Testing for Jump Spillovers Without Testing for Jumps

Valentina Corradi

University of Surrey
Guilford GU2 7XH, Surrey
United Kingdom
`v.corradi@surrey.ac.uk`

Walter Distaso

Imperial College Business School
South Kensington, London SW7 2AZ
United Kingdom
`w.distaso@imperial.ac.uk`

Marcelo Fernandes

Sao Paulo School of Economics, FGV
01322-000 Sao Paulo, SP
Brazil
`marcelo.fernandes@fgv.br`

Authors' Footnote

Valentina Corradi is Professor of Econometrics at the University of Surrey, Guilford GU2 7XH, Surrey, UK (e-mail: v.corradi@surrey.ac.uk); Walter Distaso is Professor of Financial Econometrics at the Imperial College Business School, London SW7 2AZ, UK (e-mail: w.distaso@qmul.ac.uk); and Marcelo Fernandes is Professor of Finance at the Sao Paulo School of Economics, FGV, 01456-000 Sao Paulo, SP, Brazil (e-mail: marcelo.fernandes@fgv.br). This work was partially supported by the ESRC (RES-062-23-0311), FAPESP (2013/22930-0), and CNPq (302272/2014-3). The authors thank the associate editor and anonymous referee as well as Yacine Aït-Sahalia, Dante Amengual, Jesus Gonzalo, Yongmiao Hong, Aureo de Paula, Bradley Paye, Enrique Sentana, and Stefan Sperlich for valuable comments. They are indebted to seminar and conference participants at Carlos III, CEMFI, Essex Business School, ETH Zurich, Frontiers of Finance (Warwick Business School, 2012), SKBI Annual Conference on Financial Econometrics (Singapore, 2012), Symposium of Econometric Theory and Applications (Shanghai, 2012), Fifth Italian Congress of Econometrics and Empirical Economics (Genova, 2013), International Conference on Systemic Risk, Contagion and Jumps (Cass Business School, 2013), Meetings of the Brazilian Finance Society (Recife, 2014), New Methods for the Empirical Analysis of Financial Markets (Santander Financial Institute, 2017), Southampton Finance and Econometrics Workshop (University of Southampton, 2017), Third International Workshop on Financial Econometrics (Arraial d'Ajuda, 2017), and SoFiE annual conference (Lugano, 2018).

Testing for Jump Spillovers Without Testing for Jumps

Abstract

This paper develops statistical tools for testing conditional independence among the jump components of the daily quadratic variation, which we estimate using intraday data. To avoid sequential bias distortion, we do not pretest for the presence of jumps. If the null is true, our test statistic based on daily integrated jumps weakly converges to a Gaussian random variable if both assets have jumps. If instead at least one asset has no jumps, then the statistic approaches zero in probability. We show how to compute asymptotically valid bootstrap-based critical values that result in a consistent test with asymptotic size equal to or smaller than the nominal size. Empirically, we study jump linkages between US futures and equity index markets. We find not only strong evidence of jump cross-excitation between the SPDR exchange-traded fund and E-mini futures on the S&P 500 index, but also that integrated jumps in the E-mini futures during the overnight period carry relevant information.

Keywords

conditional independence, jump intensity, kernel smoothing, quadratic variation, realized measure

1 Introduction

Introducing jump spillovers is an effective means to model systemic risk and, accordingly, financial contagion. Jump spillovers may generate asymmetric dependence across securities as well as a diversification breakdown. In the event that a downward jump occurs, negative returns spread across markets, implying a higher correlation across a large number of assets in bear markets (Das and Uppal, 2004). Due to the high correlation, systemic risk not only reduces the benefits of diversification, but also increases the likelihood of larger losses for leveraged portfolios. Aït-Sahalia, Cacho-Diaz and Hurd (2009a) study portfolio choice and diversification in the presence of simultaneous common jumps (or, for short, co-jumps). They show that the gain from diversification breaks down, and the optimal portfolio offers as much protection against common jumps as a nondiversified portfolio. Although co-jumping captures cross-sectional dependence across markets, it does not explain jumps clustering (see, e.g., Maehu and McCurdy, 2004). A natural way to capture both stylized facts is to model the jump intensity as a function of past jumps in the asset itself (self-excitation) and/or in other assets (cross-excitation). Aït-Sahalia, Cacho-Diaz and Laeven (2015) extends, for instance, Bowsher’s (2007) Hawkes jump-diffusion model to accommodate both self- and cross-excitation in the jump processes.

