neq m'$. Since the jumps have finite activity and $k_n\Delta_n\to0$, $\mathbb{P}(\Omega_n)\to1$. Therefore, we can restrict the calculation below to Ω_n without loss of generality. Below, we denote $I_{n,m} = \lfloor T_m/\Delta_n\rfloor$, which is an $\mathcal{H}_0$-measurable random integer.

Recall the definition of $\mathcal{J}_n'^*$ from (17). We complement the definition (19) with

$$M_n'^*(\mathbf{h}) = \frac{1}{k_n\Delta_n^{q/4}}\sum_{i\in\mathcal{J}_n'^*}\rho\left(\bar{Y}'_{n,i} - (\boldsymbol{\beta}^* + \Delta_n^{1/4}\mathbf{h})^\top\bar{\mathbf{Z}}'_{n,i}\right).$$

In step 2 below, we shall show that for each $\mathbf{h}$,

$$M'_n(\mathbf{h}) - M_n'^*(\mathbf{h}) = o_p(1). \tag{SA.1}$$

We then proceed to derive the finite-dimensional stable convergence in law of $M_n'^*(\cdot)$.

We denote $\mathbf{X}'^c = (Y'^c, \mathbf{Z}'^c) = \mathbf{X}^c + \boldsymbol{\chi}'$. Observe that

$$\begin{aligned}
M_n'^*(\mathbf{h}) &= \sum_{m\geq 1: T_m \leq T} 1_{\{\Delta \mathbf{Z}_{T_m} \neq 0\}} \frac{1}{k_n} \sum_{j=0}^{k_n - 1} \rho\left(\Delta_n^{-1/4}(\bar{Y}'^c_{n,I_{n,m}-j} - \boldsymbol{\beta}^{*\top} \bar{\mathbf{Z}}'^c_{n,I_{n,m}-j}) - \mathbf{h}^\top \bar{\mathbf{Z}}'_{n,I_{n,m}-j}\right) \\
&= \sum_{m\geq 1: T_m \leq T} 1_{\{\Delta \mathbf{Z}_{T_m} \neq 0\}} \int_0^1 \rho\left(\tilde{\zeta}_{n,m}(s) + \tilde{\zeta}'_{n,m}(s) - \mathbf{h}^\top \bar{\mathbf{Z}}'_{n,I_{n,m}-\lfloor k_n s \rfloor}\right) ds, \qquad \text{(SA.2)}
\end{aligned}$$

where we define, for $s \in [0, 1]$,

$$\begin{cases} \tilde{\zeta}_{n,m}(s) \equiv \Delta_n^{-1/4}(1, -\boldsymbol{\beta}^{*\top})\bar{\mathbf{X}}^c_{n,I_{n,m}-\lfloor k_n s\rfloor}, \\ \tilde{\zeta}'_{n,m}(s) \equiv \Delta_n^{-1/4}(1, -\boldsymbol{\beta}^{*\top})\bar{\boldsymbol{\chi}}'_{n,I_{n,m}-\lfloor k_n s\rfloor}. \end{cases} \qquad \text{(SA.3)}$$

In step 3 below, we show

$$(\tilde{\zeta}_{n,m}(\cdot), \tilde{\zeta}'_{n,m}(\cdot))_{m\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\tilde{\zeta}_m(\cdot), \tilde{\zeta}'_m(\cdot))_{m\geq 1} \qquad \text{(SA.4)}$$

under the product Skorokhod topology. Estimates used in step 3 also imply $\bar{\mathbf{Z}}'^c_{n,I_{n,m}-\lfloor k_n \cdot\rfloor} = o_{pu}(1)$. Hence, $\bar{\mathbf{Z}}'_{n,I_{n,m}-\lfloor k_n \cdot\rfloor} = g_n(\lfloor k_n \cdot\rfloor + 1)\Delta \mathbf{Z}_{T_m} + o_{pu}(1)$. Since the weight function $g(\cdot)$ is Lipschitz continuous, we further deduce that

$$\bar{\mathbf{Z}}'_{n,I_{n,m}-\lfloor k_n \cdot\rfloor} = g(\cdot)\Delta \mathbf{Z}_{T_m} + o_{pu}(1). \qquad \text{(SA.5)}$$

We now note that the limiting processes in (SA.4) and (SA.5) have continuous paths. By Propositions VI.1.17 and VI.1.23 in [6], as well as the continuous mapping theorem, we deduce $M_n'^*(\cdot) \xrightarrow{\mathcal{L}\text{-}s} M'(\cdot)$ on finite dimensions. The second assertion of Theorem 3 then follows from the convexity argument used in the proof of Theorem 1.

Step 2. We show (SA.1) in this step. Fix $\mathbf{h} \in \mathbb{R}^{d-1}$. We denote $\rho_{n,i} \equiv \rho(\bar{Y}'_{n,i} - (\boldsymbol{\beta}^* + \Delta_n^{1/4}\mathbf{h})^\top \bar{\mathbf{Z}}'_{n,i})$ and decompose $M_n'(\mathbf{h}) - M_n'^*(\mathbf{h}) = R_{1,n} + R_{2,n}$, where

$$R_{1,n} \equiv \frac{1}{k_n \Delta_n^{q/4}} \sum_{i \in \mathcal{J}'_n \setminus \mathcal{J}'^*_n} \rho_{n,i}, \quad R_{2,n} \equiv \frac{1}{k_n \Delta_n^{q/4}} \sum_{i \in \mathcal{J}'^*_n \setminus \mathcal{J}'_n} \rho_{n,i}.$$

It remains to show that both $R_{1,n}$ and $R_{2,n}$ are $o_p(1)$.

