Supplement to
"Testing for Self-Excitation in Jumps"

H. Peter Boswijk[*]
Amsterdam School of Economics
University of Amsterdam
and Tinbergen Institute

Roger J. A. Laeven[†]
Amsterdam School of Economics
University of Amsterdam, EURANDOM
and CentER

Xiye Yang[‡]
Department of Economics
Rutgers University

November 3, 2017

This text serves as an appendix to the paper "Testing for Self-Excitation in Jumps." For context, notation and definitions, see the paper. In the first section we provide some examples and additional results of our Monte Carlo experiments. The second section describes some empirical results. In the third section we provide some preliminary technical results. Finally, in the fourth section we provide the proofs of our limit theorems.

[*]University of Amsterdam and Tinbergen Institute, PO Box 15867, Amsterdam, 1001 NJ, The Netherlands. Email: H.P.Boswijk@uva.nl, Phone: +31 (0)20 5254316.

[†]University of Amsterdam, EURANDOM and CentER, PO Box 15867, Amsterdam, 1001 NJ, The Netherlands. Email: R.J.A.Laeven@uva.nl, Phone: +31 (0)20 5254219.

[‡]Corresponding author. Rutgers University, 75 Hamilton Street, New Jersey Hall, New Brunswick, NJ 08901, USA. Email: xiyeyang@econ.rutgers.edu, Phone: +1 848 932 8655.

A Miscellaneous

A.1 Examples

We begin this section with an example of self-excitation in jumps. In this example, only those jumps with size in $\mathscr{E} = (-\infty, -\epsilon) \cup (\epsilon, \infty)$ can excite the jump intensity process.

Example A.1. *Let ξ and ϵ be two positive constants and consider*

$$\delta' = \delta'(t, x) = \xi \cdot 1_{\{|x| > \epsilon\}}.$$

Then, for any $t \in (0, T]$,

$$\sum_{0 \le t \le T} \Delta\lambda_t(\omega)\, 1_{\{\Delta\lambda_t(\omega) \ge 0,\, \Delta X_t(\omega) \ne 0\}} = \sum_{0 \le t \le T} \xi \cdot 1_{\{|\Delta X_t| > \epsilon\}}.$$

Therefore, in this example, an outcome $\omega \in \Omega$ exhibits self-excitation in jumps on $(0, T]$ if there is a log-price jump with magnitude larger than ϵ on $(0, T]$.

This example demonstrates clearly that we define the concept of self-excitation in a path-wise fashion. Note that δ' above is a nonnegative function of x and yields (strictly) positive values for $|x| > \epsilon$. Thus, this specification allows to generate paths with self-excitation in jumps; but it does not guarantee that, with probability one, an outcome $\omega \in \Omega$ has such property. In general, $\mathbb{P}(\omega : \forall\, t \in (0, T], |\Delta X_t(\omega)| \le \epsilon) > 0$. In this example, only jumps with absolute sizes larger than some threshold trigger the self-excitation effect. In fact, this is quite a plausible assumption since one would expect that large log-price jumps, especially negative ones, are more likely to provoke chain reactions than small ones.

Below are some examples of auxiliary functions g that satisfy condition (b) in Assumption 3 and that are employed in the paper. These examples are based on Jing et al. (2012), where a similar auxiliary function is used for the purpose of *integrated* (rather than *spot*) jump intensity estimation.

Example A.2. *Let $p > 2$ be an even integer. An example of a function g satisfying condition (b) in Assumption 3 is given by*

$$g_{1,p}(x) = \begin{cases} |x|^p, & |x| \le 1, \\ 1, & |x| > 1. \end{cases}$$

Note that with this choice $g_{1,p}(x) \to g_0(x) := 1_{\{x > 1\}}$ as $p \to \infty$.

Another example is

$$g_{2,p}(x) = \begin{cases} c^{-1}|x|^p & |x| \le a, \\ c^{-1}\left(a^p + \dfrac{pa^{p-1}}{2(b-a)}\big((b-a)^2 - (|x|-b)^2\big)\right), & a \le |x| \le b, \\ 1, & |x| > b, \end{cases}$$

where $0 < a < b < \infty$ *are two constants, and* $c = a^p + pa^{p-1}(b-a)/2$*. Note that this is a smoother version of the previous example and equivalence arises if* $a = b = 1$*.*

Below are some examples of the H function used for testing. First, consider the following smooth function:

$$H(x_1, x_2, y_1, y_2) = |x_2 - x_1|^p \, h_2(y_1, y_2),$$

with

$$h_2(y_1, y_2) = \begin{cases} \exp\big(-1/(y_2 - y_1)\big) & \text{if } y_2 > y_1; \\ 0 & \text{if } y_2 \le y_1. \end{cases}$$

Here and in the remainder of this section, $p \ge 2$. This function satisfies $H \in C^n$ in (y_1, y_2) for any integer $n > 2$ and any derivative w.r.t. y_1 and y_2 is zero on the area $\{y_1 \ge y_2\}$. In sum, it would be very hard, if not impossible, to derive a CLT with non-degenerate limiting processes when condition (3.3) in the paper is true. Therefore, we will avoid using a function H satisfying this condition when it comes to testing.

To give a different example of a function H satisfying the degeneracy condition (3.12) and Assumption 5, consider the following function:

$$\begin{aligned} H(x_1, x_2, y_1, y_2) = {} & |x_2 - x_1|^p \\ & \times \left(2 \cdot \log\left(\frac{y_1 + y_2}{2}\right) - \log(y_1) - \log(y_2)\right). \end{aligned} \tag{A.1}$$

Additionally, it also satisfies (3.2).

An example of a function H for which the degeneracy condition fails is simply

$$H(x_1, x_2, y_1, y_2) = |x_2 - x_1|^p (y_2 - y_1). \tag{A.2}$$

It satisfies Assumption 4. Additionally, it also satisfies (3.4).

A.2 Additional simulation results

Finally, we provide some additional details and results on the Monte Carlo study. We choose the value of λ_∞ in such a way that the following tail probability is 0.25%:

$$\mathbb{P}(|\lambda_\infty \Delta_i^n Y| \ge \alpha \Delta_n^{\varpi}) = \frac{2 c_\beta \lambda_\infty^\beta \Delta_n}{\beta(\alpha \Delta_n^{\varpi})^\beta} + \text{smaller order term},$$

where

$$c_\beta = \frac{\Gamma(\beta + 1)}{2\pi} \sin\left(\frac{\pi \beta}{2}\right).$$

Here, we set $\alpha = 5\sqrt{\theta}$ and recall that β is the jump activity index. The calibrated value of λ_∞ is around 20 when $\beta = 1.25$.

When we set $\xi = 0$, there is no self-excitation in jumps on any sample path. There may, but need not, be self-excitation in jumps whenever $\xi > 0$. In Figure I we plot a simulated sample path of the jump intensity process, adopting a benchmark parameter specification of $\xi = 50$. It is clearly visible that intensity jumps occur on days 2 and 4. (Each label along the x-axis indicates the end of a trading day.)

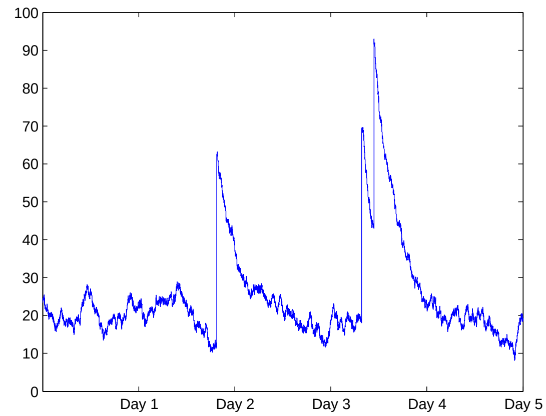

Figure I: A simulated sample path of the jump intensity process λ specified in (5.1). Each label along the x-axis indicates the end of a trading day. We observe intensity jumps on days 2 and 4. (The parameter values are given in the text.)

Remark I. *Observe that*

$$\frac{\mathbb{P}(|(\lambda_\infty + \xi)\Delta_i^n Y| \geq \alpha \Delta_n^{\varpi})}{\mathbb{P}(|\lambda_\infty \Delta_i^n Y| \geq \alpha \Delta_n^{\varpi})} \approx \Big(\frac{\lambda_\infty + \xi}{\lambda_\infty}\Big)^\beta = (1 + \xi/\lambda_\infty)^\beta. \tag{A.3}$$

With $\xi = 50$, the ratio of tail probabilities is around 4.79, implying an economically significant change in the tail probability following the occurrence of a (large, threshold-exceeding, self-exciting) jump. Instead, if one would choose a small value of ξ, for instance $\xi = 1$, then the above ratio is 1.06. In this case, the self-excitation effect will only increase the tail probability by 6% (in relative terms), an increment too small to be economically significant.

Remark II. *Although we do not require the jump intensity process to be stationary in Assumption 2, we include a mean reverting drift term in the above simulation models, specifically in the third equation. The reason is that, without a mean reversion term, the intensity process could be explosive (when $\xi > 0$) and its path might look like a deterministic process. Consequently, log-price jumps would become larger and larger (in absolute value), which is a phenomenon that is at odds with what we observe in practice.*

In Tables A.1 and A.2, we present additional Monte Carlo simulation results for the various alternative sets of parameter values displayed in Table 1 in the paper: one with

10-min sampling frequency and 1 year horizon, and the other with 5-sec sampling frequency and 1 week horizon.

To assess the performance of our tests under various data generating processes, we provide further simulation results in Tables A.3 and A.4. We consider two major scenarios, either with ("SE" scenario) or without ("NO" scenario) self-excitation in jumps. In each scenario, we further consider four different data generating processes: in case I, price jumps follow the stable process as in Assumption 2 and the volatility process does not jump; in case II, we allow for price and volatility co-jumps; case III is the same as case I, except that price jumps are now generated from a CGMY process; whereas case IV includes both CGMY type price jumps and volatility jumps. For simplicity, we fix the size of the volatility jumps to be θ, i.e. the long-run mean when there are no volatility jumps. To avoid simulating an exploding volatility process, we increase the mean-reversion parameter to make it a stationary process and put a threshold on price jumps for them to trigger volatility co-jumps (the average number of price and volatility co-jumps is around 10). As for the specification of the CGMY process, we note that Andersen et al. (2016a) estimate a tempered stable process, which is an extension of the CGMY process. The estimated values for the two parameters controlling negative and positive jumps are around 10 and 65 (larger values mean that it is more likely to have small jumps). For simplicity, we set both to 70 (note that the larger this value is, the more we deviate from the stable process).

The first columns of Tables A.3 and A.4 indicate the type of tests with the specific choices of g and H functions: "NO" and "SE" stand for the No-Self-Excitation test and the Self-Excitation test, respectively. Hence, the upper-left panel gives the size of the No-Self-Excitation test, while the upper-right panel gives its power. On the other hand, the lower-left panel shows the power of the Self-Excitation test, while the lower-right panel shows its size.

Strictly speaking, the identification of the intensity and jump size related parameters in the case of tempered stable processes is not the same as in the current setting (of stable processes). Even in the simplest CGMY setting, there is one additional parameter that controls the jump size distribution. Hence, the estimation procedure is more involved in such cases. To the best of our knowledge, Andersen et al. (2016a,b) estimate this process using additional options data. Yet these authors still find that the α (β here) parameter therein can not be estimated with high precision. From a theoretical point of view, it remains very challenging to design (no) self-excitation test(s) in such a general setting. In particular, it is not clear what the limiting distributions of our test statistics would be if the price jumps followed a general tempered stable process. However, according to the simulation results in Tables A.3 and A.4, it seems safe to say that the testing functions we used in the empirical study ($H(0,1)$ for the "NO" test and $H(0,2)$ for the "SE" test) wouldn't generate false detection of self-excitation, even if the data generating process were a tempered stable process. Note that when there is no self-excitation in jumps, the rejection rates of the two $H(0,1)$- and $H(0,2)$-tests roughly remain the same when deviating from case I. When the truth is self-excitation, the "NO" test becomes less likely to reject its

Table A.1: Simulation results: 10-min sampling frequency and 1 year horizon

		Size of self-excitation test				Power of no self-excitation test			
		$H(0,2)$		$H(6,2)$		$H(0,1)$		$H(6,1)$	
		$g_{2,6}$	g_0	$g_{2,6}$	g_0	$g_{2,6}$	g_0	$g_{2,6}$	g_0
	1%	2.24%	0.24%	2.24%	0.24%	7.20%	2.46%	7.20%	2.46%
1	5%	3.94%	0.64%	3.94%	0.64%	22.14%	12.70%	22.14%	12.70%
	10%	7.28%	1.74%	7.28%	1.74%	34.38%	24.54%	34.38%	24.54%
	1%	1.94%	0.26%	1.94%	0.26%	13.64%	5.50%	13.64%	5.50%
2	5%	5.18%	1.10%	5.18%	1.10%	32.58%	21.24%	32.58%	21.24%
	10%	10.58%	2.94%	10.58%	2.94%	46.44%	35.20%	46.44%	35.20%
	1%	2.34%	0.28%	2.34%	0.28%	5.36%	1.46%	5.36%	1.46%
3	5%	3.72%	0.54%	3.72%	0.54%	19.34%	11.16%	19.34%	11.16%
	10%	6.16%	1.42%	6.16%	1.42%	30.72%	22.04%	30.72%	22.04%
	1%	2.44%	0.14%	2.44%	0.14%	6.74%	2.22%	6.74%	2.22%
4	5%	3.46%	0.32%	3.46%	0.32%	25.78%	14.26%	25.78%	14.26%
	10%	5.96%	0.90%	5.96%	0.90%	39.78%	28.12%	39.78%	28.12%
	1%	1.66%	0.34%	1.66%	0.34%	7.42%	3.08%	7.42%	3.08%
5	5%	4.34%	1.12%	4.34%	1.12%	21.48%	13.22%	21.48%	13.22%
	10%	8.84%	2.92%	8.84%	2.92%	32.98%	24.06%	32.98%	24.06%
	1%	1.98%	0.18%	1.98%	0.18%	10.88%	4.16%	10.88%	4.16%
6	5%	4.22%	0.82%	4.22%	0.82%	29.58%	18.44%	29.58%	18.44%
	10%	8.08%	2.30%	8.08%	2.30%	42.80%	32.60%	42.80%	32.60%
	1%	2.44%	0.20%	2.44%	0.20%	3.64%	1.36%	3.64%	1.36%
7	5%	3.84%	0.54%	3.84%	0.54%	17.72%	9.68%	17.72%	9.68%
	10%	6.12%	1.58%	6.12%	1.58%	28.94%	20.24%	28.94%	20.24%
	1%:	1.98%	0.22%	1.98%	0.22%	9.04%	3.04%	9.04%	3.04%
8	5%:	4.08%	0.82%	4.08%	0.82%	24.96%	16.12%	24.96%	16.12%
	10%:	8.20%	2.14%	8.20%	2.14%	37.08%	26.96%	37.08%	26.96%
	1%:	2.10%	0.24%	2.10%	0.24%	6.56%	2.36%	6.56%	2.36%
9	5%:	3.62%	0.40%	3.62%	0.40%	22.10%	12.48%	22.10%	12.48%
	10%:	5.58%	1.32%	5.58%	1.32%	34.54%	24.70%	34.54%	24.70%

Note: The first column indicates which parameter set of Table 1 is considered. The second column displays the significance level. The remaining columns display the size or power of the relevant test, distinguishing between the employed g and H functions.