This paper develops statistical tools for testing for jump spillovers across assets/markets. We estimate jump contributions to the quadratic variation using realized measures based on high-frequency data. This amounts to model-free estimation of the jump component as it does not require functional form assumptions either on the continuous component of the process (drift and variance) or on the jump component of the process (intensity and jump size distribution). We are also agnostic about whether volatility is stochastic or a function of past asset prices, allowing us to easily accommodate both the affine jump-diffusion model of Duffie, Pan and Singleton (2000) and the Hawkes jump-diffusion specification of Aït-Sahalia et al. (2015). In addition, we ensure that our realized jump estimators are robust to the presence of microstructure noise through the use of pre-averaging methods (Podolskij and Vetter, 2009; Hautsch and Podolskij, 2013).

Our testing procedure focuses on the marginal relevance of conditioning the jump distribution of asset A on (contemporaneous or past) jumps in asset B after controlling for its own past quadratic variation components. The asymptotic theory proceeds in two steps. First, we derive the limiting distribution of the infeasible statistic based on the unobservable quadratic variation components. This is nontrivial because the jump contribution to the quadratic variation is a random variable that takes value zero with positive

probability. In particular, the realized jumps that appear in the argument and in the conditioning set of the conditional distributions are censored from below at zero. This differs not only from Tobit-type nonparametric regressions, in which censoring affects only the dependent variable (e.g., Chen, Dahl and Khan, 2005), but also from nonparametric regressions with mixed continuous and categorical conditioning variables (e.g., Li and Racine, 2008). In the second step, we provide a set of sufficient conditions under which the feasible statistic based on noisy measures of the jump component and of the integrated variance is asymptotically equivalent to its infeasible counterpart. Simulations show that a bootstrap-based implementation of our testing procedures not only exhibits virtually no size distortion, but also entails excellent power against jump spillovers.

A very nice feature of our test for jump transmission is that it does not require pre-testing for jumps. In the absence of jumps in at least one asset, the test statistic automatically converges in probability to zero, ensuring the nonrejection of the null of no jump spillovers. Simulations confirm that this property holds even for relatively small samples. This is extremely convenient for two reasons. First, we do not have to deal with misclassification errors that do not shrink to zero as the sample size increases (see discussions in Lee and Mykland, 2008; Bajgrowicz, Scaillet and Treccani, 2016). Second, standard nonparametric tests for jumps (e.g., Barndorff-Nielsen, Shephard and Winkel, 2006; Lee and Mykland, 2008; Podolskij and Vetter, 2009) aim to learn about the presence of jumps over a finite time span only. The same applies to Dungey, Erdemlioglu, Mateia and Yang's (2017) test for jump cross-excitation. However, the sequential implementation of these tests over rolling time spans induces severe size distortions. Altogether, this means it is possible to circumvent pre-testing issues by using high-frequency data to examine jump transmission at a lower frequency (say, daily).

Note that we depart from the recent literature by focusing on daily jump spillovers. Most papers restrict attention to simultaneous common jumps, whereas we entertain more general forms of jump spillovers, but at a lower frequency. Bollerslev, Law and Tauchen (2008) identify co-jumps by looking at the cross-covariance of asset returns implied by the jump realizations in the individual assets and in their equal-weighted index. Jacod and Todorov (2009) develop tests for the null hypotheses of common and disjoint jumps, identifying as a jump any price movement whose magnitude exceeds a given truncation level that shrinks to zero at a certain known rate. Jacod and Todorov (2010) extend their approach to examine whether prices and volatility jump together.

Mancini and Gobbi (2012) and Aït-Sahalia and Xiu (2016) develop similar frameworks to separate continuous correlations from co-jumps using realized measures of the continuous and discontinuous components of the quadratic variation. The latter authors examine whether the recent increase in the correlation among asset prices is due to increases in the Brownian correlation and/or in jump comovements. They find that both components are relevant, though with co-jumps contributing relatively less. They also show by means of linear regressions that macroeconomic news announcements indeed drive (co-)jumps even after controlling for the continuous component of the quadratic (co-)variation.¹ All of the above papers restrict attention to the simultaneous arrival of jumps, and so where they originate is not an issue. Dungey et al. (2017) take a different approach by studying jump transmission across assets/markets at the high frequency. To this end, they extend Boswijk, Laeven and Yang's (2018) test for jump self-excitation to a bivariate setting in order to test for jump cross-excitation in high frequency data. They identify cross-excitation through common jumps between asset price A and the jump intensity in asset B 's price process. In this paper, we consider jump spillovers at a lower frequency by focusing on contemporaneous integrated jumps over a day as well as on lagged jump spillovers from one asset/market to another.