We first consider $R_{1,n}$. Note that for $i \notin \mathcal{J}_n'^*$, $\bar{\mathbf{Z}}_{n,i}' = \bar{\mathbf{Z}}_{n,i}'^c$. Set $\underline{u}_n' = \min_j u_{j,n}'$ and observe $\mathcal{J}_n' \subseteq \{i : \|\bar{\mathbf{Z}}_{n,i}'\| > \underline{u}_n'\}$. Hence,

$$R_{1,n} \leq \frac{1}{k_n \Delta_n^{q/4}} \sum_{i \in \mathcal{I}_n'} \rho_{n,i} 1_{\{\|\bar{\mathbf{Z}}_{n,i}'^c\| > \underline{u}_n'\}}.$$

Under Assumption 1, $\mathbb{E}\,|\rho_{n,i}|^2 \leq K\mathbb{E}\|\bar{\mathbf{X}}_{n,i}'\|^{2q}$, where the majorant side is bounded. By (5.39) of [4], for any $v > 0$,

$$\mathbb{E}[\|\bar{\mathbf{X}}_{n,i}'^c\|^v | \mathcal{H}_0] \leq K_v \Delta_n^{1/4}. \tag{SA.6}$$

Hence, $\mathbb{P}(\|\bar{\mathbf{Z}}_{n,i}'^c\| > \underline{u}_n') \leq K_v \Delta_n^{(1/4-\varpi')v}$ by Markov's inequality. Note that $k_n^{-1}\Delta_n^{-q/4} \leq K$. By the Cauchy–Schwarz inequality, we further deduce $\mathbb{E}\left[R_{1,n}\right] \leq K_v \Delta_n^{(1/4-\varpi')v/2-1}$ for any $v > 0$. Setting $v > 2/(1/4 - \varpi')$, we deduce $R_{1,n} = o_p(1)$.

Turning to $R_{2,n}$, we first observe that $\bar{Y}_{n,i}' - \boldsymbol{\beta}^{*\top}\bar{\mathbf{Z}}_{n,i}' = \bar{Y}_{n,i}'^c - \boldsymbol{\beta}^{*\top}\bar{\mathbf{Z}}_{n,i}'^c$ for all $i \in \mathcal{J}_n'^*$ in restriction to Ω_n. Therefore, we can rewrite $R_{2,n} = k_n^{-1}\sum_{i \in \mathcal{J}_n'^* \setminus \mathcal{J}_n'} \tilde{\rho}_{n,i}$, where $\tilde{\rho}_{n,i} \equiv \rho(\Delta_n^{-1/4}(\bar{Y}_{n,i}'^c - \boldsymbol{\beta}^{*\top}\bar{\mathbf{Z}}_{n,i}'^c) - \mathbf{h}^\top \bar{\mathbf{Z}}_{n,i}')$. For each $m \geq 1$, we consider the positive process $(f_{n,m}(s))_{s \in [0,1]}$ given by

$$f_{n,m}(s) = \tilde{\rho}_{n,I_{n,m}-\lfloor k_n s \rfloor} 1_{\left\{\|\bar{\mathbf{Z}}_{n,I_{n,m}-\lfloor k_n s \rfloor}'\| \leq \sum_{j=1}^{d-1} u_{j,n}'\right\} \cap \{\Delta \mathbf{Z}_{T_m} \neq 0\}}.$$

We then bound $R_{2,n}$ as follows

$$\begin{aligned} R_{2,n} &= \frac{1}{k_n} \sum_{m \geq 1 : T_m \leq T} \sum_{j=0}^{k_n - 1} \tilde{\rho}_{n,I_{n,m}-j} 1_{\{I_{n,m}-j \notin \mathcal{J}_n'\}} 1_{\{\Delta \mathbf{Z}_{T_m} \neq 0\}} \\ &\leq \sum_{m \geq 1 : T_m \leq T} \int_0^1 f_{n,m}(s) ds. \end{aligned} \tag{SA.7}$$

Recall that the random integer $I_{n,m}$ is $\mathcal{H}_0$-measurable. Hence,

$$\mathbb{E}\left[f_{n,m}(s)^2 | \mathcal{H}_0\right] \leq \mathbb{E}[\tilde{\rho}_{n,I_{n,m}-\lfloor k_n s \rfloor}^2 | \mathcal{H}_0] \leq K, \tag{SA.8}$$

where the second inequality follows from $|\tilde{\rho}_{n,i}|^2 \leq K + K\|\Delta_n^{-1/4}\bar{\mathbf{X}}'^c_{n,i}\|^{2q}$ and (SA.6).