Table A.2: Simulation results: 5-sec sampling frequency and 1 week horizon

		Size of self-excitation test				Power of no self-excitation test			
		$H(0,2)$		$H(6,2)$		$H(0,1)$		$H(6,1)$	
		$g_{2,6}$	g_0	$g_{2,6}$	g_0	$g_{2,6}$	g_0	$g_{2,6}$	g_0
	1%	3.14%	0.88%	2.92%	0.98%	73.80%	58.48%	67.16%	49.66%
1	5%	11.00%	3.98%	8.62%	3.56%	87.98%	80.96%	84.40%	75.48%
	10%	18.64%	9.48%	15.72%	7.40%	92.70%	88.60%	90.32%	84.74%
	1%	4.56%	2.24%	4.06%	2.52%	79.48%	66.08%	70.22%	54.56%
2	5%	11.70%	5.82%	10.36%	5.14%	90.86%	85.56%	86.42%	78.80%
	10%	19.00%	10.72%	16.78%	9.34%	94.98%	91.82%	91.74%	87.48%
	1%	3.16%	0.92%	2.42%	0.74%	68.22%	52.14%	60.66%	43.86%
3	5%	11.02%	4.08%	8.58%	2.80%	84.26%	76.04%	79.66%	69.68%
	10%	19.68%	9.22%	16.38%	6.94%	90.72%	85.86%	87.50%	81.42%
	1%	3.04%	0.92%	2.86%	1.06%	73.88%	59.38%	66.70%	49.12%
4	5%	10.56%	3.64%	8.66%	3.02%	88.64%	81.64%	85.18%	76.34%
	10%	18.28%	8.60%	15.90%	7.34%	93.58%	89.38%	91.32%	85.86%
	1%	3.20%	0.96%	2.50%	1.04%	73.20%	58.76%	64.80%	48.28%
5	5%	10.38%	4.22%	8.40%	3.32%	87.14%	79.86%	82.96%	73.62%
	10%	19.06%	9.26%	16.26%	7.18%	92.42%	87.96%	89.40%	83.86%
	1%	2.24%	0.68%	2.30%	1.08%	59.42%	41.26%	50.78%	31.98%
6	5%	6.84%	2.34%	5.94%	2.38%	79.20%	68.32%	73.48%	60.12%
	10%	13.56%	5.80%	11.64%	5.00%	86.88%	80.20%	82.66%	74.40%
	1%	4.50%	1.72%	3.80%	1.70%	83.66%	71.94%	78.42%	63.76%
7	5%	14.06%	6.10%	11.24%	4.62%	93.08%	88.18%	90.60%	84.20%
	10%	22.78%	12.32%	20.26%	9.64%	96.26%	93.58%	94.56%	91.18%
	1%	4.36%	1.52%	3.32%	1.50%	79.34%	66.42%	69.94%	54.78%
8	5%	13.20%	5.84%	10.50%	4.40%	91.26%	85.14%	86.20%	77.72%
	10%	22.10%	11.94%	17.94%	9.20%	94.90%	92.16%	91.62%	87.52%
	1%	2.58%	0.66%	2.20%	0.66%	68.24%	51.46%	61.56%	42.60%
9	5%	8.26%	2.84%	6.88%	2.54%	85.08%	76.26%	81.06%	71.22%
	10%	16.02%	6.92%	13.32%	6.14%	91.12%	85.92%	88.72%	82.38%

Note: The first column indicates which parameter set of Table 1 is considered. The second column displays the significance level. The remaining columns display the size or power of the relevant test, distinguishing between the employed g and H functions.

Table A.3: Robustness check with volatility jumps and the CGMY model (10-min)

		No self-excitation in jumps				Self-excitation in jumps			
		NO: I	NO: II	NO: III	NO: IV	SE: I	SE: II	SE: III	SE: IV
NO: $g_{2,6}$ $H(0,1)$	1%	0.50%	0.52%	0.80%	0.84%	7.20%	6.08%	3.60%	3.12%
	5%	3.08%	3.26%	4.42%	4.56%	22.14%	21.16%	17.86%	17.64%
	10%	6.88%	6.78%	9.06%	8.86%	34.38%	33.78%	30.38%	30.06%
NO: g_0 $H(0,1)$	1%	0.18%	0.12%	0.96%	0.90%	2.46%	2.20%	1.70%	1.80%
	5%	1.58%	1.58%	4.84%	4.80%	12.70%	12.08%	13.38%	13.46%
	10%	4.54%	4.26%	9.92%	10.04%	24.54%	24.36%	25.84%	25.44%
NO: $g_{2,6}$ $H(6,1)$	1%	0.20%	0.20%	0.44%	0.50%	7.20%	6.08%	3.60%	3.12%
	5%	2.16%	2.02%	3.56%	3.48%	22.14%	21.16%	17.86%	17.64%
	10%	5.60%	5.10%	7.16%	7.78%	34.38%	33.78%	30.38%	30.06%
NO: g_0 $H(6,1)$	1%	0.08%	0.08%	0.52%	0.76%	2.46%	2.20%	1.70%	1.80%
	5%	1.04%	1.00%	3.78%	3.88%	12.70%	12.08%	13.38%	13.46%
	10%	3.70%	3.08%	7.94%	8.62%	24.54%	24.36%	25.84%	25.44%
SE: $g_{2,6}$ $H(0,1)$	1%	84.00%	84.02%	70.44%	70.04%	2.24%	2.06%	3.62%	3.38%
	5%	93.54%	93.18%	86.90%	87.00%	3.94%	3.50%	4.76%	4.54%
	10%	95.94%	95.76%	92.32%	92.68%	7.28%	6.28%	5.86%	5.80%
SE: g_0 $H(0,1)$	1%	71.38%	70.28%	71.74%	71.02%	0.24%	0.34%	0.24%	0.20%
	5%	88.16%	87.50%	88.04%	87.92%	0.64%	0.72%	0.68%	0.68%
	10%	92.62%	93.10%	92.64%	92.54%	1.74%	1.32%	1.22%	1.28%
SE: $g_{2,6}$ $H(6,1)$	1%	81.50%	80.68%	66.12%	65.10%	2.24%	2.06%	3.62%	3.38%
	5%	92.54%	92.14%	84.94%	84.46%	3.94%	3.50%	4.76%	4.54%
	10%	95.28%	95.06%	91.06%	91.24%	7.28%	6.28%	5.86%	5.80%
SE: g_0 $H(6,1)$	1%	66.96%	65.92%	67.98%	66.96%	0.24%	0.34%	0.24%	0.20%
	5%	86.08%	85.50%	86.16%	85.34%	0.64%	0.72%	0.68%	0.68%
	10%	91.56%	91.68%	91.34%	91.18%	1.74%	1.32%	1.22%	1.28%

Note: The first column indicates the type of tests with the specific choices of g and H functions: "NO" and "SE" stand for the No-Self-Excitation test and the Self-Excitation test, respectively.

Table A.4: Robustness check with volatility jumps and the CGMY model (5-sec)

		No self-excitation in jumps				Self-excitation in jumps			
		NO: I	NO: II	NO: III	NO: IV	SE: I	SE: II	SE: III	SE: IV
NO: $g_{2,6}$ $H(0,1)$	1%	0.90%	0.72%	0.72%	0.54%	82.72%	82.92%	57.32%	58.16%
	5%	4.22%	3.82%	3.86%	3.76%	92.94%	92.98%	80.86%	81.72%
	10%	8.28%	8.16%	8.24%	7.90%	96.36%	96.16%	89.52%	89.88%
NO: g_0 $H(0,1)$	1%	0.18%	0.16%	0.14%	0.18%	70.66%	71.74%	39.54%	41.04%
	5%	1.70%	1.66%	2.14%	2.22%	88.04%	88.10%	70.74%	70.72%
	10%	5.22%	5.06%	5.90%	5.60%	93.86%	93.58%	83.54%	84.42%
NO: $g_{2,6}$ $H(6,1)$	1%	0.62%	0.72%	0.06%	0.02%	76.12%	76.92%	56.08%	57.22%
	5%	3.62%	3.68%	1.04%	1.08%	89.64%	90.00%	80.28%	81.10%
	10%	7.86%	7.66%	3.82%	3.52%	94.34%	94.40%	89.08%	89.54%
NO: g_0 $H(6,1)$	1%	0.02%	0.14%	0.02%	0.04%	61.64%	62.02%	38.62%	39.72%
	5%	1.64%	1.84%	0.80%	0.58%	83.24%	82.96%	69.74%	70.20%
	10%	4.86%	4.72%	2.82%	2.40%	90.92%	90.78%	82.96%	83.76%
SE: $g_{2,6}$ $H(0,2)$	1%	99.62%	99.02%	99.18%	98.96%	4.84%	4.86%	1.74%	1.84%
	5%	99.86%	99.66%	99.68%	99.56%	14.36%	14.04%	4.84%	5.34%
	10%	99.94%	99.80%	99.88%	99.78%	23.54%	22.66%	9.64%	10.22%
SE: g_0 $H(0,2)$	1%	99.18%	98.16%	98.68%	98.60%	1.80%	1.82%	0.84%	0.76%
	5%	99.72%	99.46%	99.62%	99.48%	6.40%	6.16%	2.26%	2.52%
	10%	99.84%	99.72%	99.74%	99.70%	12.84%	12.44%	4.52%	5.24%
SE: $g_{2,6}$ $H(6,2)$	1%	47.46%	43.96%	78.70%	76.66%	4.12%	3.84%	1.68%	1.82%
	5%	66.90%	63.40%	91.78%	90.58%	11.68%	10.76%	4.46%	5.14%
	10%	76.40%	72.38%	95.38%	94.76%	20.62%	18.78%	9.40%	10.02%
SE: g_0 $H(6,2)$	1%	32.16%	28.86%	67.04%	65.76%	1.80%	1.78%	0.80%	0.84%
	5%	55.08%	51.30%	85.96%	85.32%	5.04%	4.76%	2.26%	2.38%
	10%	66.32%	62.88%	92.64%	91.78%	10.24%	9.30%	4.30%	4.96%

Note: The first column indicates the type of tests with the specific choices of g and H functions: "NO" and "SE" stand for the No-Self-Excitation test and the Self-Excitation test, respectively.

null, which is "no self-excitation in jumps", while the "SE" test remains to give adequate sizes.

Furthermore, the impact of volatility jumps on the size and power of the tests appears to be limited. Intuitively speaking, if the volatility jumps at the same time with the price but the intensity doesn't, the tests won't suggest self-excitation. The reason is as follows. We adopt an adaptive threshold level $\alpha\Delta_n^{\varpi}$, where α is chosen to be 5 times the estimated annualized local volatility. If the volatility jumps together with the price, then the value we choose for α will also increase. That is to say, although price increments will be larger after a volatility jump, most of them will still be smaller than the new threshold level, hence will not lead to an artificial jump in the estimated intensity process.

B Empirical Application

Supported by the Monte Carlo results, we now apply our tests to real equity data. In our empirical application, we analyze equity returns data on 30 individual stocks, which cover all 11 sectors of the Global Industry Classification Standard (GICS): Consumer Discretionary (CD), Consumer Staples (CS), Energy, Financials, Health Care (HC), Industrials, Information Technologies (IT), Materials, Real Estate (RE) and Utilities. Many of these stocks are constituents of the Dow Jones Industrial Average and those that are not also have large market capitalization; they are all very liquid. We downloaded 1-min price data of these stocks from TAQ, filtered out obvious price outliers and then computed 10-min returns from 2006 to 2013. The lower sampling frequency should reduce the risk of microstructure noise considerably, and indeed we find no evidence of noise in our data. We then implement the tests introduced in Sections 4.1 and 4.2 on a yearly basis to each individual stock.

Our choices of the constants and sequence ϖ, ρ, k_n and w exactly mimic our Monte Carlo study, to which we refer for details. Furthermore, throughout this section, $\alpha\Delta_n^{\varpi}$ is taken to be the square root of the integrated volatility over $[0, T]$ times $5\sqrt{\Delta_n/T}$, and β is again estimated using Jing et al. (2012).

For each individual stock, we picked up the largest K number of jumps in each year and tested sequentially (with Benjamini-Hochberg corrections for multiple hypothesis testing) if each such jump exhibits self-excitation or not. The choice of K is inspired by Li et al. (2017). In their paper, the number of identified large market (SPY) jumps over 2007 to 2012 ranges from 8 to 20. Since individual stocks may jump more frequently than the market index, we set $K = 20$ for simplicity (but the overall pattern is similar for $K = 10, 20, 30$ and 40).

We note that this number K is not the sample size. It is more like the number of regressors in a linear regression. The true sample size is the local window size, i.e. k_n, which goes to infinity as the sampling frequency increases. In linear regressions, it is possible that some of the regressors have significant coefficients while others don't. What we are testing here is somewhat similar to testing how many of those regressors are significant. Since we are testing multiple hypotheses at the same time, we employed the Benjamini-Hochberg procedure to correct for the multiple testing bias.

The results are summarized in Table B.1. We conclude that the jump being tested shows evidence of self-excitation (SE) if the self-excitation test does not reject its null, and the no self-excitation test does reject the null hypothesis, both at the 5% significance level (with Benjamini-Hochberg corrections). This classification is similar to Jacod and Todorov (2009, 2010). By design, our procedure is less likely to yield false detection of self-excitation than using only one test. Moreover, the finite-sample simulation results show that the power of the no self-excitation test can be rather low at the 10-min frequency, which implies that even if there is self-excitation, the probability of finding it by our procedure could be as low as 20%. Despite this, there is still quite substantial evidence for the existence of self-excitation: All individual stocks exhibit self-excitation during the sample period. These results reveal that price jumps exert a positive feedback effect and such dependence should be taken into consideration in both theory and practice.

Our results also show that the numbers of identified jumps that excite the intensity (or scale) during the two crisis periods (the financial crisis of 2007-2008 and the peak of the European debt crisis of 2010-2011) are substantially larger than those during the other years. This overall pattern remains broadly the same for K ranging from 10 to 40.

Finally, we also implemented the price and volatility co-jump tests introduced by Jacod and Todorov (2010). At the 5% significance level (with Benjamini-Hochberg corrections), we identified in total 146 price jumps associated with simultaneous volatility jumps, and only 25 of these coincide with the self-excitation jumps identified above. This finding, which is consistent with the corresponding simulation results (in this online appendix), confirms that the above evidence of self-excitation is not a spurious result due to price and volatility co-jumps.