As for our empirical contribution, we test for jump spillovers between the SPDR exchange-traded fund (ETF) and E-mini futures on the S&P 500 index. The literature argues that derivatives markets should lead the price discovery relative to cash markets because they allow informed traders to get more leverage and trade at lower costs (see, among others, Kawaller, Koch and Kock, 1987; Stoll and Whaley, 1990; Fleming, Ost diek and Whaley, 1996).² However, it is not clear whether futures markets offer lower trading costs than ETF platforms (Wermers and Xue, 2015; Aldrich and Lee, 2016; Ben-David, Franzoni and Moussawi, 2017). For instance, Hasbrouck (2003) finds that ETFs lead futures for the S&P 400 midcap index, whereas E-mini futures dominate the price discovery for both the S&P 500 and Nasdaq 100 indices relative to regular futures and ETFs. Using more recent intraday data, Wallace, Kalev and Lian (2014) shows that E-mini futures indeed lead the information flow, but only until January 2006. The SPDR exchange-traded fund then becomes as relevant as E-mini futures, eventually driving most of the price discovery after the subprime crisis. They nonetheless show that E-mini futures are more informative than the ETF at times of high expected volatility. This is in line with the stylized fact that ETF liquidity drastically reduces at times of

¹ See the recent works by Li, Todorov and Tauchen (2017a,b) and Li, Todorov, Tauchen and Chen (2017c) for a more formal framework to handle jump regressions.

² Stephan and Whaley (1990), Easley, O'Hara and Srinivas (1998) and Chakravarty, Gulen and Mayhew (2004) offer some counterexamples in stock option markets and Yang (2009) for currency markets, though.

market stress, causing price discovery to move away from ETF markets (Ben-David et al., 2017).

Not surprisingly, we unveil strong evidence of contemporaneous integrated jump spillovers between the SPDR exchange-traded fund and E-mini futures. We also find that jumps in the ETF today help excite jumps in E-mini futures tomorrow (regardless of whether we control for the past realized variance or not). According to Aldrich and Lee (2016), this means that, even if cross-market arbitrage opportunities were on average in favor of E-mini futures market makers due to their larger minimum price increments, they were not enough to countervail the presumably lower trading costs of ETFs. Further, conditioning on the jump contribution to the quadratic variation of the E-mini futures during the overnight period seems very informative in view that the jump components during NYSE trading hours in both assets carry very similar information. For instance, test results indicate the presence of jump spillovers from after-hours E-mini futures to the exchange-traded fund, especially for more recent samples.

The rest of this paper ensues as follows. Section 2 illustrates the channels through which jump spillovers may arise using a simple bivariate jump-diffusion example. Section 3 first discusses the null hypothesis of no jump spillovers and then establish the limiting distribution of the infeasible statistic. Section 4 derives the conditions under the feasible and infeasible statistics are asymptotically equivalent. Section 5 shows how to compute asymptotic-valid critical values via bootstrap, whereas Section 6 assesses size and power of the resulting bootstrap-based test through Monte Carlo experiments. Section 7 investigates jump spillovers in the US equity index markets. Section 8 offers some concluding remarks. We collect all technical proofs in the Appendix.

2 Jump transmission: Setup

In this section, we discuss how to analyze jump spillovers through a nonparametric test of conditional independence. For notational simplicity, we restrict attention to the case of two assets, say, A and B . It is straightforward to consider more than two assets, though the usual concern with the curse of dimensionality applies.