We now claim that, for each $s \in (0,1)$,

$$\mathbb{P}\left(\left\{\left\|\bar{\mathbf{Z}}'_{n,I_{n,m}-\lfloor k_n s\rfloor}\right\| \leq \sum_{j=1}^{d-1} u'_{j,n}\right\} \cap \{\Delta \mathbf{Z}_{T_m} \neq 0\}\middle| \mathcal{H}_0\right) = o_p(1). \tag{SA.9}$$

To see this, we first note that $\bar{\mathbf{Z}}'_{n,I_{n,m}-\lfloor k_n s\rfloor} = (g(s) + O(k_n^{-1}))\Delta \mathbf{Z}_{T_m} + \bar{\mathbf{Z}}'^c_{n,I_{n,m}-\lfloor k_n s\rfloor}$. Since $g(s) > 0$ for $s \in (0,1)$ by our maintained assumption on the weight function, $\|(g(s) + O(k_n^{-1}))\Delta \mathbf{Z}_{T_m}\|$ is bounded below by $g(s)\|\Delta \mathbf{Z}_{T_m}\|/2 > 0$ for large n when $\Delta \mathbf{Z}_{T_m} \neq 0$. On the other hand, by the maximal inequality (see, e.g., Lemma 2.2.2 of [11]) and the L_v-bound given by (SA.6), we deduce $\max_{i\in\mathcal{I}'_n} \|\bar{\mathbf{Z}}'^c_{n,i}\| = O_p(\Delta_n^{1/4-\iota})$ under the $\mathcal{H}_0$-conditional probability for any fixed but arbitrarily small $\iota > 0$. From these estimates, the claim (SA.9) readily follows.

From (SA.8), we also see $\tilde{\rho}_{n,I_{n,m}-\lfloor k_n s\rfloor} = O_p(1)$ under the $\mathcal{H}_0$-conditional probability. Then, by (SA.9), $f_{n,m}(s) = o_p(1)$ for each $s \in (0,1)$. Note that (SA.8) implies that, for m and s fixed, the sequence $(f_{n,m}(s))_{n\geq 1}$ is uniformly integrable. Therefore, $\mathbb{E}\left[f_{n,m}(s)|\mathcal{H}_0\right] = o_p(1)$. By Fubini's theorem and the bounded convergence theorem, we further deduce for each $m \geq 1$,

$$\mathbb{E}\left[\int_0^1 f_{n,m}(s)ds\middle| \mathcal{H}_0\right] = o_p(1). \tag{SA.10}$$

Finally, note that the cadinality of $\{m : T_m \leq T\}$ is finite almost surely. It then follows from (SA.7) and (SA.10) that $R_{2,n} = o_p(1)$. The proof of (SA.1) is now complete.

Step 3. We show (SA.4) in this step. By Lemma A3 of [10], which is a functional extension of Proposition 5 of [3], it suffices to show that $(\tilde{\zeta}_{n,m}(\cdot))_{m\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\tilde{\zeta}_m(\cdot))_{m\geq 1}$ and $\mathcal{L}[(\tilde{\zeta}'_{n,m}(\cdot))_{m\geq 1}|\mathcal{F}] \xrightarrow{\mathbb{P}} \mathcal{L}[(\tilde{\zeta}'_m(\cdot))_{m\geq 1}|\mathcal{F}]$, where $\mathcal{L}[\,\cdot\,|\mathcal{F}]$ denotes the $\mathcal{F}$-conditional law and the latter convergence is under any metric for the weak convergence of probability measures.

We first show $(\tilde{\zeta}_{n,m}(\cdot))_{m\geq 1} \xrightarrow{\mathcal{L}\text{-}s} (\tilde{\zeta}_m(\cdot))_{m\geq 1}$. Recall the definitions in (SA.3). Since $g(\cdot)$

is supported on $[0,1]$, we can rewrite

$$\tilde{\zeta}_{n,m}\left(s\right)=\Delta_{n}^{-1/4}\sum_{i=-(k_n-1)}^{k_n-1}g_n\left(i+\lfloor k_n s\rfloor\right)\Delta_{I_{n,m}+i}^{n}U^{*c}, \tag{SA.11}$$

where U^{*c} is the continuous component of the process U^*, that is, $U_t^{*c}=(1,-\boldsymbol{\beta}^{*\top})\mathbf{X}_t^c$. We consider an approximation of $\tilde{\zeta}_{n,m}\left(s\right)$ given by

$$\begin{aligned}\tilde{\zeta}_{n,m}^{c}\left(s\right)=\Delta_{n}^{-1/4}&\sum_{i=-(k_n-1)}^{-1}g_n\left(i+k_n s\right)(1,-\boldsymbol{\beta}^{*\top})\boldsymbol{\sigma}_{T_m-k_n\Delta_n}\Delta_{I_{n,m}+i}^{n}\mathbf{W}\\&+\Delta_{n}^{-1/4}\sum_{i=1}^{k_n-1}g_n\left(i+k_n s\right)(1,-\boldsymbol{\beta}^{*\top})\boldsymbol{\sigma}_{T_m}\Delta_{I_{n,m}+i}^{n}\mathbf{W}.\end{aligned} \tag{SA.12}$$

We now show that

$$\tilde{\zeta}_{n,m}\left(\cdot\right)-\tilde{\zeta}_{n,m}^{c}\left(\cdot\right)=o_{pu}(1). \tag{SA.13}$$

To see this, we first note that $\Delta_n^{-1/4}g(\lfloor k_n s\rfloor/k_n)\Delta_{I_{n,m}}^n\mathbf{X}^c=O_p(\Delta_n^{1/4})$ uniformly in s, so the summand in (SA.11) with $i=0$ is $o_{pu}(1)$. Since $g(\cdot)$ is Lipschitz continuous, $|g_n\left(i+\lfloor k_n s\rfloor\right)-g_n\left(i+k_n s\right)|\leq Kk_n^{-1}$ uniformly in s. Since $\mathbb{E}\|\Delta_{I_{n,m}+i}^n\mathbf{X}^c\|\leq K\Delta_n^{1/2}$, the difference resulted from replacing $\lfloor k_n s\rfloor$ with $k_n s$ in (SA.11) is $O_p(\Delta_n^{1/4})$ uniformly in s. Hence,