C Preliminary Technical Results

Proposition C.1. *Suppose that Assumptions 1 and 2 hold. Furthermore, let the function $g(\cdot)$ satisfy Assumption 3. Then, for some $\alpha > 0$ and $\varpi \in (0, 1/2)$,*

$$\widehat{\Lambda}_t^n = \Delta_n^{\varpi\beta} \sum_{j=1}^{[t/\Delta_n]} g\left(\frac{|\Delta_j^n X|}{\alpha \Delta_n^{\varpi}}\right) \frac{\alpha^\beta}{C_\beta(1)} \xrightarrow{\mathbb{P}} \int_0^t \lambda_{s-} ds := \Lambda_t,$$

with

$$C_\beta(k) = \int_0^\infty \left(g(x)\right)^k / x^{1+\beta} dx,$$

for any integer k. Note that $C_\beta(k) = 1$ for $g = g_0$.

Proposition C.1 synthesizes the results of the two relevant papers: Aït-Sahalia and Jacod (2009) employ the indicator function $g_0(x) = 1_{\{x>1\}}$, i.e., case (a) of Assumption 3. Jing et al. (2012) suggest to use a smoothed and bounded function g satisfying case (b) of Assumption 3 in order to extract information from relatively small increments, while keeping the contribution from the Brownian part asymptotically negligible.

However, the above result is not sufficient for the purpose of our paper. We need to derive the asymptotic properties of the spot intensity estimator given in (3.6). In the sequel, we are going to accomplish this goal in several steps.

Table B.1: Number of self-exciting jumps per year for 10-min individual equity returns data

Symbol	Sector	2006	2007	2008	2009	2010	2011	2012	2013	Sum
DIS	CD		1		3	1	1		1	7
HD	CD	1	1	1		3				6
MCD	CD	2		1		2	1			6
PG	CS	1				3	2		1	7
WMT	CS	2	1	1	1			1	1	7
WMB	Energy			1	2	2	4		1	10
XOM	Energy		3	1		1	2	1		8
AXP	Financials		2				4			6
BAC	Financials	1	1	3	5		6		1	17
JPM	Financials		1	4	1	1	3	1	1	12
WFC	Financials			2			4	1	1	8
PFE	HC	1	1	1	1		1		1	6
UNH	HC	1		3	1	1	1	2	1	10
BA	Industrials			2		1	1		2	6
CAT	Industrials		4			2	3			9
GE	Industrials	1	1	3	2		1			8
MMM	Industrials	2	1	2		1	1	1		8
UTX	Industrials	1	1	1		3				6
AAPL	IT		3	1	2	3	4		2	15
CSCO	IT		2	1		3	2	1	1	10
INTC	IT	1	2	2		3	3			11
DD	Materials	1	1			1	1			4
FCX	Materials	2	4		1	2	3		1	13
AMT	RE	2		1			2	1	2	8
CBG	RE		1	1			1	1	3	7
CCI	RE	2	1	1		1	2			7
T	TS	2		1		3		1	1	8
VZ	TS	1	1	2	2	3				9
AEP	Utilities		1	2	1	2	1	1		8
PCG	Utilities		1	2		5				8
Sum		24	35	40	22	47	54	12	21	255

Note: The integers in the third to tenth columns represent the number of jumps among the 20 largest jumps for each stock and year that is identified with self-excitation (SE) by both tests. That is, we conclude that the jump being tested shows evidence of self-excitation (SE) if the self-excitation test does not reject its null, while the no self-excitation test does reject its null, both at the 5% significance level. We use $H(6,1)$ with $g_{2,6}$ for the self-excitation test and $H(6,2)$ with $g_{2,6}$ for the no self-excitation test. Since we are testing multiple hypotheses at the same time, we use the Benjamini-Hochberg procedure to correct for the multiple testing bias. The total number of SE jumps in the un-adjusted case would be 789.

By localization, a standard technique used in high frequency analysis (see e.g., Jacod and Protter (2011) and Aït-Sahalia et al. (2016)), we may replace the local boundedness assumptions in Assumptions 1 and 2 by boundedness. We state this stronger assumption as follows:

Assumption C.1. *The processes b and σ are both bounded by some constant L, and Assumption 2 holds with $L_t(\omega) = L$, hence λ is also bounded by L. Furthermore, $\delta(\omega, t, x) \leq \gamma(x)$, where $\gamma(\cdot)$ is a nonnegative bounded function satisfying $\int_{\mathbb{R}} \big(\gamma(x)^\beta \wedge 1\big) F_t(dx) < \infty$, uniformly in t.*

Throughout this appendix, we suppose this assumption holds and primarily consider functions g satisfying part (b) of Assumption 3. As for part (a) of Assumption 3, upon noticing that $g_{1,p}(x) \to g_0(x)$ as $p \to \infty$, the results we derive below can easily be extended to cover this case, too, by applying the dominated convergence theorem. Also, throughout this appendix we denote by K a positive constant, which we allow to change from line to line. Finally, we denote by $\mathbb{E}_{j-1}^n$ and $\mathbb{E}_t$ the conditional expectations with respect to $\mathcal{F}_{t_{j-1}^n}$ and $\mathcal{F}_t$, respectively.

We start by decomposing X_t as $X_t = X_t^c + X_t^d$, where $X_t^c = X_0 + \int_0^t b_s ds + \int_0^t \sigma_s dW_s$, and

$$
\begin{aligned}
X_t^d &= x\, 1_{\{|x| \leq 1\}} * (\mu - \nu)_t + x\, 1_{\{|x|>1\}} * \mu_t \\
&= x\, 1_{\{|x|>\delta\}} * \mu_t + x\, 1_{\{|x| \leq \delta\}} * (\mu - \nu)_t - x\, 1_{\{\delta < |x| \leq 1\}} * \nu_t \\
&=: \; X_t^d(\delta)' + X_t^d(\delta) + B(\delta).
\end{aligned}
$$

Next, denoting $\delta_n = \alpha \Delta_n^{\varpi}$ and $g_n(x) = g(x/\delta_n)$ and recalling (3.6), we consider the following decomposition:

$$
\widehat{\lambda}(k_n)_{t_i^n} - \lambda_{t_i^n} = \frac{\alpha^\beta}{C_\beta(1)} (I_0 + I_1 + I_2 + I_3),
$$

where I_0 is the partial sum of a martingale difference involving $g_n(\Delta_j^n X)$; the summand in I_1 is the conditional expectation of the deviation of $g_n(\Delta_j^n X)$ from its compensator over each small time interval; I_2 represents the discretization error of the jump intensity process over each small time interval; and I_3 represents the discretization error of the jump intensity process over the entire local window. Formally:

$$
\begin{aligned}
I_0 &= \frac{\Delta_n^{\varpi\beta}}{k_n \Delta_n} \sum_{j=i+1}^{i+k_n} \left(g_n(\Delta_j^n X) - \mathbb{E}_{j-1}^n \left(g_n(\Delta_j^n X) \right) \right), \\
I_1 &= \frac{\Delta_n^{\varpi\beta}}{k_n \Delta_n} \sum_{j=i+1}^{i+k_n} \mathbb{E}_{j-1}^n \left(g_n(\Delta_j^n X) - \int_{t_{j-1}^n}^{t_j^n} \int_{\mathbb{R}} g_n(x) F_s(dx) ds \right), \\
I_2 &= \frac{1}{k_n} \sum_{j=i+1}^{i+k_n} \left(\Delta_n^{\varpi\beta - 1} \mathbb{E}_{j-1}^n \int_{t_{j-1}^n}^{t_j^n} \int_{\mathbb{R}} g_n(x) F_s(dx) ds - \lambda_{t_{j-1}^n} \frac{C_\beta(1)}{\alpha^\beta} \right), \\
I_3 &= \frac{1}{k_n} \sum_{j=i+1}^{i+k_n} \lambda_{t_{j-1}^n} \frac{C_\beta(1)}{\alpha^\beta} - \lambda_{t_i^n} \frac{C_\beta(1)}{\alpha^\beta}.
\end{aligned}
\tag{C.1}
$$

(We use the notation $\hat{\lambda}(k_n)_{t_i^n}$ and $\hat{\lambda}(k_n)_i$ interchangeably.) For notational convenience, we let

$$
C_j^n = \int_{t_{j-1}^n}^{t_j^n} \int_{\mathbb{R}} g_n(x) F_s(dx) ds. \tag{C.2}
$$

The remainder of Appendix C is organized as follows. First, we provide in Lemmas C.2 and C.3 upper bounds to $|I_1|$ and $|I_2|$ given in (C.1), respectively. Next, we prove in Lemma C.4 the asymptotic negligibility of the sum $|I_1+I_2+I_3|$ as well as a conditional independence result. These three lemmas will prove very useful in the proofs of the consistency and the associated CLT of the spot jump intensity estimator $\hat{\lambda}$ given in (3.6). The stable convergence of I_0, which, in light of Lemma C.4, is equivalent to the CLT of the spot jump intensity estimator $\hat{\lambda}$, is proved in Proposition C.6. Finally, for the consistency of the spot jump intensity estimator $\hat{\lambda}$, we refer to the first part of the proof of Theorem 1. The following lemma is adapted from Jing et al. (2012).

Lemma C.2. *Let $\phi = \frac{1}{2}(1-\varpi\beta) \wedge \varpi(\beta-\beta') \wedge \varpi\gamma$ and $\phi' = \phi \wedge (p(1/2-\varpi-\epsilon) - (1-\varpi\beta)) \wedge (1/2-\varpi-\epsilon)$. If $p \geq 2$, we have*

$$
|\mathbb{E}_{j-1}^n (g_n(\Delta_j^n X^d) - C_j^n)| \leq K \Delta_n^{1-\varpi\beta+\phi}. \tag{C.3}
$$

$$
|\mathbb{E}_{j-1}^n (g_n(\Delta_j^n X) - C_j^n)| \leq K \Delta_n^{1-\varpi\beta+\phi'}. \tag{C.4}
$$

Furthermore, let $h = 1-\varpi\beta/2$ and $h' = (-\varpi\beta) \wedge (-\varpi\beta - \varpi + \frac{1-\varpi\beta}{2})$. Then, for any bounded martingale M and $0 < s < \Delta_n$,

$$
\begin{aligned}
&|\mathbb{E}_t (M_{t+s} - M_t) g_n(X_{t+s} - X_t)| \\
&\quad \leq K(\Delta_n^{(1-\varpi\beta+\phi')\wedge h} + \epsilon_n \Delta_n^h + \Delta_n^{h'} s \sqrt{\mathbb{E}_t (M_{t+s} - M_t)^2}),
\end{aligned}
\tag{C.5}
$$

where $\epsilon_n \downarrow 0$ as $\Delta_n \to 0$.

Proof. See the proofs of Lemmas 1, 2 and 4 in Jing et al. (2012). □

Lemma C.3. *Let* $\phi'' = \frac{1}{2} \wedge \varpi(\beta - \beta') \wedge (\varpi\gamma)$. *Then,*

$$\left| \Delta_n^{\varpi\beta-1} \mathbb{E}_{j-1}^n (C_j^n) - \lambda_{t_{j-1}^n} \frac{C_\beta(1)}{\alpha^\beta} \right| \leq K \Delta_n^{\phi''}. \tag{C.6}$$

Proof. Recall that Assumption C.1 includes Assumption 2. We consider the following decomposition:

$$\Delta_n^{\varpi\beta-1} \mathbb{E}_{j-1}^n (C_j^n) - \lambda_{t_{j-1}^n} \frac{C_\beta(1)}{\alpha^\beta} = H_1 + H_2 + H_3,$$

where

$$\begin{aligned} H_1 &= \Delta_n^{\varpi\beta-1} \mathbb{E}_{j-1}^n \int_{t_{j-1}^n}^{t_j^n} \int_{\mathbb{R}} \frac{g_n(x)\lambda_s}{|x|^{1+\beta}} dx ds - \lambda_{t_{j-1}^n} \frac{C_\beta(1)}{\alpha^\beta}, \\ H_2 &= \Delta_n^{\varpi\beta-1} \mathbb{E}_{j-1}^n \int_{t_{j-1}^n}^{t_j^n} \int_{\mathbb{R}} \frac{|x|^\gamma h(s,x)}{|x|^{1+\beta}} g_n(x) \lambda_s dx ds, \\ H_3 &= \Delta_n^{\varpi\beta-1} \mathbb{E}_{j-1}^n \int_{t_{j-1}^n}^{t_j^n} \int_{\mathbb{R}} g_n(x) F_s''(dx) ds. \end{aligned}$$

Recall that $\delta_n = \alpha \Delta_n^\varpi$ and $g_n(x) = g(x/\delta_n)$. By the change-of-variable technique ($x \longmapsto x\delta_n$) and Jensen's inequality, we obtain

$$\begin{aligned} |H_1| &= \left| \Delta_n^{\varpi\beta-1} \mathbb{E}_{j-1}^n \int_{t_{j-1}^n}^{t_j^n} \delta_n^{-\beta} \lambda_s \int_{\mathbb{R}} \frac{g(x)}{|x|^{1+\beta}} dx ds - \lambda_{t_{j-1}^n} \frac{C_\beta(1)}{\alpha^\beta} \right| \\ &= \frac{C_\beta(1)}{\alpha^\beta \Delta_n} \left| \int_{t_{j-1}^n}^{t_j^n} \mathbb{E}_{j-1}^n (\lambda_s - \lambda_{t_{j-1}^n}) ds \right| \\ &\leq \frac{C_\beta(1)}{\alpha^\beta \Delta_n} \int_{t_{j-1}^n}^{t_j^n} \left(\mathbb{E}_{j-1}^n (\lambda_s - \lambda_{t_{j-1}^n})^2 \right)^{1/2} ds \\ &\leq \frac{K}{\Delta_n} \int_0^{\Delta_n} \sqrt{s}\, ds < K\sqrt{\Delta_n}. \end{aligned}$$

Next, from Assumption C.1 and the properties of g and F'', it is easy to verify that $|H_2| \leq K\Delta_n^{\varpi\gamma}$ and $|H_3| \leq K\Delta_n^{\varpi(\beta-\beta')}$. Then the result readily follows. □

Lemma C.4. *Assume condition* (3.7) *holds together with* (3.10). *Then,*

$$\sqrt{\frac{k_n \Delta_n}{\Delta_n^{\varpi\beta}}} \left| \frac{\Delta_n^{\varpi\beta}}{k_n \Delta_n} \sum_{j=i+1}^{i+k_n} \mathbb{E}_{j-1}^n \left(g_n(\Delta_j^n X) \right) - \lambda_{t_i^n} \frac{C_\beta(1)}{\alpha^\beta} \right| \xrightarrow{\mathbb{P}} 0. \tag{C.7}$$

Furthermore, if M is a bounded martingale, we have

$$\sqrt{\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}}\sum_{j=i+1}^{i+k_n}\left|\mathbb{E}_{j-1}^n\left(g_n(\Delta_j^nX)\Delta_j^nM\right)\right| \xrightarrow{\mathbb{P}} 0. \tag{C.8}$$

Proof. The left-hand side of (C.7) is smaller than

$$\sqrt{k_n}\Delta_n^{(1-\varpi\beta)/2}\Big(|I_1|+|I_2|+|I_3|\Big).$$

The first two terms in the expression above converge to zero in probability, as direct consequences of (C.4) and (C.6). From the Markov and Jensen's inequalities, for any $\epsilon > 0$,

$$\begin{aligned}
\mathbb{P}\left(\sqrt{k_n}\Delta_n^{(1-\varpi\beta)/2}|I_3|>\epsilon\right) &\leq \frac{K}{\epsilon}\Delta_n^{\frac{1}{2}(1-\varpi\beta-\rho)}\frac{1}{k_n}\sum_{j=i+1}^{i+k_n}\mathbb{E}(|\lambda_{t_{j-1}^n}-\lambda_{t_i^n}|)\\
&\leq \frac{K}{\epsilon}\Delta_n^{\frac{1}{2}(1-\varpi\beta-\rho)}\frac{1}{k_n}\sum_{j=i+1}^{i+k_n}\left(\mathbb{E}(|\lambda_{t_{j-1}^n}-\lambda_{t_i^n}|^2)\right)^{1/2}\\
&\leq \frac{K}{\epsilon}\Delta_n^{\frac{1}{2}(1-\varpi\beta-\rho)}\frac{1}{k_n}k_n\sqrt{k_n\Delta_n} \leq \frac{K}{\epsilon}\Delta_n^{1-\varpi\beta-\rho+\frac{\varpi\beta}{2}} \xrightarrow{\mathbb{P}} 0,
\end{aligned}$$

according to condition (3.10). Thus, (C.7) readily follows.