$$\tilde{\zeta}_{n,m}\left(\cdot\right)-\tilde{\zeta}_{n,m}^{c}\left(\cdot\right)=R_{n,1}(\cdot)+R_{n,2}(\cdot)+o_{pu}(1) \tag{SA.14}$$

where

$$\begin{aligned}R_{n,1}(s)&\equiv\Delta_{n}^{-1/4}\sum_{i=-(k_n-1)}^{-1}g_n\left(i+k_n s\right)\left(\Delta_{I_{n,m}+i}^{n}U^{*c}-(1,-\boldsymbol{\beta}^{*\top})\boldsymbol{\sigma}_{T_m-k_n\Delta_n}\Delta_{I_{n,m}+i}^{n}\mathbf{W}\right),\\R_{n,2}(s)&\equiv\Delta_{n}^{-1/4}\sum_{i=1}^{k_n-1}g_n\left(i+k_n s\right)\left(\Delta_{I_{n,m}+i}^{n}U^{*c}-(1,-\boldsymbol{\beta}^{*\top})\boldsymbol{\sigma}_{T_m}\Delta_{I_{n,m}+i}^{n}\mathbf{W}\right).\end{aligned}$$

For each $s \in [0,1]$,

$$\begin{aligned}
&\mathbb{E}\left[\left|R_{n,1}(s)\right|^2 \middle| \mathcal{H}_0\right] \\
&\quad \leq K\Delta_n^{-1/2} \sum_{i=-(k_n-1)}^{-1} g_n\left(i+k_n s\right)^2 \\
&\qquad \times \mathbb{E}\left[\left|\Delta_{I_{n,m}+i}^n U^{*c} - (1,-\boldsymbol{\beta}^{*\top})\boldsymbol{\sigma}_{T_m-k_n\Delta_n}\Delta_{I_{n,m}+i}^n \mathbf{W}\right|^2 \middle| \mathcal{H}_0\right] \\
&\quad \leq K\Delta_n^{-1/2} \sum_{i=-(k_n-1)}^{-1} \left(\Delta_n^2 + \int_{(I_{n,m}+i-1)\Delta_n}^{(I_{n,m}+i)\Delta_n} \mathbb{E}\left[\left|\Sigma_s^{1/2} - \Sigma_{T_m-k_n\Delta_n}^{1/2}\right|^2 \middle| \mathcal{H}_0\right] ds\right) \\
&\quad = o_p(1),
\end{aligned}$$

where the first inequality is derived using the fact that $R_{n,1}(s)$ is formed as a sum of martingale differences; the second inequality follows from the boundedness of the drift and Itô's isometry; the last line follows from $k_n \asymp \Delta_n^{-1/2}$ and the fact that the process Σ is càdlàg. Therefore, $R_{n,1}(s) = o_p(1)$ for fixed s. We further verify that $R_{n,1}(\cdot)$ is tight. To this end, we note that for $s,t \in [0,1]$

$$\begin{aligned}
&\mathbb{E}\left[\left|R_{n,1}(s) - R_{n,2}(t)\right|^2 \middle| \mathcal{H}_0\right] \\
&\quad \leq K\Delta_n^{-1/2} \sum_{i=-(k_n-1)}^{-1} \left(g_n\left(i+k_n s\right) - g_n\left(i+k_n t\right)\right)^2 \\
&\qquad \times \mathbb{E}\left[\left|\Delta_{I_{n,m}+i}^n U^{*c} - (1,-\boldsymbol{\beta}^{*\top})\boldsymbol{\sigma}_{T_m-k_n\Delta_n}\Delta_{I_{n,m}+i}^n \mathbf{W}\right|^2 \middle| \mathcal{H}_0\right] \\
&\quad \leq K\left|s-t\right|^2.
\end{aligned} \tag{SA.15}$$

From here, the tightness of $R_{n,1}(\cdot)$ readily follows. Hence, $R_{n,1}(\cdot) = o_{pu}(1)$. Similarly, we can show $R_{n,2}(\cdot) = o_{pu}(1)$. Recalling (SA.14), we deduce (SA.13) as claimed.

We note that $\tilde{\zeta}_{n,m}^c(\cdot)$ is continuous and, for each s, $\tilde{\zeta}_{n,m}^c(s)$ is formed as a sum of martingale differences. We derive the finite-dimensional convergence of $\tilde{\zeta}_{n,m}^c(\cdot)$ towards $\tilde{\zeta}_m(\cdot)$ by using the central limit theorem given by Theorem IX.7.28 in [6]. In particular, it is easy to verify using a Riemann approximation that the asymptotic covariance between

$\tilde{\zeta}^{c}_{n,m}(s)$ and $\tilde{\zeta}^{c}_{n,m'}(t)$ is

$$\left(\theta\Sigma_{T_m-}\int_{-1}^{0} g(s+u)\,g(t+u)\,du+\theta\Sigma_{T_m}\int_{0}^{1} g(s+u)\,g(t+u)\,du\right)1_{\{m=m'\}}.$$

Since $g(\cdot)$ is Lipschitz continuous, we deduce $\mathbb{E}|\tilde{\zeta}^{c}_{n,m}(s)-\tilde{\zeta}^{c}_{n,m}(t)|^2\le K|s-t|^2$ for $s,t\in[0,1]$ using estimates similar to (SA.15). Therefore, the continuous processes $\tilde{\zeta}^{c}_{n,m}(\cdot)$ form a tight sequence. From here, we deduce $(\tilde{\zeta}^{c}_{n,m}(\cdot))_{m\ge1}\xrightarrow{\mathcal{L}\text{-}s}(\tilde{\zeta}_{m}(\cdot))_{m\ge1}$ under the product topology induced by the uniform metric. By (SA.13) and Corollary VI.3.33 of [6], we deduce $(\tilde{\zeta}_{n,m}(\cdot))_{m\ge1}\xrightarrow{\mathcal{L}\text{-}s}(\tilde{\zeta}_{m}(\cdot))_{m\ge1}$ as wanted.

Next, we show $\mathcal{L}[(\tilde{\zeta}'_{n,m}(\cdot))_{m\ge1}|\mathcal{F}]\xrightarrow{\mathbb{P}}\mathcal{L}[(\tilde{\zeta}'_{m}(\cdot))_{m\ge1}|\mathcal{F}]$. Recall (SA.3) and $g'_n(j)\equiv g_n(j)-g_n(j-1)$. To simplify notations, we denote $\tilde{\chi}_t=\left(1,-\boldsymbol{\beta}^{*\top}\right)\boldsymbol{\chi}'_t$ and note that $\mathbb{E}[\tilde{\chi}_t^2|\mathcal{F}]=A_t$. It is elementary to rewrite $\tilde{\zeta}'_{n,m}(s)$ as

$$\begin{aligned}\tilde{\zeta}'_{n,m}(s) &= -\Delta_n^{-1/4}\sum_{j=1}^{k_n} g'_n(j)\,\tilde{\chi}_{(I_{n,m}-\lfloor k_n s\rfloor+j-1)\Delta_n}\\ &= -\Delta_n^{-1/4}\sum_{i=-k_n}^{k_n-1} g'_n(i+1+\lfloor k_n s\rfloor)\,\tilde{\chi}_{(I_{n,m}+i)\Delta_n}.\end{aligned}$$

We approximate $\tilde{\zeta}'_{n,m}(\cdot)$ with the continuous process $\tilde{\zeta}'^{c}_{n,m}(\cdot)$ given by

$$\tilde{\zeta}'^{c}_{n,m}(s)=-\Delta_n^{-1/4}\sum_{i=-k_n}^{-1} g'_n(i+1+k_n s)\,\tilde{\chi}_{(I_{n,m}+i)\Delta_n}-\Delta_n^{-1/4}\sum_{i=1}^{k_n-1} g'_n(i+1+k_n s)\,\tilde{\chi}_{(I_{n,m}+i)\Delta_n}.$$

Since $g(\cdot)$ and $g'(\cdot)$ are Lipschitz continuous, we have

$$\sup_{s\in[0,1]}|g'_n(\lfloor k_n s\rfloor+1)|\le Kk_n^{-1},\quad \sup_{s\in[0,1]}|g'_n(i+k_n s)-g'_n(i+\lfloor k_n s\rfloor)|\le Kk_n^{-2}.$$

From these estimates, we deduce

$$\sup_{s\in[0,1]}\left|\tilde{\zeta}'_{n,m}(s)-\tilde{\zeta}'^{c}_{n,m}(s)\right|=O_p(\Delta_n^{-1/4}k_n^{-1})=o_p(1). \tag{SA.16}$$

By applying a central limit theorem under the $\mathcal{F}$-conditional probability, we derive the finite-dimensional convergence of $\tilde{\zeta}_{n,m}^{\prime c}(\cdot)$ towards $\tilde{\zeta}_{m}^{\prime}(\cdot)$. In particular, it is easy to verify using a Riemann approximation that the asymptotic covariance between $\tilde{\zeta}_{n,m}^{\prime c}(s)$ and $\tilde{\zeta}_{n,m'}^{\prime c}(t)$ is given by

$$\left(\frac{A_{T_m-}}{\theta}\int_{-1}^{0} g'(s+u)\,g'(t+u)\,du + \frac{A_{T_m}}{\theta}\int_{0}^{1} g'(s+u)\,g'(t+u)\,du\right)1_{\{m=m'\}}.$$

We now verify that the processes $\tilde{\zeta}_{n,m}^{\prime c}(\cdot)$ form a tight sequence. Note that, for $s,t\in[0,1]$, $|g_n'(i+k_n s)-g_n'(i+k_n t)|\leq Kk_n^{-1}|t-s|$. Since the variables $\left(\tilde{\chi}_{(I_{n,m}+i)\Delta_n}\right)$ are $\mathcal{F}$-conditionally independent with bounded second moments, we further derive that

$$\mathbb{E}\left|\tilde{\zeta}_{n,m}^{\prime c}(s)-\tilde{\zeta}_{n,m}^{\prime c}(t)\right|^2 \leq K\Delta_n^{-1/2}k_n^{-1}|t-s|^2 \leq K|t-s|^2.$$

From here, the tightness of $\tilde{\zeta}_{n,m}^{\prime c}(\cdot)$ follows. Therefore, $\mathcal{L}[(\tilde{\zeta}_{n,m}^{\prime c}(\cdot))_{m\geq 1}|\mathcal{F}] \xrightarrow{\mathbb{P}} \mathcal{L}[(\tilde{\zeta}_{m}^{\prime}(\cdot))_{m\geq 1}|\mathcal{F}]$. By (SA.16), the conditional law of $(\tilde{\zeta}_{n,m}^{\prime}(\cdot))_{m\geq 1}$ converges to the same limit as claimed. This finishes the proof of (SA.4). □