It remains to show (C.8). The following result readily follows from (C.5) (recall that, in (C.5), $h = 1-\varpi\beta/2$ and $h' = (-\varpi\beta)\wedge(-\varpi\beta-\varpi+\frac{1-\varpi\beta}{2})$):

$$\begin{aligned}
\sqrt{\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}}\sum_{j=i+1}^{i+k_n}\left|\mathbb{E}_{j-1}^n\left(g_n(\Delta_j^nX)\Delta_j^nM\right)\right| &\leq K\Delta_n^{-\frac{1}{2}(1-\varpi\beta-\rho)-\rho}\left(\Delta_n^{(1-\varpi\beta+\phi')\wedge h}+\epsilon_n\Delta_n^h\right)\\
&\quad + K\Delta_n^{-\frac{1}{2}(1-\varpi\beta-\rho)}\sum_{j=i+1}^{i+k_n}\Delta_n^{1+h'}\sqrt{\mathbb{E}_{j-1}^n(\Delta_j^nM)^2}\\
&\leq K\Delta_n^{\frac{1}{2}(1-\varpi\beta-\rho)}\Big(\Delta_n^{\phi'\wedge\frac{1}{2}\varpi\beta}+\epsilon_n\Delta_n^{\frac{1}{2}\varpi\beta}+\Delta_n^{\rho-\frac{1}{2}+\varpi\beta+h'}\sum_{j=i+1}^{i+k_n}\sqrt{\Delta_n\mathbb{E}_{j-1}^n(\Delta_j^nM)^2}\Big).
\end{aligned}$$

By the Cauchy-Schwarz and Jensen's inequalities, we have

$$\sum_{j=i+1}^{i+k_n}\sqrt{\Delta_n\mathbb{E}_{j-1}^n(\Delta_j^nM)^2} \leq \sqrt{k_n\Delta_n}\Big(\sum_{j=i+1}^{i+k_n}\mathbb{E}_{j-1}^n(\Delta_j^nM)^2\Big)^{1/2} \leq K\Delta_n^{1-\rho}.$$

Hence, it is straightforward to verify that the left-hand side of (C.8) is smaller than

$$K\Delta_n^{\frac{1}{2}(1-\varpi\beta-\rho+(2\phi'\wedge\varpi\beta))}+K\epsilon_n\Delta_n^{\frac{1}{2}(1-\rho)}+K\Delta_n^{\frac{1}{2}[(1-\varpi\beta+1-\rho)\wedge(1-\varpi\beta-\rho+1-\varpi\beta+1-2\varpi)]}.$$

Note that $2\phi' \leq (1-\varpi\beta)$ (by its definition), $1 > \varpi\beta$, $1 > \rho$ and $(1-2\varpi) > 0$ (by condition (3.7)). Then, under the assumptions of this lemma, condition (3.10) in particular, it is easy to verify that all powers of Δ_n in the above expression are positive, hence that (C.8) is true. □

Corollary C.5. *We have* $\frac{\Delta^{\varpi\beta}}{k_n\Delta_n}\sum_{j=i+1}^{j=i+k_n}\mathbb{E}_{j-1}^n\left(g_n^2(\Delta_j^n X)\right) \xrightarrow{\mathbb{P}} \lambda_{t_i^n}\frac{C_\beta(2)}{\alpha^\beta}$.

Proof. Replace $g_n(\cdot)$ by its square $g_n^2(\cdot)$, and $C_\beta(1)$ by $C_\beta(2)$ in (C.7). □

Proposition C.6. *Given any finite number of time points $\{t_p\}_{p=1}^P$, the following vector of random variables:*

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\left(\widehat{\lambda}(k_n)_{t_p} - \lambda_{t_p}\right)_{1\leq p\leq P}, \tag{C.9}$$

converges stably in law to a vector of Gaussian random variables $(W_{t_p})_{p=1}^P$ independent of $\mathcal{F}$, with

$$\mathbb{E}_{t_p}(W_{t_p}^2) = \lambda_{t_p}\frac{\alpha^\beta C_\beta(2)}{\left(C_\beta(1)\right)^2} \quad \textit{and} \quad \mathbb{E}_{t_{p\wedge q}}(W_{t_p}W_{t_q}) = 0.$$

Proof. Let $i_p = [t_p/\Delta_n] + 1$ and

$$\zeta_j^n = \sqrt{\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}}\frac{\alpha^\beta}{C_\beta(1)}\left(g_n(\Delta_j^n X) - \mathbb{E}_{j-1}^n\left(g_n(\Delta_j^n X)\right)\right). \tag{C.10}$$

Note that ζ_j^n is a martingale sequence for fixed n. In view of (C.7) in Lemma C.4, it suffices to show that, for any $p \neq q$,

$$\begin{cases} \displaystyle\sum_{j=i_p+1}^{i_p+k_n}\mathbb{E}_{j-1}^n(\zeta_j^n)^2 \xrightarrow{\mathbb{P}} \mathbb{E}_{t_p}(W_{t_p}^2), \\ \displaystyle\sum_{j=i_q+1}^{i_q+k_n}\mathbb{E}_{j-1}^n(\zeta_j^n)^2 \xrightarrow{\mathbb{P}} \mathbb{E}_{t_q}(W_{t_q}^2), \\ \displaystyle\sum_{j=i_p+1}^{i_p+k_n}\sum_{k=i_q+1}^{i_q+k_n}\mathbb{E}_{j\wedge k-1}^n(\zeta_j^n\zeta_k^n) \xrightarrow{\mathbb{P}} 0, \end{cases} \tag{C.11}$$

$$\begin{cases} \displaystyle\sum_{j=i_p+1}^{i_p+k_n}\mathbb{E}_{j-1}^n(\zeta_j^n\Delta_j^n M) \xrightarrow{\mathbb{P}} 0, \\ \displaystyle\sum_{j=i_q+1}^{i_q+k_n}\mathbb{E}_{j-1}^n(\zeta_j^n\Delta_j^n M) \xrightarrow{\mathbb{P}} 0, \end{cases} \tag{C.12}$$

and

$$\begin{cases} \sum_{j=i_p+1}^{i_p+k_n} \mathbb{E}_{j-1}^n(\zeta_j^n)^4 \xrightarrow{\mathbb{P}} 0, \\ \sum_{j=i_q+1}^{i_q+k_n} \mathbb{E}_{j-1}^n(\zeta_j^n)^4 \xrightarrow{\mathbb{P}} 0. \end{cases} \tag{C.13}$$

First of all, from (C.10), we have

$$\mathbb{E}_{j-1}^n(\zeta_j^n)^2 = \frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big(\frac{\alpha^\beta}{C_\beta(1)}\Big)^2\Big(\mathbb{E}_{j-1}^n\left(g_n^2(\Delta_j^n X)\right) - \left(\mathbb{E}_{j-1}^n(g_n(\Delta_j^n X))\right)^2\Big).$$

Using Corollary C.5 and the assumption that λ has right-continuous paths almost surely, we can conclude that

$$\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big(\frac{\alpha^\beta}{C_\beta(1)}\Big)^2 \sum_{j=i_p+1}^{i_p+k_n} \mathbb{E}_{j-1}^n\left(g_n^2(\Delta_j^n X)\right) \xrightarrow{\mathbb{P}} \lambda_{i_p}\frac{\alpha^\beta C_\beta(2)}{\left(C_\beta(1)\right)^2} \xrightarrow{\mathbb{P}} \lambda_{t_p}\frac{\alpha^\beta C_\beta(2)}{\left(C_\beta(1)\right)^2}.$$

Furthermore, from (C.4) and (C.6), we obtain

$$\begin{aligned} &\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big(\frac{\alpha^\beta}{C_\beta(1)}\Big)^2 \sum_{j=i_p+1}^{i_p+k_n} \left(\mathbb{E}_{j-1}^n(g_n(\Delta_j^n X))\right)^2 \\ &\leq K\Delta_n^{1-\varpi\beta}\frac{1}{k_n}\sum_{j=i_p+1}^{i_p+k_n}\lambda^2_{t_{j-1}^n} + K\Delta_n^{1-\varpi\beta+2(\phi'\wedge\phi'')} \leq K\Delta_n^{1-\varpi\beta} \xrightarrow{\mathbb{P}} 0. \end{aligned}$$

Hence, the first two statements in (C.11) are true.

Next, note that for n large enough, we have $k_n\Delta_n < \min_{p\neq q}|t_p - t_q|$, or equivalently $k_n < \min_{p\neq q}|i_p - i_q|$. This fact yields the last equation of (C.11). Moreover, it is easy to see that (C.12) follows directly from (C.8).

As for (C.13), we have

$$\begin{aligned} \mathbb{E}_{j-1}^n(\zeta_j^n)^4 = \Big(\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big)^2\Big(\frac{\alpha^\beta}{C_\beta(1)}\Big)^4\Big(&\mathbb{E}_{j-1}^n(g_n^4(\Delta_j^n X)) - 4\big(\mathbb{E}_{j-1}^n(g_n^3(\Delta_j^n X))\big)\left(\mathbb{E}_{j-1}^n(g_n(\Delta_j^n X))\right) \\ &+ 6\big(\mathbb{E}_{j-1}^n(g_n^2(\Delta_j^n X))\big)\left(\mathbb{E}_{j-1}^n(g_n(\Delta_j^n X))\right)^2 - 3\big(\mathbb{E}_{j-1}^n(g_n(\Delta_j^n X))\big)^4\Big). \end{aligned}$$

Lemmas C.2 and C.3, together with Assumption C.1, yield

$$\left|\mathbb{E}_{j-1}^n(g_n^m(\Delta_j^n X))\right| \leq \left|\mathbb{E}_{j-1}^n(C_j^n(m))\right| + o_p(\Delta_n^{1-\varpi\beta}) \leq K\Delta_n^{1-\varpi\beta},$$

where $C_j^n(m)$ is defined by replacing $g_n(\cdot)$ by $\big(g_n(\cdot)\big)^m$ in C_j^n. As a consequence, we obtain

$$\sum_{j=i_p+1}^{i_p+k_n} \mathbb{E}_{j-1}^n(\zeta_j^n)^4 \le K\Big(\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big)^2 \sum_{j=i_p+1}^{i_p+k_n} \Delta_n^{1-\varpi\beta} = K\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n} = K\Delta_n^{\varpi\beta+\rho-1}.$$

According to condition (3.10), $\varpi\beta+\rho-1>0$. Therefore, the right-hand side converges to zero as n goes to infinity. Hence, the desired result readily follows. □

Lemma C.7. *The conclusion of Proposition C.6 also holds when* (C.9) *is replaced by its feasible counterpart*

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\,\Big(\widehat{\lambda}(\widehat{\beta},k_n)_{t_p}-\lambda_{t_p}\Big)_{1\le p\le P},$$

where $\widehat{\beta}$ is a consistent estimator of β (see Aït-Sahalia and Jacod (2009) and Jing et al. (2012)).

Proof. First, we note that it is sufficient to show that the difference between the feasible and the infeasible estimation error process is asymptotically negligible, that is, it is sufficient to prove that

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(\widehat{\lambda}(\widehat{\beta})-\lambda\Big)-\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\widehat{\lambda}(\beta)-\lambda\Big)=o_P(1).$$

Here we omit the argument k_n for ease of notation.

Simple calculations yield that

$$\begin{aligned}
&\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(\widehat{\lambda}(\widehat{\beta})-\lambda\Big)-\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\widehat{\lambda}(\beta)-\lambda\Big)\\
&=\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(\widehat{\lambda}(\widehat{\beta})-\lambda\Big)-\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\widehat{\lambda}(\beta)-\lambda\Big)+\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(\widehat{\lambda}(\beta)-\lambda\Big)-\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(\widehat{\lambda}(\beta)-\lambda\Big)\\
&=\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(\widehat{\lambda}(\widehat{\beta})-\widehat{\lambda}(\beta)\Big)-\Big(\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}-\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big)\Big(\widehat{\lambda}(\widehat{\beta})-\lambda\Big)\\
&=\sqrt{k_n\Delta_n}\sqrt{\Delta_n^{\varpi(\beta-\widehat{\beta})}}\,\frac{1}{\sqrt{\Delta_n^{\varpi\beta}}}\Big(\widehat{\lambda}(\widehat{\beta})-\widehat{\lambda}(\beta)\Big)+\Big(\sqrt{\Delta_n^{\varpi(\beta-\widehat{\beta})}}-1\Big)\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\widehat{\lambda}(\beta)-\lambda\Big).
\end{aligned}$$

Hence, it is sufficient to show that the above two terms are of order $o_P(1)$.

Second, Proposition C.6 shows that

$$\sqrt{\frac{k_n \Delta_n}{\Delta_n^{\varpi\beta}}}\left(\widehat{\lambda}(\beta)-\lambda\right)=O_P(1).$$

Therefore, if $\Delta_n^{\varpi(\beta-\widehat{\beta})}\xrightarrow{\mathbb{P}}1$, the second term above will be of order $o_P(1)$.

Third, the convergence rate for $\widehat{\beta}$ (see Aït-Sahalia and Jacod (2009) and Jing et al. (2012)) is $\sqrt{1/\Delta_n^{\varpi\beta}}$. Therefore,

$$\frac{1}{\sqrt{\Delta_n^{\varpi\beta}}}\left(\widehat{\beta}-\beta\right)=O_P(1).$$

From the expression of $\widehat{\lambda}$, we can see that $\widehat{\lambda}$ is a continuous function of β. Then the continuous mapping theorem yields

$$\frac{1}{\sqrt{\Delta_n^{\varpi\beta}}}\left(\widehat{\lambda}(\widehat{\beta})-\widehat{\lambda}(\beta)\right)=O_P(1).$$

Note that $k_n\Delta_n\to 0$. Consequently, if $\Delta_n^{\varpi(\beta-\widehat{\beta})}\xrightarrow{\mathbb{P}}1$, the first term above will also be of order $o_P(1)$.