SA.4 Proof of Theorem 4

Similarly as in the proof of Theorem 3, we can restrict the calculation below to the event Ω_n without loss of generality. We first establish some facts about the clusters used in Algorithm 1. Firstly, by a maximal inequality and (SA.6), $\sup_{i\notin\mathcal{J}_n^{\prime *}}\|\bar{\mathbf{Z}}_{n,i}^{\prime}\| = \sup_{i\notin\mathcal{J}_n^{\prime *}}\|\bar{\mathbf{Z}}_{n,i}^{\prime c}\| = O_p(\Delta_n^{1/4-\iota})$, where the constant $\iota>0$ can be taken to be less than ϖ'. We then deduce that, w.p.a.1., the indices outside $\mathcal{J}_n^{\prime *}$ are not selected by $\mathcal{J}_n^{\prime}$. Secondly, we observe that (SA.9) can be strengthened to

$$\mathbb{P}\left(\left\{\inf_{s\in[\varepsilon,1-\varepsilon]}\left\|\bar{\mathbf{Z}}_{n,I_{n,m}-\lfloor k_n s\rfloor}^{\prime}\right\| \leq \sum_{j=1}^{d-1}u_{j,n}^{\prime}\right\}\cap\{\Delta\mathbf{Z}_{T_m}\neq 0\}\right)\to 0,$$

for any fixed, but arbitrarily small $\varepsilon \in (0, 1/4)$. Therefore, the following holds w.p.a.1: each jump time of $\mathbf{Z}$ is matched with at least one index $i \in \mathcal{J}_n'^*$ such that $(i\Delta_n, (i + k_n)\Delta_n]$ contains this jump time, and the differences between these indices are bounded above by $k_n/4$. From these two facts, we deduce that, w.p.a.1., there is a one-to-one correspondence between each jump time τ_p of $\mathbf{Z}$ and each cluster $\mathcal{J}'_{n,p}$ such that $\tau_p \in (i\Delta_n, (i + k_n)\Delta_n]$ for each $i \in \mathcal{J}'_{n,p}$. In particular, $(\min \mathcal{J}'_{n,p} - k_n' - k_n)\Delta_n$ and $(\max \mathcal{J}'_{n,p} + k_n - 1)\Delta_n$ converge to τ_p from left and right, respectively. Then, by (B.7) of [2], we have

$$\begin{cases} \hat{\Sigma}'_{n,\min \mathcal{J}'_{n,p} - k_n' - k_n} \xrightarrow{\mathbb{P}} \Sigma_{\tau_p -}, & \hat{\Sigma}'_{n,\max \mathcal{J}'_{n,p} + k_n - 1} \xrightarrow{\mathbb{P}} \Sigma_{\tau_p}, \\ \hat{A}_{n,\min \mathcal{J}'_{n,p} - k_n' - k_n} \xrightarrow{\mathbb{P}} A_{\tau_p -}, & \hat{A}_{n,\max \mathcal{J}'_{n,p} + k_n - 1} \xrightarrow{\mathbb{P}} A_{\tau_p}. \end{cases} \tag{SA.17}$$

By an argument similar to that in step 2 of the proof of Theorem 3, we can show that $k_n^{-1} \sum'_{i \in \mathcal{J}'_{n,p}} \bar{\mathbf{Z}}'_{n,i} = k_n^{-1} \sum_{i \in \mathcal{J}'^*_{n,p}} \bar{\mathbf{Z}}'_{n,i} + o_p(1)$, where $\mathcal{J}'^*_{n,p} = \{i : \tau_p \in (i\Delta_n, i\Delta_n + k_n\Delta_n]\}$. By (SA.5), we further deduce that $k_n^{-1} \sum'_{i \in \mathcal{J}'_{n,p}} \bar{\mathbf{Z}}'_{n,i} \xrightarrow{\mathbb{P}} (\int_0^1 g(u)\, du) \Delta \mathbf{Z}_{\tau_p}$. Similarly, we have $\left|\mathcal{J}'_{n,p}\right| / k_n \xrightarrow{\mathbb{P}} 1$ and, hence,

$$\frac{1}{k_n} \sum_{j = \lfloor (k_n - |\mathcal{J}'_{n,p}|)/2 \rfloor}^{\lfloor (k_n - |\mathcal{J}'_{n,p}|)/2 \rfloor + |\mathcal{J}'_{n,p}| - 1} g(j/k_n) \xrightarrow{\mathbb{P}} \int_0^1 g(u)\, du.$$

It follows that $\Delta \hat{\mathbf{Z}}_{n,p} \xrightarrow{\mathbb{P}} \Delta \mathbf{Z}_{\tau_p}$.

We now show that the $\mathcal{F}$-conditional law of $\tilde{M}'_n(\cdot)$ converges on finite dimensions in probability to that of $M'(\cdot)$ under the topology for the weak convergence of probability measures. By a subsequence argument as in the proof of Theorem 2, it suffices to prove the convergence under the $\mathcal{F}$-conditional probability for a given path on which (SA.17) holds pathwise, $(\Delta\hat{\mathbf{Z}}_{n,p})_{p\in\mathcal{P}} \to (\Delta\mathbf{Z}_{\tau_p})_{p\in\mathcal{P}}$ and $\mathcal{P}_n = \mathcal{P}$. Then, we can rewrite

$$\begin{aligned}\tilde{M}_n'(\mathbf{h}) &= \frac{1}{k_n}\sum_{p\in\mathcal{P}}\sum_{i=0}^{k_n-1}\rho\left(\left(\Delta_n^{-1/4}\sum_{j=1}^{k_n-1}g_n(j)\,\tilde{r}_{n,p,j-i}'\right)-g_n(i)\,\mathbf{h}^\top\Delta\hat{\mathbf{Z}}_{n,p}\right)\\ &= \sum_{p\in\mathcal{P}}\int_0^1\rho\left(\left(\Delta_n^{-1/4}\sum_{j=1}^{k_n-1}g_n(j)\,\tilde{r}_{n,p,j-\lfloor k_n s\rfloor}'\right)-g_n(\lfloor k_n s\rfloor)\,\mathbf{h}^\top\Delta\hat{\mathbf{Z}}_{n,p}\right)ds.\end{aligned}$$