What remains now is to show that $\Delta_n^{\varpi(\beta-\widehat{\beta})}\xrightarrow{\mathbb{P}}1$. In fact, we can prove that $\lim_{x\to 0+}x^{cx^a}=1$ for some $a>0$ and $|c|<\infty$. Let $y(x)=x^{cx^a}$ and note that $\ln y(x)=cx^a\ln x=\frac{c\ln x}{x^{-a}}$. Then, according to l'Hôpital's rule, we have

$$\lim_{x\to 0+}\log y(x)=\lim_{x\to 0+}\frac{c\ln x}{x^{-a}}=\lim_{x\to 0+}\frac{c/x}{-ax^{-a-1}}=-\frac{c}{a}\lim_{x\to 0+}x^a=0.$$

Therefore, $\lim_{x\to 0+}y(x)=\lim_{x\to 0+}x^{cx^a}=1$. This result further implies that

$$\lim_{\Delta_n\to 0+}\Delta_n^{\varpi(\beta-\widehat{\beta})}\xrightarrow{\mathbb{P}}1,$$

by observing that $\widehat{\beta}-\beta=O_P(1)\,\Delta_n^{\varpi\beta/2}$. The conclusion then follows. □

D Proofs of the Limit Theorems

Before proceeding to the proofs of the main theorems, we introduce some additional notation used in the sequel; this notation is more commonly used in high frequency analysis with jumps, see e.g., Jacod and Todorov (2010) and Aït-Sahalia et al. (2016). First, for each integer $m \geq 1$, we define $(S(m,q) : q \geq 1)$ to be the successive jump times of the Poisson random measure given by

$$\mu^X \left([0,t] \times \left\{ x : \frac{1}{m} < \gamma(x) \leq \frac{1}{m-1} \right\} \right).$$

(Please refer to (2.1) for μ^X and to Assumption C.1 for $\gamma(\cdot)$.) One may relabel the two-parameter sequence $(S(m,q) : m,q \geq 1)$ to become a single sequence $(T_p : p \geq 1)$ that exhausts the jumps of X. Next, we define

$$\begin{gathered} A_m = \{z : \gamma(z) \leq 1/m\}, \quad \gamma_{m,t} = \int_{A_m} \big(\gamma(x)\big)^\beta F_t(dx), \\ \mathcal{T}_m^t = \{p : \exists\, p \geq 1 \text{ and } m' \in \{1, \cdots, m\} \text{ s.t. } T_p = S(m',q) \leq [t/\Delta_n]\Delta_n\}, \\ i(n,p) = \text{the unique integer such that } T_p \in (t_{i-1}^n, t_i^n], \\ J(n,m,t) = \{i(n,p) : p \in \mathcal{T}_m^t\}, \quad J'(n,m,t) = \{1, \cdots, [t/\Delta_n]\} \setminus J(n,m,t), \\ T_-(n,p) = t_{i(n,p)-1}^n, \quad T_+(n,p) = t_{i(n,p)}^n, \\ \Omega_{n,m,t} = \bigcap_{p \neq q, p,q \in \mathcal{T}_m} \{T_p < [t/\Delta_n]\Delta_n - 3k_n\Delta_n \text{ or } T_p > 3k_n\Delta_n; \text{ and } |T_p - T_q| > 6k_n\Delta_n\}. \end{gathered}$$

Observe that, for any m,

$$\lim_{n \to \infty} \mathbb{P}(\Omega_{n,m,t}) = 1.$$

Hence, it is sufficient to prove all the desired results on this set. Furthermore, we sometimes decompose X in an alternative way, using A_m, as follows:

$$X_t = X'(m)_t + L(m)_t + J(m)_t,$$

where

$$b(m)_t = \begin{cases} b_t - \displaystyle\int_{(A_m)^c} x F_t(dx), & \text{if } \beta > 1, \\ b_t, & \text{if } \beta \leq 1, \end{cases} \qquad L(m)_t = \begin{cases} \displaystyle\int_0^t \int_{A_m} x(\mu - \nu)(ds,dx), & \text{if } \beta > 1, \\ \displaystyle\int_0^t \int_{A_m} x\mu(ds,dx), & \text{if } \beta \leq 1, \end{cases}$$

$$X'(m) = X_0 + \int_0^t b'(m)_s ds + \int_0^t \sigma_s dW_s, \quad J(m)_t = \int_0^t \int_{(A_m)^c} x\mu(ds,dx).$$

In the sequel, we first prove each theorem with true β and in the last subsection we show that the same conclusions hold for the feasible estimators as well.

D.1 Proof of Theorem 1

We first prove consistency of the spot jump intensity estimator $\widehat{\lambda}(k_n)_{t_i}$ given in (3.6). Recall the first equation in (C.1) and equation (C.10). One easily verifies that

$$I_0 = \frac{C_\beta(1)}{\alpha^\beta}\sqrt{\frac{\Delta^{\varpi\beta}}{k_n\Delta_n}}\sum_{j=i+1}^{i+k_n}\zeta_j^n.$$

Because $\mathbb{E}_{j-1}^n(\zeta_j^n) = 0$, in order to show $I_0 \xrightarrow{\mathbb{P}} 0$, it is sufficient to prove that

$$\frac{\Delta^{\varpi\beta}}{k_n\Delta_n}\sum_{j=i+1}^{i+k_n}\mathbb{E}_{j-1}^n(\zeta_j^n)^2 \xrightarrow{\mathbb{P}} 0.$$

Lemmas C.2 and C.3 yield that

$$\mathbb{E}_{j-1}^n(g_n^2(\Delta_j^n X)) = O_p(\Delta_n^{1-\varpi\beta}),$$
$$\big(\mathbb{E}_{j-1}^n(g_n^2(\Delta_j^n X))\big)^2 = O_p(\Delta_n^{2(1-\varpi\beta)}).$$

Hence,

$$\begin{aligned}\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big|\sum_{j=i+1}^{i+k_n}\mathbb{E}_{j-1}^n(\zeta_j^n)^2\Big| &\le \Big(\frac{\Delta_n^{\varpi\beta}}{k_n\Delta_n}\Big)^2\sum_{j=i+1}^{i+k_n}\Big|\mathbb{E}_{j-1}^n(g_n^2(\Delta_j^n X)) - \big(\mathbb{E}_{j-1}^n(g_n^2(\Delta_j^n X))\big)^2\Big| \\ &\le K\big(\Delta_n^{\varpi\beta+\rho-1} + \Delta_n^{\rho}\big).\end{aligned}$$

The right-hand side of the expression above converges to zero as soon as $\rho > 0$ (condition (3.7)) and $\rho > 1 - \varpi\beta$. Thus, in view of (the proof of) Lemma C.4, one readily gets the desired consistency result: $\widehat{\lambda}(k_n)_{t_i} \xrightarrow{\mathbb{P}} \lambda_{t_i^n}$.

Next, on the set $\Omega_{n,m,t}$, we decompose $U(H, k_n)$ into two parts, as follows:

$$U(H, k_n) = \widetilde{U}^n(m) + \overline{U}^n(m),$$

with

$$\begin{cases}\widetilde{U}^n(m)_t = \displaystyle\sum_{p\in\mathcal{T}_m^t} H(X_{i(n,p)-1}, X_{i(n,p)}, \widehat{\lambda}(k_n)_{p-}, \widehat{\lambda}(k_n)_{p+})\, 1_{\{|\Delta_{i(n,p)}^n X|>\alpha\Delta_n^{\varpi}\}}, \\ \overline{U}^n(m)_t = \displaystyle\sum_{j\in J'(n,m,t)} H(X_{j-1}, X_j, \widehat{\lambda}(k_n)_{j-k_n-1}, \widehat{\lambda}(k_n)_j)\, 1_{\{|\Delta_j^n X|>\alpha\Delta_n^{\varpi}\}}.\end{cases}$$

Observe that, for $p \in J(n, m, t)$, we have

$$X_{i(n,p)} = X_{i(n,p)-1} + \Delta X_{T_p} + O_p(\sqrt{\Delta_n}).$$

Furthermore, by (3.8), one easily verifies that $\mathbb{P}(\Delta X_{T_p} \in R) = 1$. Since H is continuous on $D(R) \times \mathbb{R}^2$ and $\widehat{\lambda}(k_n)_{p\pm}$ is a consistent estimator of $\lambda_{T_{p\pm}}$, the continuous mapping theorem yields that for any m and t,

$$\widetilde{U}^n(m)_t \xrightarrow{\mathbb{P}} \widetilde{U}(m)_t := \sum_{p \in \mathcal{T}_m^t} H(X_{T_p-}, X_{T_p+}, \lambda_{T_p-}, \lambda_{T_p+})\, 1_{\{\Delta X_{T_p} \neq 0\}}.$$

Consider case (a) stated in the theorem. For that given ϵ, choose $m > 2/\epsilon$. Then, for any $j \in J'(n, m, t)$, $|\Delta_j^n X| \leq \epsilon/2$ on $\Omega_{n,m,t}$ with sufficiently large n. Therefore, we have $\overline{U}^n(m)_u = 0$ for all $u \leq t$. On the other hand, $U(m) \equiv U(H)$. Hence, the result readily follows.

Next, we turn to case (b) stated in the theorem. By the assumption, for any s such that $|\Delta X_s| \leq \epsilon$, we have

$$H(X_{s-}, X_{s+}, \lambda_{s-}, \lambda_{s+}) \leq K|\Delta X_s|^\beta.$$

(Recall that by Assumption C.1, λ is bounded.) Furthermore, because β is the jump activity index, $\sum_{s \leq t} |\Delta X_s|^\beta < \infty$. Then, by the dominated convergence theorem, one readily gets $U(m) \to U(H)$ a.s., locally uniformly in time, as m goes to infinity. Thus, what is left, is to prove that, for all $t > 0$ and $\eta > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\big(\sup_{u \leq t} |\overline{U}^n(m)_u| > \eta\big) = 0.$$

First let m be large enough, next let n be sufficiently large, so we have $|\Delta_j^n X| = |\Delta_j^n(X'(m) + J(m))| \leq \epsilon/2$ for any $j \in J'(n, m, t)$ on $\Omega_{n,m,t}$. In addition, the Markov inequality (in the extended version for monotonically increasing functions) yields

$$\mathbb{P}\big(|\Delta_j^n X'(m)| > \alpha \Delta_n^\varpi/2\big) \leq \frac{\mathbb{E}\big(|\Delta_j^n X'(m)|^k\big)}{(\alpha\Delta_n^\varpi/2)^k} \leq K\Delta_n^{k(1/2-\varpi)},$$

for any integer $k \geq 1$. Henceforth, we choose $k > \frac{1}{1/2-\varpi}$. Now consider the set

$$\Omega_{n,m,t}^\varpi := \Omega_{n,m,t} \bigcap_{i=1}^{[t/\Delta_n]} \{|\Delta_j^n X'(m)| \leq \alpha\Delta_n^\varpi/2\}. \tag{D.1}$$

It is easy to see that if $\varpi < 1/2$,

$$\begin{aligned}\lim_{n\to\infty} \mathbb{P}(\Omega_{n,m,t}^\varpi) &= 1 - \big(1 - \lim_{n\to\infty} \mathbb{P}\big(\Omega_{n,m,t}\big)\big) - \lim_{n\to\infty} \sum_{i=1}^{[t/\Delta_n]} \mathbb{P}\big(|\Delta_j^n X'(m)| > \alpha\Delta_n^\varpi/2\big) \\ &\geq 1 - 0 - \lim_{n\to\infty} \sum_{i=1}^{[t/\Delta_n]} K\Delta_n^{k(1/2-\varpi)} = 1 - \lim_{n\to\infty} K\Delta_n^{k(1/2-\varpi)-1} = 1.\end{aligned}$$

Furthermore, if $|x'| \leq \alpha\Delta_n^{\varpi}/2$, $|x''| \leq \epsilon/2$ and $\alpha\Delta_n^{\varpi} < \epsilon$ (easily satisfied with n sufficiently large), then

$$|x' + x''| > \alpha\Delta_n^{\varpi} \Longrightarrow |x'| + |x''| > \alpha\Delta_n^{\varpi} \Longrightarrow |x'| \leq \alpha\Delta_n^{\varpi}/2 < |x''|.$$

As a consequence, the assumed property of H yields that

$$|H(x_1, x_1 + x' + x'', y_1, y_2)|\, 1_{\{|x'+x''|>\alpha\Delta_n^{\varpi}\}} \leq K|x''|^{\beta}(1 + y_1 + y_2).$$

Therefore, on the set $\Omega_{n,m,t}^{\varpi}$ we obtain

$$|\overline{U}^n(m)_t| \leq K \sum_{j \in J'(n,m,t)} |\Delta_j^n J(m)|^{\beta}(1 + \widehat{\lambda}(k_n)_{j-k_n-1} + \widehat{\lambda}(k_n)_j).$$

On the other hand,

$$\mathbb{E}\big(|\Delta_j^n J(m)|^{\beta}(1 + \widehat{\lambda}(k_n)_{j-k_n-1} + \widehat{\lambda}(k_n)_j)\big) \leq K\gamma_m\Delta_n,$$

where $\gamma_m := \sup_t \gamma_{m,t} < \infty$, according to Assumption C.1. Observe that as m goes to infinity, $\gamma_m \longrightarrow 0$. Then, by first letting n and next letting m go to infinity, the desired result readily follows from the Markov inequality. This completes the proof of Theorem 1.

D.2 Proof of Theorem 2

We continue to use the notation $\widetilde{U}^n(m)$ and $\overline{U}^n(m)$ as in the proof of Theorem 1 when the bandwidth is k_n, and we write $\widetilde{U}'^n(m)$ and $\overline{U}'^n(m)$ instead when dealing with wk_n. In addition to $\widetilde{U}(m)$, we also define

$$\overline{U}(m)_t := U_t - \widetilde{U}(m)_t = \sum_{p \notin \mathcal{T}_m^t} H(X_{T_p-}, X_{T_p+}, \lambda_{T_p-}, \lambda_{T_p+})\, 1_{\{\Delta X_{T_p} \neq 0\}}.$$

That is, $\overline{U}(m)$ represents the part of the limiting process that is associated with "small" log-price jumps, while $\widetilde{U}(m)$ represents the part associated with "large" jumps. Our proof is subdivided into 7 steps.

Step 1. We first focus on the approximation error to $\widetilde{U}(m)$. Denote

$$\widetilde{U}^n(m)_t - \widetilde{U}(m)_t = \sum_{p \in \mathcal{T}_m^t} \sum_{j=1}^{3} \tilde{\chi}(m,j)_p^n, \quad \text{and} \quad \widetilde{U}'^n(m)_t - \widetilde{U}(m)_t = \sum_{p \notin \mathcal{T}_m^t} \sum_{j=1}^{3} \tilde{\chi}'(m,j)_p^n,$$

where

$$\tilde{\chi}(m,1)_p^n = H(X_{i(n,p)-1}, X_{i(n,p)}, \widehat{\lambda}(k_n)_{p-}, \widehat{\lambda}(k_n)_{p+})\left(1_{\{|\Delta_{i(n,p)}^n X|>\alpha\Delta_n^{\varpi}\}} - 1_{\{\Delta X_{T_p} \neq 0\}}\right),$$

$$\tilde{\chi}(m,2)_p^n = \big(H(X_{i(n,p)-1}, X_{i(n,p)}, \widehat{\lambda}(k_n)_{p-}, \widehat{\lambda}(k_n)_{p+}) - H(X_{T_p-}, X_{T_p+}, \widehat{\lambda}(k_n)_{p-}, \widehat{\lambda}(k_n)_{p+})\big)\, 1_{\{\Delta X_{T_p} \neq 0\}},$$

$$\tilde{\chi}(m,3)_p^n = \big(H(X_{T_p-}, X_{T_p+}, \widehat{\lambda}(k_n)_{p-}, \widehat{\lambda}(k_n)_{p+}) - H(X_{T_p-}, X_{T_p+}, \lambda_{T_p-}, \lambda_{T_p+})\big)\, 1_{\{\Delta X_{T_p} \neq 0\}}.$$

In addition, we also set

$$\begin{aligned}\tilde{\chi}(m,4)_p^n = \big(&H_3'(X_{T_p-}, X_{T_p+}, \lambda_{T_p-}, \lambda_{T_p+})(\widehat{\lambda}(k_n)_{T_p-} - \lambda_{T_p-}) \\ &+ H_4'(X_{T_p-}, X_{T_p+}, \lambda_{T_p-}, \lambda_{T_p+})(\widehat{\lambda}(k_n)_{T_p+} - \lambda_{T_p+})\big)\, 1_{\{\Delta X_{T_p}\neq 0\}}.\end{aligned}$$

Similarly, we define each $\tilde{\chi}'(m,j)_p^n$ in the same way, except that the bandwidth k_n is replaced by wk_n.

When $p \in \mathcal{T}_m^t$, we must have $\Delta X_{T_p} \neq 0$. In this case,

$$\begin{aligned}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\mathbb{E}\big(|1_{\{|\Delta_{i(n,p)}^n X|>\alpha\Delta_n^{\varpi}\}} - 1_{\{\Delta X_{T_p}\neq 0\}}|\big) &= \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\mathbb{P}\big(|\Delta_{i(n,p)}^n X| \leq \alpha\Delta_n^{\varpi}\big) \\ &\leq K\sqrt{\Delta_n^{1-\rho-\varpi\beta}}\,\mathbb{P}\big(|\Delta_{i(n,p)}^n X - \Delta X_{T_p}| \geq 1/m - \alpha\Delta_n^{\varpi}\big) \\ &\leq K\sqrt{\Delta_n^{1-\rho-\varpi\beta}}\,\frac{\mathbb{E}\big(|\Delta_{i(n,p)}^n X - \Delta X_{T_p}|^2\big)}{(1/m-\alpha\Delta_n^{\varpi})^2} = \frac{K\sqrt{\Delta_n^{1-\rho-\varpi\beta}}\,\Delta_n}{(1/m-\alpha\Delta_n^{\varpi})^2}.\end{aligned}$$

For any integer m, the right-hand side term will converge to zero as n goes to infinity. Then, as convergence in L^1-norm implies convergence in probability, we can conclude that, for any $p \in \mathcal{T}_m^t$,

$$\lim_{m\to\infty}\limsup_{n\to\infty}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\tilde{\chi}(m,1)_p^n \xrightarrow{\mathbb{P}} 0. \tag{D.2}$$

Next, in view of condition (c) of Assumption 4, we obtain

$$\begin{aligned}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,|\tilde{\chi}(m,2)_p^n| &\leq \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}K\big(|X_{i(n,p)-1} - X_{T_p-}|^{q_c'} + |X_{i(n,p)} - X_{T_p+}|^{q_c'}\big)|\Delta X_{T_p}|^{q_d'} \\ &\leq K\sqrt{k_n\Delta_n^{1-\varpi\beta}}\,O_p\big(\sqrt{\Delta_n}\big) = O_p\big(\sqrt{\Delta_n^{1-\rho}\Delta_n^{1-\varpi\beta}}\big) \xrightarrow{\mathbb{P}} 0.\end{aligned}$$

The last line results from the assumptions that $q_c', q_d' \geq 1$, $1-\rho > 0$ and $1-\varpi\beta > 0$, and the fact that $|\Delta X_{T_p}|$ ($p \in \mathcal{T}_m^t$), $\frac{1}{\sqrt{\Delta_n}}|X_{i(n,p)-1} - X_{T_p-}|$ and $\frac{1}{\sqrt{\Delta_n}}|X_{i(n,p)} - X_{T_p+}|$ are bounded in probability. Hence, (D.2) also holds for $\tilde{\chi}(m,2)_p^n$.

Moreover, by Taylor expansion, we also have

$$\begin{aligned}&\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\big(\tilde{\chi}(m,3)_p^n - \tilde{\chi}(m,4)_p^n\big) \\ &= \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,O_p\Big((\widehat{\lambda}(k_n)_{T_p-} - \lambda_{T_p-})^2 + (\widehat{\lambda}(k_n)_{T_p+} - \lambda_{T_p+})^2\Big) \xrightarrow{\mathbb{P}} 0.\end{aligned}$$

The last convergence result follows from the fact that $\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}(\widehat{\lambda}(k_n)_{T_p\pm} - \lambda_{T_p\pm})$ is bounded in probability and that $\widehat{\lambda}(k_n)_{T_p\pm} \xrightarrow{\mathbb{P}} \lambda_{T_p\pm}$.

Note that the same line of argument also applies to the case where the bandwidth is wk_n. Thus, we have proved that for each $p \in \mathcal{T}_m^t$,

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\sum_{j=1}^{3}\tilde{\chi}(m,j)_p^n - \tilde{\chi}(m,4)_p^n\Big) \xrightarrow{\mathbb{P}} 0 \text{ and } \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\sum_{j=1}^{3}\tilde{\chi}'(m,j)_p^n - \tilde{\chi}'(m,4)_p^n\Big) \xrightarrow{\mathbb{P}} 0.$$

The above convergence results also hold when we aggregate over $p \in \mathcal{T}_m^t$.

Step 2. Now we turn to study the asymptotic properties of $(\tilde{\chi}(m,4)_p^n, \tilde{\chi}'(m,4)_p^n)$. For the set of finite number time points $\{T_p\}_{p=1}^P$, the conclusion of Proposition C.6 can be generalized to include different values of bandwidths. Recall that, for any finite m, the set $\mathcal{T}_m^t$ is finite for any given t. Consequently, one can verify that

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\sum_{p\in\mathcal{T}_m^t}(\tilde{\chi}(m,4)_p^n, \tilde{\chi}'(m,4)_p^n) \xrightarrow{\mathscr{L}_{st}} \Big(\mathcal{U}(m)_t, \frac{1}{w}(\mathcal{U}(m)_t + \sqrt{w-1}\,\mathcal{U}'(m)_t)\Big),$$

where $\mathcal{U}(m)_t$ and $\mathcal{U}'(m)_t$ are defined just like $\mathcal{U}_t$ and $\mathcal{U}'_t$ in (3.9), except that the sum is taken over $p \in \mathcal{T}_m^t$, instead of over $p \in \mathcal{T}^t$.

Together with the results from the previous Step 1, for any integer m, we obtain the following result:

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\widetilde{U}^n(m)_t - \widetilde{U}(m)_t, \widetilde{U}'^n(m)_t - \widetilde{U}(m)_t\Big) \xrightarrow{\mathscr{L}_{st}} \Big(\mathcal{U}(m)_t, \frac{1}{w}(\mathcal{U}(m)_t + \sqrt{w-1}\,\mathcal{U}'(m)_t)\Big).$$

Furthermore, it is straightforward to verify that

$$\begin{aligned}\widetilde{\mathbb{E}}\big(|\mathcal{U}_t - \mathcal{U}(m)_t|^2|\mathcal{F}\big) &= \frac{\alpha^\beta C_\beta(2)}{\big(C_\beta(1)\big)^2}\sum_{s\le t}\big(\lambda_{s-}H_3'(X_{s-},X_s,\lambda_{s-},\lambda_s)^2 \\ &\quad + \lambda_s H_4'(X_{s-},X_s,\lambda_{s-},\lambda_s)^2\big)\,1_{\{|\Delta X_s|\le 1/m\}} \\ &\le K\sum_{s\le t}|\Delta X_s|^{2q}\,1_{\{|\Delta X_s|\le 1/m\}} \le K\int_0^t \gamma_{m,s}ds.\end{aligned}$$

Thus, as m goes to infinity, the right-hand side shrinks to zero a.s. because $\sup_{s\le t}\gamma_{m,t} \to 0$ as m goes to infinity (by assumption). This implies that $\mathcal{U}(m) \xrightarrow{u.c.p.} \mathcal{U}$ (convergence in probability, uniformly in time). Following the same line of reasoning, we also obtain $\mathcal{U}'(m) \xrightarrow{u.c.p.} \mathcal{U}'$. Hence, so far, we have proved that

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(\widetilde{U}^n(m)_t - \widetilde{U}(m)_t, \widetilde{U}'^n(m)_t - \widetilde{U}(m)_t\Big) \xrightarrow{\mathscr{L}_{st}} \Big(\mathcal{U}_t, \frac{1}{w}(\mathcal{U}_t + \sqrt{w-1}\,\mathcal{U}'_t)\Big).$$

Note that if condition (a) in the theorem is satisfied with some $\epsilon > 0$, then one can set $m \geq 2/\epsilon$. Then, $\overline{U}^n(m) \equiv 0$ for sufficiently large n and $\overline{U}(m) \equiv 0$. This completes the proof of part (a) of the theorem.

Step 3. Now let us turn to study the approximation error to $\overline{U}(m)$. What is left to be proved is that for all $t < \infty$ and $\eta > 0$,

$$\lim_{m\to\infty} \limsup_{n\to\infty} \mathbb{P}\Big(\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\big|\overline{U}^n(m)_t - \overline{U}(m)_t\big| > \eta\Big) = 0, \tag{D.3}$$

and the same for $\overline{U}'^n(m)_t - \overline{U}(m)_t$. It is sufficient to prove (D.3) only.

Define $Y(m) = X'(m) + L(m)$, i.e., the sum of the continuous component and the "small" jumps of X. Hence, for any $i \in J'(n,m,t)$, we have $\Delta_i^n Y(m) = \Delta_i^n X$. As a consequence, we may rewrite $\overline{U}^n(m)_t$ as

$$\overline{U}^n(m)_t = \sum_{i\in J'(n,m,t)} H(X_{i-1}, X_{i-1} + \Delta_i^n Y(m), \widehat{\lambda}(k_n)_{j-k_n-1}, \widehat{\lambda}(k_n)_j)\, 1_{\{|\Delta_i^n Y(m)|>\alpha\Delta_n^{\varpi}\}}.$$

Define $I(n,i) = ((i-1)\Delta_n, i\Delta_n]$ and $J(n,t) = (0, k_n\Delta_n] \cup (([t/\Delta_n] - k_n)\Delta_n, t]$. Observe that on the set $\Omega(n,m,t)$, although "large" jumps do not fall in $J(n,t)$, this does not hold for "small" jumps. Another important issue to note is that for each $p \in \mathcal{T}_m^t$, "small" jumps may still occur within the interval $I(n, i(n,p))$, hence contribute to $\overline{U}(m)_t$. For simplicity, let $I_p^n := I(n, i(n,p))$ and $\widehat{\lambda}(k_n)_{i-} = \widehat{\lambda}(k_n)_{i-k_n-1}$.

In view of the discussion above, we can decompose $\overline{U}^n(m)_t - \overline{U}(m)_t$ as

$$\overline{U}^n(m)_t - \overline{U}(m)_t = \sum_{i\in J'(n,m,t)} \sum_{j=1}^{3} \bar{\chi}(m,j)_i^n + \sum_{j=1}^{2} \overline{V}(m,j)_t^n,$$

where

$$\overline{V}(m,1)_t^n = -\sum_{p\in\mathcal{T}_m^t} \sum_{s\in I_p^n} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda_{s-}, \lambda_s)\, 1_{\{\Delta Y(m)_s \neq 0\}},$$
$$\overline{V}(m,2)_t^n = -\sum_{s\in J(n,t)} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda_{s-}, \lambda_s)\, 1_{\{\Delta Y(m)_s \neq 0\}},$$

and

$$\begin{aligned}\bar{\chi}(m,1)_i^n &= \big(H(X_{i-1}, X_{i-1}+\Delta_i^n Y(m), \widehat{\lambda}(k_n)_{j-k_n-1}, \widehat{\lambda}(k_n)_j) \\ &\quad - H(X_{i-1}, X_{i-1}+\Delta_i^n Y(m), \lambda_{i-1}, \lambda_i)\big)\, 1_{\{|\Delta_i^n Y(m)|>\alpha\Delta_n^{\varpi}\}}, \\ \bar{\chi}(m,2)_i^n &= H(X_{i-1}, X_{i-1}+\Delta_i^n Y(m), \lambda_{i-1}, \lambda_i)\, 1_{\{|\Delta_i^n Y(m)|>\alpha\Delta_n^{\varpi}\}} \\ &\quad - \sum_{s\in I(n,i)} H(X_{s-}, X_s+\Delta Y(m)_s, \lambda_{i-1}, \lambda_i)\, 1_{\{\Delta Y(m)_s\neq 0\}}, \\ \bar{\chi}(m,3)_i^n &= \sum_{s\in I(n,i)} \big(H(X_{s-}, X_s+\Delta Y(m)_s, \lambda_{i-1}, \lambda_i) \\ &\quad - H(X_{s-}, X_s+\Delta Y(m)_s, \lambda_{s-}, \lambda_s)\big)\, 1_{\{\Delta Y(m)_s\neq 0\}}.\end{aligned}$$

Step 4. We are going to prove that

$$\lim_{m\to\infty}\limsup_{n\to\infty}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\mathbb{E}\big(\big|\overline{V}(m,j)_t^n\big|\big)=0, \tag{D.4}$$

for $j=1,2$.

By condition (b) of Assumption 4, we have

$$\sum_{s\in I(n,i)}\big|H(X_{s-}, X_{s-}+\Delta Y(m)_s, \lambda_{s-}, \lambda_s)\big| \le K|\Delta Y(m)_s|^q =: a(n,i).$$

Moreover, by the property of $Y(m)$, we obtain

$$\mathbb{E}\big(a(n,i)\big) \le K\int_{t_{j-1}^n}^{t_j^n}\int_{A_m}\mathbb{E}\big((\gamma(x))^q F_t(dx)\big) \le K\gamma_m\Delta_n.$$

Recall that for any m, $\{p\in\mathcal{T}_m^t\}$ is a finite set. Then, it is straightforward to obtain that

$$\limsup_{n\to\infty}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\mathbb{E}\big(\big|\overline{V}(m,j)_t^n\big|\big) \le \limsup_{n\to\infty}\sum_{p\in\mathcal{T}_m^t}K\gamma_m\Delta_n^{(1-\rho+1-\varpi\beta+1)/2}=0.$$

The last equality readily follows from condition (3.7). Hence, (D.4) holds for $j=1$.