Observe that

$$\begin{aligned}&\Delta_n^{-1/4}\sum_{j=1}^{k_n-1}g_n(j)\,\tilde{r}_{n,p,j-\lfloor k_n s\rfloor}'\\ &= \Delta_n^{-1/4}\sum_{j=1}^{k_n-1}g_n(j)\,\tilde{r}_{n,p,j-\lfloor k_n s\rfloor}-\Delta_n^{-1/4}\sum_{j=1}^{k_n}g_n'(j)\,\tilde{\chi}_{n,p,j-1-\lfloor k_n s\rfloor}'\\ &= \Delta_n^{-1/4}\sum_{i=-(k_n-1)}^{k_n-1}g_n(i+\lfloor k_n s\rfloor)\,\tilde{r}_{n,p,i}-\Delta_n^{-1/4}\sum_{i=-k_n}^{k_n-1}g_n'(i+1+\lfloor k_n s\rfloor)\,\tilde{\chi}_{n,p,i}'.\end{aligned}$$

The two terms on the right-hand side of the last equality are $\mathcal{F}$-conditionally independent by construction. Then, similarly as in step 3 of the proof of Theorem 3, we can derive the $\mathcal{F}$-conditional convergence in law of $(\Delta_n^{-1/4}\sum_{j=1}^{k_n-1}g_n(j)\,\tilde{r}_{n,p,j-\lfloor k_n\cdot\rfloor}')_{p\in\mathcal{P}}$ towards $(\varsigma_p(\cdot))_{p\in\mathcal{P}}$. By the continuous mapping theorem, we further deduce the finite-dimensional $\mathcal{F}$-conditional convergence in law of $\tilde{M}_n'(\cdot)$ towards $M'(\cdot)$. By a convexity argument used in the proof of Theorem 1, we deduce that $\tilde{\mathbf{h}}_n'$ converges in $\mathcal{F}$-conditional law to $\hat{\mathbf{h}}'$ as asserted. $\square$

Supplemental Appendix B: The effect of uncorrected noise: an example

In this appendix, we provide a concrete analytical example to illustrate the adverse effect of measurement errors (i.e., noise) on the jump regression procedure discussed in Section 2 of the main text. Since this procedure ignores the measurement errors when they are present, we shall refer to it as the "naive" procedure below. We consider the setting with $d = 2$, so $\mathbf{Z}$ is scalar-valued and will be denoted by Z. We shall impose some simplifying assumptions so as to make the argument in this appendix self-contained with minimal technicality. Below, we show that the naive estimator is biased *even* under these (favorable) simplifying assumptions.

In order to highlight the effect of the measurement errors on the naive procedure, we shut down the other complexities in our setting by assuming: (i) there is no drift (i.e., $\mathbf{b} = 0$); (ii) the volatility process $\boldsymbol{\sigma}$ is constant and nonrandom; (iii) the Brownian motion $\mathbf{W}$ is independent of the jump process; (iv) the error terms $(\boldsymbol{\chi}'_{i\Delta_n})_{i\geq}$ are independent of other processes; (v) Y does not have idiosyncratic jumps.

In addition, we recall that the naive procedure involves two steps. The first step is the jump detection and the second step is the regression using detected jumps. The adverse effect of noise on jump detection has been demonstrated in prior work (see, e.g., [1], [2], [8]). Therefore, here, we shall focus on the effect of noise in the regression step by considering the (favorable) infeasible setting in which the jump times are observed up to the precision of the sampling interval.

We now turn to the analysis of the naive estimator. For each jump time τ, let $i(\tau)$ denote the unique integer i such that $\tau \in ((i-1)\Delta_n, i\Delta_n]$. The noisy return vector over this sampling interval is given by

$$\begin{pmatrix} \Delta^n_{i(\tau)} Y' \\ \Delta^n_{i(\tau)} Z' \end{pmatrix} = \underbrace{\begin{pmatrix} \beta^* \Delta Z_\tau \\ \Delta Z_\tau \end{pmatrix}}_{\text{Jump part}} + \underbrace{\boldsymbol{\sigma} \Delta^n_{i(\tau)} \mathbf{W}}_{\text{Diffusive part}} + \underbrace{\Delta^n_{i(\tau)} \boldsymbol{\chi}'}_{\text{Noise part}}. \tag{SB.18}$$

For notational simplicity, we denote

$$D_\tau \equiv \begin{pmatrix} D_{Y,\tau} \\ D_{Z,\tau} \end{pmatrix} \equiv \frac{\boldsymbol{\sigma}\Delta^n_{i(\tau)}\mathbf{W}}{\Delta_n^{1/2}}, \quad N_\tau \equiv \begin{pmatrix} N_{Y,\tau} \\ N_{Z,\tau} \end{pmatrix} \equiv \Delta^n_{i(\tau)}\boldsymbol{\chi}',$$

so that we can rewrite (SB.18) as

$$\begin{pmatrix} \Delta^n_{i(\tau)}Y' \\ \Delta^n_{i(\tau)}Z' \end{pmatrix} = \begin{pmatrix} \beta^*\Delta Z_\tau \\ \Delta Z_\tau \end{pmatrix} + \Delta_n^{1/2}\begin{pmatrix} D_{Y,\tau} \\ D_{Z,\tau} \end{pmatrix} + \begin{pmatrix} N_{Y,\tau} \\ N_{Z,\tau} \end{pmatrix}. \tag{SB.19}$$

We note that the terms on the right-hand side of (SB.19) are mutually independent and independent across different τ's. In addition, D_τ and N_τ are non-degenerate when $\Delta_n \to 0$.