As for $j=2$, we have the following result:

$$\begin{aligned}\limsup_{n\to\infty}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\,\mathbb{E}\big(\big|\overline{V}(m,j)_t^n\big|\big) &\le \limsup_{n\to\infty}\Delta_n^{(1-\varpi\beta-\rho)/2}k_nK\gamma_m\Delta_n \\ &= \limsup_{n\to\infty}K\gamma_m\Delta_n^{\frac{3}{2}(1-\rho-\varpi\beta+\frac{2}{3}\varpi\beta)}=0.\end{aligned}$$

The last equality follows from the assumption that $\rho < 1 - \varpi\beta + \frac{1}{2}\varpi\beta$ (recall condition (3.10)). Hence, (D.4) is also true when $j = 2$.

Step 5. What we will show in this step is that

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}} \sum_{i\in J'(n,m,t)} \mathbb{E}\Big(\big|\bar{\chi}(m,j)_i^n\big|\Big) = 0, \tag{D.5}$$

holds for $j = 1$.

First, condition (b) of Assumption 4 yields

$$\big|\bar{\chi}(m,1)_i^n\big| \le K|\Delta_i^n Y(m)|^q \big(|\widehat{\lambda}(k_n)_{i-} - \lambda_{i-1}| + |\widehat{\lambda}(k_n)_i - \lambda_i|\big)\, 1_{\{|\Delta_i^n Y(m)| > \alpha\Delta_n^{\varpi}\}}.$$

Next, Proposition C.6 implies that

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\, \big|\widehat{\lambda}(k_n)_{i-} - \lambda_{i-1}\big| \quad \text{and} \quad \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\, \big|\widehat{\lambda}(k_n)_i - \lambda_i\big|$$

are bounded in probability. Moreover, following the same argument as in the previous subsection, we obtain that for sufficiently large n

$$|\Delta_i^n Y(m)|^q\, 1_{\{|\Delta_i^n Y(m)| > \alpha\Delta_n^{\varpi}\}} \le K|\Delta_i^n L(m)|^q$$

on the set $\Omega_{n,m,t}^{\varpi}$ (see (D.1)). Consequently, on this set it is easy to obtain the following evaluation by successive conditioning:

$$\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}} \sum_{i\in J'(n,m,t)} \mathbb{E}\Big(\big|\bar{\chi}(m,1)_i^n\big|\Big) \le \sum_{i\in J'(n,m,t)} K\gamma_m\Delta_n \le Kt\gamma_m.$$

Since $\lim_{m\to\infty}\gamma_m = 0$, we readily deduce (D.5) for $j = 1$.

Step 6. Now we prove (D.5) for $j = 3$ on the set $\Omega_{n,m,t}^{\varpi}$. One can verify that

$$|\bar{\chi}(m,3)_i^n| \le K \sum_{s\in I(n,i)} |\Delta Y(m)_s|^q \big(|\lambda_{s-} - \lambda_{i-1}| + |\lambda_i - \lambda_s|\big).$$

By successive conditioning, we obtain

$$\begin{aligned}\mathbb{E}\Big(\sum_{s\in I(n,i)} |\Delta Y(m)_s|^q \big(|\lambda_{s-} - \lambda_{i-1}|\big) &\le \gamma_m \int_{I(n,i)} \mathbb{E}\big(|\lambda_{s-} - \lambda_{i-1}|\big)ds \\ &\le \gamma_m \int_{I(n,i)} \sqrt{\mathbb{E}\big(|\lambda_{s-} - \lambda_{i-1}|^2\big)}ds \le K\gamma_m\Delta_n^{3/2},\end{aligned}$$

and

$$\mathbb{E}\Big(\sum_{s\in I(n,i)} |\Delta Y(m)_s|^q \left(|\lambda_i - \lambda_s|\right)\Big) \leq \int_{I(n,i)} \mathbb{E}\big(|\Delta Y(m)_s|^q \sqrt{s}\big) ds \leq K\gamma_m \Delta_n^{3/2}.$$

These results yield that

$$\begin{aligned}\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}} \sum_{i\in J'(n,m,t)} \mathbb{E}\Big(\big|\bar{\chi}(m,3)_i^n\big|\Big) &\leq \sum_{i\in J'(n,m,t)} K\gamma_m\Delta_n\Delta_n^{(1-\rho+1-\varpi\beta)/2} \\ &\leq Kt\gamma_m\Delta_n^{(1-\rho+1-\varpi\beta)/2}.\end{aligned}$$

Then the desired result readily follows from the assumption that $\rho < 1$, $\varpi < 1/2$ and $\beta < 2$.

Step 7. Now we turn to the case $j = 2$. We introduce a new random variable:

$$\begin{aligned}\bar{\chi}(m,4)_i^n &= H(X_{i-1}, X_{i-1} + \Delta_i^n Y(m), \lambda_{i-1}, \lambda_i)\, 1_{\{|\Delta_i^n Y(m)| > \alpha\Delta_n^{\varpi}\}} \\ &\quad - \sum_{s\in I(n,i)} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda_{i-1}, \lambda_i)\, 1_{\{|\Delta Y(m)_s| > \alpha\Delta_n^{\varpi}\}}.\end{aligned}$$

Observe that

$$\big|\bar{\chi}(m,2)_i^n - \bar{\chi}(m,4)_i^n\big| = \sum_{s\in I(n,i)} H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda_{i-1}, \lambda_i)\, 1_{\{|\Delta Y(m)_s| \leq \alpha\Delta_n^{\varpi}\}}.$$

For any m, we can always have $\alpha\Delta_n^{\varpi} < 1/m$ when n is sufficiently large. Hence, we deduce from Assumption 4 that

$$\begin{aligned}\mathbb{E}\big(\big|\bar{\chi}(m,2)_i^n - \bar{\chi}(m,4)_i^n\big|\big) &\leq K \sum_{s\in I(n,i)} \mathbb{E}\big(|\Delta Y(m)_s|^q 1_{\{|\Delta Y(m)_s| \leq \alpha\Delta_n^{\varpi}\}}\big) \\ &\leq K \int_{I(n,i)} \int_{\gamma(x)\leq\alpha\Delta_n^{\varpi}} \mathbb{E}\big((\gamma(x))^q F_t(dx) ds\big) \\ &\leq K\Delta_n^{\varpi(q-\beta)} \int_{I(n,i)} \int_{A_m} \mathbb{E}\big((\gamma(x))^\beta F_t(dx) ds\big) \\ &\leq K\gamma_m\Delta_n^{1+\varpi(q-\beta)}.\end{aligned} \tag{D.6}$$

So condition (b) in the theorem implies that, for any m,

$$\limsup_{n\to\infty} \sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}} \sum_{i\in J'(n,m,t)} \mathbb{E}\Big(\big|\bar{\chi}(m,2)_i^n - \bar{\chi}(m,4)_i^n\big|\Big) = \limsup_{n\to\infty} K\gamma_m\Delta_n^{\frac{1}{2}(1-\rho-\varpi\beta)+\varpi(q-\beta)} = 0.$$

To conduct a detailed analysis of $\bar{\chi}(m,4)_i^n$, we need additional notation. The basic idea consists in further distinguishing "small jumps" according to whether they are "close" to the truncation level. For some $l \in (1, 1/(2\varpi\beta))$, we denote $q_n = [(\alpha\Delta_n^{\varpi})^{-l}]$ and we suppose that n is sufficiently large so that $1/q_n < \alpha\Delta_n^{\varpi} < 1/m$. Next, define

$$A'_n = A_m \cap (A_{q_n}^c), \quad N_t^n = \mu([0,t] \times A'_n),$$
$$L'_t = \int_0^t \int_{A'_n} \delta(\omega, u, z)\mu(du, dz), \quad L(q_n) = L(m) - L',$$
$$G(n,i) = \{|\Delta_i^n X'(m)| \leq \alpha\Delta_n^{\varpi}/4\} \cap \{|\Delta_i^n L(q_n)| \leq \alpha\Delta_n^{\varpi}/4\} \cap \{\Delta_i^n N^n \leq 1\}.$$

Accordingly, let $Y(q_n) = X'(m) + L(q_n)$.

We then evaluate the probability of $\omega \in G(n,i)$. First, for $q \geq \beta$ it is easy to see that

$$\int_{A_{q_n}} \delta(t,x)^q F_t(dx) = \int_{A_{q_n}} \delta^{\beta}\delta^{q-\beta} F_t(dx) \leq K\Delta_n^{l\varpi(q-\beta)} \int_{A_{q_n}} \delta^{\beta} F_t(dx) \leq K\Delta_n^{l\varpi(q-\beta)}\gamma_m.$$

And, for any $\varrho \geq \beta$, we have

$$\mathbb{E}\big(|\Delta_i^n X'(m)|^{\varrho}\big) \leq K(\Delta_n^{\varrho/2}) \ \text{ and } \ \mathbb{E}\big(|\Delta_i^n L(q_n)|^{\varrho}\big) \leq \Delta_n^{1+l\varpi(\varrho-\beta)}\gamma_m.$$

These results, applied with $\varrho = \frac{4}{1-2\varpi}$ to the first term and with $\varrho = \frac{1+l\beta\varpi}{\varpi(l-1)}$ to the second one, and the Markov inequality yield

$$\mathbb{P}(|\Delta_i^n X'(m)| > \alpha\Delta_n^{\varpi}/4) \leq K\Delta_n^2 \ \text{ and } \ \mathbb{P}(|\Delta_i^n L(q_n)| > \alpha\Delta_n^{\varpi}/4) \leq K\Delta_n^2.$$

Second, N^n is a Poisson process with parameter $\lambda(A'_n) \leq K\gamma_m q_n^{\beta}$. We therefore obtain

$$\mathbb{P}(\Delta_i^n N^n = 1) \propto \Delta_n^{1-l\varpi\beta}\gamma_m \quad \text{and} \quad \mathbb{P}(\Delta_i^n N^n \geq 2) \leq K\Delta_n^{2-2l\varpi\beta}\gamma_m^2.$$

Let $\Omega(G)_{n,t} = \bigcap_{1\leq i\leq[t/\Delta_n]} G(n,i)$. Then we have

$$\mathbb{P}(\Omega(G)_{n,t}^c) \leq \sum_{i=1}^{[t/\Delta_n]} \mathbb{P}(G(n,i)^c) \leq tK\Delta_n^{1-2l\varpi\beta}\gamma_m^2 \longrightarrow 0.$$

Hence, it is sufficient to prove the desired results on the intersection of $\Omega_{n,m,t}^{\varpi}$ and $\Omega(G)_{n,t}$. Observe that on the set $G(n,i)$ we have

$$|\Delta_i^n Y(q_n)| \leq |\Delta_i^n X'(m)| + |\Delta_i^n L(q_n)| \leq \alpha\Delta_n^{\varpi}/2 < \alpha\Delta_n^{\varpi}.$$

Therefore, for any $i \in J'(n,m,t)$, if $\Delta_i^n N^n = 0$, i.e., $|\Delta Y(m)_s| \leq q_n$ for all $s \in I(n,i)$, then

$$|\Delta_i^n L(m)| = |\Delta_i^n L(q_n)| \leq \alpha\Delta_n^{\varpi}/4,$$
$$|\Delta_i^n X| = |\Delta_i^n Y(m)| = |\Delta_i^n Y(q_n)| \leq |\Delta_i^n X'(m)| + |\Delta_i^n L(q_n)| \leq \alpha\Delta_n^{\varpi}/2.$$

Then it is obvious that, conditional on $\Delta_i^n N^n = 0$, $\bar{\chi}(m,4)_i^n = 0$. When $\Delta_i^n N^n = 1$, there is one and only one $s \in I(n,i)$ such that $|\Delta Y(m)_s| > q_n$. In view of these results, define

$$\bar{\chi}(m,5)_i^n = \sum_{s\in I(n,i)} \Big(H(X_{s-}, X_{s-} + \Delta_i^n Y(m), \lambda_{i-1}, \lambda_i)\, 1_{\{|\Delta_i^n Y(m)|>\alpha\Delta_n^{\varpi}, |\Delta Y(m)_s|>q_n\}} \\ - H(X_{s-}, X_{s-} + \Delta Y(m)_s, \lambda_{i-1}, \lambda_i)\, 1_{\{|\Delta Y(m)_s|>\alpha\Delta_n^{\varpi}\}} \Big).$$

Again, condition (c) in Assumption 4 implies that

$$\begin{aligned} &\mathbb{E}\big(\bar{\chi}(m,4)_i^n - \bar{\chi}(m,5)_i^n\big) \\ &\le K\,\mathbb{E}\Big(\big(|X_{s-} - X_{i-1}|^{q'_c} + |X_s - X_{i-1}|^{q'_c}\big) \cdot |\Delta_i^n Y(m)|^{q'_d} \cdot 1_{\{|\Delta Y(m)_s|>q_n\}} \Big) \\ &\le K\Big(\mathbb{E}\big((|X_{s-} - X_{i-1}|^{q'_c} + |X_s - X_{i-1}|^{q'_c})^2 \big) \cdot \mathbb{E}\big(|\Delta_i^n Y(m)|^{2q'_d}\big) \Big)^{1/2} \\ &\le K\Big(\big(\Delta_n^{q'_c} + \Delta_n^{1+\varpi(2q'_c-\beta)}\gamma_m\big) \Delta_n^{1+\varpi(2q'_d-\beta)}\gamma_m \Big)^{1/2} \\ &\le K\Delta_n\big(\Delta_n^{(q'_c-1+\varpi(2q'_d-\beta))} + \Delta_n^{2\varpi(q'_c+q'_d-\beta)}\gamma_m\big)^{1/2}\sqrt{\gamma_m}. \end{aligned} \tag{D.7}$$

Therefore, as long as $1-\rho-\varpi\beta + (q'_c - 1 + \varpi(2q'_d-\beta)) \wedge 2\varpi(q'_c + q'_d - \beta) \ge 0$, we obtain (on the set $\Omega_{n,m,t} \cap \Omega(G)_{n,t}$) the following result:

$$\lim_{m\to\infty} \limsup_{n\to\infty} \sqrt{\frac{k_n \Delta_n}{\Delta_n^{\varpi\beta}}} \sum_{i\in J'(n,m,t)} \mathbb{E}\Big(\big|\bar{\chi}(m,4)_i^n - \bar{\chi}(m,5)_i^n\big| \Big) = 0.$$

Hence, it is sufficient to prove (D.5) for $j = 5$.

For any i and any $s \in I(n,i)$, define the following function:

$$f(x)_s^i := H(X_{s-}, X_{s-} + x, \lambda_{i-1}, \lambda_i).$$

So Assumption 4 implies that $f(x)_s^i \le K|x|^q$ and $|f(x+y)_s^i - f(x)_s^i| \le K|y|^{q'_c}|x|^{q'_d}$. Consequently, some elementary calculations show that

$$\begin{aligned} \big|\bar{\chi}(m,5)_i^n\big| &\le \sum_{s\in I(n,i)} \big| f(\Delta_i^n Y(m))_s^i \cdot 1_{\{|\Delta_i^n Y(m)|>\alpha\Delta_n^{\varpi}, |\Delta Y(m)_s|>q_n\}} \\ &\quad - f(\Delta Y(m)_s)_s^i \cdot 1_{\{|\Delta Y(m)_s|>\alpha\Delta_n^{\varpi}\}} \big| \\ &\le K \sum_{s\in I(n,i)} \Big(|\Delta_i^n Y(m)|^q \cdot 1_{\{q_n<|\Delta Y(m)_s|\le\alpha\Delta_n^{\varpi}\}} + |\Delta Y(m)_s|^q \cdot 1_{\{|\Delta Y(m)_s|>\alpha\Delta_n^{\varpi}\}} \\ &\quad \cdot 1_{\{|\Delta_i^n Y(m)|\le\alpha\Delta_n^{\varpi}\}} + |\Delta_i^n Y(m) - \Delta Y(m)_s|^{q'_c} \cdot |\Delta Y(m)_s|^{q'_d} \cdot 1_{\{|\Delta Y(m)_s|>\alpha\Delta_n^{\varpi}\}} \Big) \\ &\le K\big(|\Delta_i^n Y(m)| \wedge 2\alpha\Delta_n^{\varpi}\big)^q + K \sum_{s\in I(n,i)} \Big(\big(|\Delta Y(m)_s| \wedge 2\alpha\Delta_n^{\varpi}\big)^q \\ &\quad + |\Delta_i^n Y(m) - \Delta Y(m)_s|^{q'_c} \cdot |\Delta Y(m)_s|^{q'_d} \Big) \cdot 1_{\{|\Delta Y(m)_s|>\alpha\Delta_n^{\varpi}\}}. \end{aligned}$$

Recall that on the set $G(n,i)$, there is at most one $s \in I(n,i)$ such that $|\Delta Y(m)_s| > \alpha\Delta_n^{\varpi}$. For simplicity, let $1(Y)_s^n := 1_{\{|\Delta Y(m)_s|>\alpha\Delta_n^{\varpi}\}}$ and $\sum_{i,s} := \sum_{i=1}^{[t/\Delta_n]}\sum_{s\in I(n,i)}$. In view of the previous calculations, we obtain

$$\sum_{i=1}^{[t/\Delta_n]} \mathbb{E}\Big(\big(|\Delta_i^n Y(m)| \wedge 2\alpha\Delta_n^{\varpi}\big)^q\Big) \le \sum_{i=1}^{[t/\Delta_n]} \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}\int_{2\alpha\Delta_n^{\varpi}} \delta(s,x)^q F_s(dx)ds \le K\Delta_n^{\varpi(q-\beta)}\gamma_m t,$$

$$\sum_{i,s} \mathbb{E}\Big(\big(|\Delta Y(m)_s| \wedge 2\alpha\Delta_n^{\varpi}\big)^q \cdot 1(Y)_s^n\Big) \le \int_0^t \mathbb{E}\int_{2\alpha\Delta_n^{\varpi}} \delta(s,x)^q F_s(dx)ds \le K\Delta_n^{\varpi(q-\beta)}\gamma_m t.$$

Moreover, we obtain

$$\begin{aligned}
&\sum_{i,s} \mathbb{E}\Big(|\Delta_i^n Y(m) - \Delta Y(m)_s|^{q_c'} \cdot |\Delta Y(m)_s|^{q_d'} \cdot 1(Y)_s^n\Big) \\
&\le \sum_{i,s} \sqrt{\mathbb{E}\big(|\Delta_i^n Y(m) - \Delta Y(m)_s|^{2q_c'} \cdot 1(Y)_s^n\big)\, \mathbb{E}\big(|\Delta Y(m)_s|^{2q_d'} \cdot 1(Y)_s^n\big)} \\
&\le \Big(\sum_{i,s} \mathbb{E}\big(|\Delta_i^n Y(m) - \Delta Y(m)_s|^{2q_c'} \cdot 1(Y)_s^n\big) \cdot \sum_{i,s} \mathbb{E}\big(|\Delta Y(m)_s|^{2q_d'} \cdot 1(Y)_s^n\big)\Big)^{1/2} \\
&\le \Big(\sum_{i=1}^{[t/\Delta_n]} \int_{(i-1)\Delta_n}^{i\Delta_n} \mathbb{E}\int_{q_n} \delta(s,x)^{2q_c'} F_s(dx)ds \cdot \int_0^t \mathbb{E}\int_{q_n} \delta(s,x)^{2q_d'} F_s(dx)ds\Big)^{1/2} \\
&\le K\Delta_n^{\varpi(q_c'+q_d'-\beta)}\gamma_m t.
\end{aligned}$$

As a consequence, we have the following estimate:

$$\sum_{i\in J'(n,m,t)} \mathbb{E}\Big(\big|\bar{\chi}(m,5)_i^n\big|\Big) \le Kt\gamma_m\Delta_n^{\varpi(q\wedge(q_c'+q_d')-\beta)}. \tag{D.8}$$

Therefore, when condition (b) of Theorem 2 is satisfied, we have (D.5) for $j = 5$. This completes the proof.

D.3 Proof of Theorem 3

The adopted notation is the same as in the proof of Theorem 2, except that $\tilde{\chi}(m,4)_p^n$ will be replaced by

$$\begin{aligned}
\bar{\xi}_p^n = \frac{k_n\Delta_n}{2\Delta_n^{\varpi\beta}}\Big(&H_{33}''(X_{T_p-}, X_{T_p}, \lambda_{T_p-}, \lambda_{T_p})(\widehat{\lambda}(k_n)_{T_p-} - \lambda_{T_p-})^2 \\
&+ H_{44}''(X_{T_p-}, X_{T_p}, \lambda_{T_p-}, \lambda_{T_p})(\widehat{\lambda}(k_n)_{T_p} - \lambda_{T_p})^2 \\
&+ 2H_{34}''(X_{T_p-}, X_{T_p}, \lambda_{T_p-}, \lambda_{T_p})(\widehat{\lambda}(k_n)_{T_p-} - \lambda_{T_p-})(\widehat{\lambda}(k_n)_{T_p} - \lambda_{T_p})\Big).
\end{aligned} \tag{D.9}$$

Our proof is subdivided into 3 steps.

Step 1. Using the calculations we performed in Step 1 of the proof of Theorem 2, we obtain

$$\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\,\mathbb{E}\big(|1_{\{|\Delta^n_{i(n,p)}X|>\alpha\Delta_n^{\varpi}\}}-1_{\{\Delta X_{T_p}\neq 0\}}|\big)\leq \frac{K\Delta_n^{(1-\rho)+(1-\varpi\beta)}}{(1/m-\alpha\Delta_n^{\varpi})^2},$$
$$\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\,\mathbb{E}\big(\big|\tilde\chi(m,2)^n_p\big|\big)\leq K\Delta_n^{1-\rho-\varpi\beta+1/2}.$$

Then the following conclusion readily follows from the assumption that $\rho<1$, $\varpi\beta<1$ and condition (3.10):

$$\lim_{m\to\infty}\limsup_{n\to\infty}\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\,\tilde\chi(m,j)^n_p\xrightarrow{\mathbb{P}}0, \tag{D.10}$$

for $j=1$ and $j=2$.

Taking the Taylor expansion of the function H up to the second order, we obtain

$$\begin{aligned}&H(x_1,x_2,\hat y_1,\hat y_2)-H(x_1,x_2,y_1,y_2)\\&\approx H_3'(x_1,x_2,y_1,y_2)(\hat y_1-y_1)+H_4'(x_1,x_2,y_1,y_2)(\hat y_2-y_2)\\&\quad+\frac{1}{2}\Big(H''_{33}(x_1,x_2,y_1,y_2)(\hat y_1-y_1)^2+H''_{44}(x_1,x_2,y_1,y_2)(\hat y_2-y_2)^2\\&\qquad+2H''_{34}(x_1,x_2,y_1,y_2)(\hat y_1-y_1)(\hat y_2-y_2)\Big).\end{aligned}$$

Now set $x_1=X_{T_p-}$, $x_2=X_{T_p}$, $y_1=\lambda_{T_p-}$ and $y_2=\lambda_{T_p}$. The degeneracy condition induces that the first order terms are all zero. Thus, we obtain

$$\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\big(\tilde\chi(m,3)^n_p-\bar\xi^n_p\big)=\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\,O_p\Big((\widehat\lambda(k_n)_{T_{p-}}-\lambda_{T_{p-}})^3+(\widehat\lambda(k_n)_{T_{p+}}-\lambda_{T_{p+}})^3\Big)\xrightarrow{\mathbb{P}}0.$$

Furthermore, we can obtain the following result using a similar argument as in Step 2 of the proof of Theorem 2: $\sum_{p\in\mathcal{T}^t_m}\bar\xi^n_p\xrightarrow{\mathscr{L}_{st}}\overline{\mathcal{U}}(m)_t$. Similarly, from Assumption 5, we obtain $\widetilde{\mathbb{E}}(|\overline{\mathcal{U}}(m)_t-\overline{\mathcal{U}}_t|\mid\mathcal{F})\leq K\sum_{s\leq t}|\Delta X_s|^q 1_{\{|\Delta X_s|\leq m\}}\leq K\gamma_m t$. Then, invoking the same line of reasoning as in Step 2 of the proof of Theorem 2, we prove part (a) of Theorem 3.

Step 2. Now we turn to part (b). First, based on the results in Step 4 of the proof of Theorem 2, it is easy to verify that, for $j=1$ and $j=2$ respectively,

$$\limsup_{n\to\infty}\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\,\mathbb{E}\big(\big|\overline{V}(m,j)^n_t\big|\big)\leq\limsup_{n\to\infty}\sum_{p\in\mathcal{T}^t_m}K\gamma_m\Delta_n^{(1-\rho)+(1-\varpi\beta)}=0,$$
$$\limsup_{n\to\infty}\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\,\mathbb{E}\big(\big|\overline{V}(m,j)^n_t\big|\big)\leq\limsup_{n\to\infty}K\gamma_m\Delta_n^{2(1-\rho-\varpi\beta+\frac{1}{2}\varpi\beta)}=0.$$

Next, condition (b) of Assumption 5 yields

$$
\begin{aligned}
\left|\bar{\chi}(m,1)_i^n\right| &\le K|\Delta_i^n Y(m)|^q\big(|\widehat{\lambda}(k_n)_{i-}-\lambda_{i-1}|^2+|\widehat{\lambda}(k_n)_i-\lambda_i|^2\big)\,1_{\{|\Delta_i^n Y(m)|>\alpha\Delta_n^{\varpi}\}},\\
|\bar{\chi}(m,3)_i^n| &\le K\sum_{s\in I(n,i)}|\Delta Y(m)_s|^q\left(|\lambda_{s-}-\lambda_{i-1}|^2+|\lambda_i-\lambda_s|^2\right).
\end{aligned}
$$

Then we readily obtain

$$
\lim_{m\to\infty}\limsup_{n\to\infty}\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\sum_{i\in J'(n,m,t)}\mathbb{E}\Big(\left|\bar{\chi}(m,1)_i^n\right|\Big)\le\lim_{m\to\infty}Kt\gamma_m=0,
$$

$$
\lim_{m\to\infty}\limsup_{n\to\infty}\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\sum_{i\in J'(n,m,t)}\mathbb{E}\Big(\left|\bar{\chi}(m,3)_i^n\right|\Big)\le\lim_{m\to\infty}\limsup_{n\to\infty}Kt\gamma_m\Delta_n^{(1-\rho)+(1-\varpi\beta)}=0.
$$

Finally, note that all the evaluations in (D.6), (D.7) and (D.8) remain valid. Therefore, we obtain

$$
\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\sum_i\mathbb{E}\big(\left|\bar{\chi}(m,2)_i^n-\bar{\chi}(m,4)_i^n\right|\big)\le Kt\gamma_m\Delta_n^{1-\rho-\varpi\beta+\varpi(q-\beta)},
$$

$$
\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\sum_i\mathbb{E}\big(\left|\bar{\chi}(m,4)_i^n-\bar{\chi}(m,5)_i^n\right|\big)\le Kt\gamma_m\Delta_n^{1-\rho-\varpi\beta+(q_c'-1+\varpi(2q_d'-\beta))/2\wedge\varpi(q_c'+q_d'-\beta)},
$$

$$
\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}\sum_i\mathbb{E}\big(\left|\bar{\chi}(m,5)_i^n\right|\big)\le Kt\gamma_m\Delta_n^{1-\rho-\varpi\beta+\varpi(q\wedge(q_c'+q_d')-\beta)},
$$

where the sum $\sum_i$ denotes $\sum_{i\in J'(n,m,t)}$. Under condition (b) of Theorem 3, the above three terms are all asymptotically negligible. This completes the proof of Theorem 3.

D.4 Feasible estimators with $\widehat{\beta}$

Note that H is a continuous function of the spot intensities, hence it is also a continuous function with respect to β. Then following the same argument as in the proof of Lemma C.7, one can show that

$$
\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\widehat{\beta}}}}\Big(H(\widehat{\lambda}(\widehat{\beta}))-H(\lambda)\Big)-\sqrt{\frac{k_n\Delta_n}{\Delta_n^{\varpi\beta}}}\Big(H(\widehat{\lambda}(\beta))-H(\lambda)\Big)=o_P(1),
$$

where we omit other arguments of H for simplicity. Therefore, the same conclusions hold for the feasible estimator as well.

References

Aït-Sahalia, Y., Fan, J., Laeven, R. J., Wang, C. D., and Yang, X. (2016). Estimation of the continuous and discontinuous leverage effects. *Journal of the American Statistical Association, forthcoming.*

Aït-Sahalia, Y. and Jacod, J. (2009). Estimating the degree of activity of jumps in high frequency data. *The Annals of Statistics*, 37:2202–2244.

Andersen, T. G., Fusari, N., and Todorov, V. (2016a). The pricing of tail risk and the equity premium: Evidence from international option markets. Working paper.

Andersen, T. G., Fusari, N., and Todorov, V. (2016b). Short-term market risk implied by weekly options. *The Journal of Finance, forthcoming.*

Jacod, J. and Protter, P. (2011). *Discretization of Processes*. Springer.

Jacod, J. and Todorov, V. (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *The Annals of Statistics*, 37:1792–1838.

Jacod, J. and Todorov, V. (2010). Do price and volatility jump together? *Annals of Applied Probability*, 20(4):1425–1469.

Jing, B.-Y., Kong, X.-B., Liu, Z., and Mykland, P. (2012). On the jump activity index for semimartingales. *Journal of Econometrics*, 166(2):213–223.

Li, J., Todorov, V., and Tauchen, G. (2017). Jump regressions. *Econometrica*, 85(1):173–195.