We consider the least-squares estimator, which admits a closed-form solution given by

$$\hat{\beta}_n^{LS} \equiv \frac{\sum_{\tau\in\mathcal{T}} \Delta^n_{i(\tau)}Y'\Delta^n_{i(\tau)}Z'}{\sum_{\tau\in\mathcal{T}}\left(\Delta^n_{i(\tau)}Z'\right)^2}.$$

Using (SB.19) and then extracting the leading term, we deduce

$$\begin{aligned} \hat{\beta}_n^{LS} - \beta^* &= \frac{\sum_{\tau\in\mathcal{T}}(\Delta_n^{1/2}(D_{Y,\tau} - \beta^* D_{Z,\tau}) + N_{Y,\tau} - \beta^* N_{Z,\tau})(\Delta Z_\tau + \Delta_n^{1/2}D_{Z,\tau} + N_{Z,\tau})}{\sum_{\tau\in\mathcal{T}}\left(\Delta Z_\tau + \Delta_n^{1/2}D_{Z,\tau} + N_{Z,\tau}\right)^2} \\ &= \underbrace{\frac{\sum_{\tau\in\mathcal{T}}(N_{Y,\tau} - \beta^* N_{Z,\tau})(\Delta Z_\tau + N_{Z,\tau})}{\sum_{\tau\in\mathcal{T}}(\Delta Z_\tau + N_{Z,\tau})^2}}_{\text{Bias}} + O_p(\Delta_n^{1/2}). \end{aligned} \tag{SB.20}$$

Clearly, the bias term would be zero if there were no measurement errors (i.e., $N_{Y,\tau} = N_{Z,\tau} = 0$), which is the case considered in Section 2.

We observe that the bias depends on a *fixed* number of measurement errors around the jump times, which, importantly, do not "average out" asymptotically. Hence, the bias is non-degenerate even in the limit.

Finally, we provide some further analytical insight by considering the case when there is only one jump realized in the sample, that is, $\mathcal{T} = \{\tau\}$. We can then rewrite the bias term in (SB.20) as

$$\begin{aligned} \text{Bias} &= \frac{N_{Y,\tau} - \beta^* N_{Z,\tau}}{\Delta Z_\tau + N_{Z,\tau}} \\ &= \left(\frac{\Delta Z_\tau}{N_{Z,\tau}} + 1\right)^{-1} \times \left(\frac{N_{Y,\tau}}{N_{Z,\tau}} - \beta^*\right). \end{aligned} \tag{SB.21}$$

The idea behind the decomposition (SB.21) is as follows. The scaling factor $\left(\frac{\Delta Z_\tau}{N_{Z,\tau}} + 1\right)^{-1}$ is strictly decreasing in the signal-to-noise ratio $\frac{\Delta Z_\tau}{N_{Z,\tau}}$. Consistently with intuition, the bias is larger when the signal-to-noise ratio is low. In the extreme case with the signal-to-noise ratio being zero, this scaling factor is 1 and the bias becomes $\frac{N_{Y,\tau}}{N_{Z,\tau}} - \beta^*$. In this sense, the second term in (SB.21) can be understood as the bias in the worst-case scenario.

References

[1] Yacine Aït-Sahalia and Jean Jacod. Testing for jumps in a discretely observed process. *Annals of Statistics*, 37:184–222, 2009.

[2] Yacine Aït-Sahalia, Jean Jacod, and Jia Li. Testing for jumps in noisy high frequency data. *Journal of Econometrics*, 168:207–222, 2012.

[3] Ole E. Barndorff-Nielsen, Peter Reinhard Hansen, Asger Lunde, and Neil Shephard. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76:1481–1536, 2008.

[4] Jean Jacod, Mark Podolskij, and Mathias Vetter. Limit theorems for moving averages of discretized processes plus noise. *Annals of Statistics*, 38:1478–1545, 2010.

[5] Jean Jacod and Philip Protter. *Discretization of Processes*. Springer, 2012.

[6] Jean Jacod and A. N. Shiryaev. *Limit Theorems for Stochastic Processes*. Springer-Verlag, New York, second edition, 2003.

[7] Keith Knight. Limit theory for autoregressive-parameter estimates in an infinite-variance random walk. *The Canadian Journal of Statistics / La Revue Canadienne de Statistique*, 17(3):pp. 261–278, 1989.

[8] Suzanne S. Lee and Per A. Mykland. Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics*, 168:396–406, 2012.

[9] Jia Li, Viktor Todorov, and George Tauchen. Jump regressions. Technical report, Duke Univeristy, 2014.

[10] Jia Li and Dacheng Xiu. Generalized methods of integrated moments for high-frequency data. Technical report, Duke University and University of Chicago, 2015.

[11] A. W. van der Vaart and J. Wellner. *Weak Convergence and Empirical Processes*. Springer-Verlag, 1996.