We start with a simple example in order to outline the channels through which price jumps in A might affect the jump component in B . As customary in financial economics, we assume that asset prices (in logs) follow a jump-diffusion process:

$$\begin{pmatrix} dp_{A,s} \\ dp_{B,s} \end{pmatrix} = \begin{pmatrix} \mu_{A,s} \\ \mu_{B,s} \end{pmatrix} ds + \begin{pmatrix} \sigma_{AA,s} & \sigma_{AB,s} \\ \sigma_{BA,s} & \sigma_{BB,s} \end{pmatrix} \begin{pmatrix} dW_{A,s} \\ dW_{B,s} \end{pmatrix} + \begin{pmatrix} \kappa_{AA,s} & \kappa_{AB,s} \\ \kappa_{BA,s} & \kappa_{BB,s} \end{pmatrix} \begin{pmatrix} dJ_{A,s} \\ dJ_{B,s} \end{pmatrix}, \quad (1)$$

where $(W_{A,s}, W_{B,s})$ are independent standard Brownian motions, $(\mu_{A,s}, \mu_{B,s})$ are predictable drift processes, and the volatility and cross-volatility components follow a multivariate *càdlàg* process regardless of whether it is stochastic or a measurable function of asset prices.

As for the jump component, $J_{A,s}$ and $J_{B,s}$ are Poisson processes with possibly time-varying intensity. In particular, $\kappa_{Aj,s} = \Delta p_{A,s} \mathbf{1}(dJ_{j,s} = 1)$ with $\Delta p_{A,s} = p_{A,s} - p_{A,s-}$ and $\kappa_{Bj,s} = \Delta p_{B,s} \mathbf{1}(dJ_{j,s} = 1)$ with $\Delta p_{B,s} = p_{B,s} - p_{B,s-}$ correspond respectively to the sizes of the price jumps in assets A and B as the Poisson process $J_{j,s}$ jumps at time s . We thus allow for a different jump size depending on which Poisson process hits the asset price (see, e.g., Chapter 5 in Cont and Tankov, 2004). Finally, $\Pr(dJ_{j,s} = 1 \mid \mathcal{F}_s) = d\lambda_{j,s}$, where $\mathcal{F}_s$ is the filtration at time s and $\lambda_{j,s}$ is the jump intensity for asset $j \in \{A, B\}$.

It is natural to decompose the quadratic variation process $[\cdot, \cdot]_t$ of a given asset price, say p_A , over the time interval $[t-1, t]$ into the part due to the discontinuous jump component $p_A^{(d)}$ and the part due to the continuous diffusive component $p_A^{(c)}$. In particular, $[p_A, p_A]_t = IV_{A,t} + A_t$, where $IV_{A,t} = [p_A^{(c)}, p_A^{(c)}] \equiv \int_{t-1}^t \sigma_{A,s}^2 ds + \int_{t-1}^t \sigma_{AB,s}^2 ds$ corresponds to the integrated variance over the time interval $[t-1, t]$ and $A_t = [p_A^{(d)}, p_A^{(d)}] \equiv \sum_{t-1 \leq s \leq t} \Delta p_{A,s}^2$. It also follows from (1) that

$$\sum_{t-1 \leq s \leq t} \Delta p_{A,s}^2 = \sum_{s=J_{A,t-1}}^{J_{A,t}} \kappa_{AA,s}^2 + \sum_{s=J_{B,t-1}}^{J_{B,t}} \kappa_{AB,s}^2 \quad (2)$$

$$\sum_{t-1 \leq s \leq t} \Delta p_{B,s}^2 = \sum_{s=J_{A,t-1}}^{J_{A,t}} \kappa_{BA,s}^2 + \sum_{s=J_{B,t-1}}^{J_{B,t}} \kappa_{BB,s}^2. \quad (3)$$

As Poisson processes are finite activity processes, in the absence of perfect correlation between $J_{A,t}$ and $J_{B,t}$, the probability that they jump together over a finite time span is zero and hence the cross-term component $\sum_{t-1 \leq s \leq t} \kappa_{BA,s} \kappa_{BB,s} \mathbf{1}(dJ_{A,s} dJ_{B,s} = 1)$ is negligible.

It is easy to appreciate from (2) and (3) that A_t does not depend on $B_\tau = [p_B^{(d)}, p_B^{(d)}]_\tau$ for any $\tau \leq t$ if and only if

- (i) $\kappa_{AB,s} = \kappa_{BA,s} = 0$ almost surely;
- (ii) $\kappa_{AA,s}$ and $\kappa_{BB,s}$ are independent; and
- (iii) $J_{A,s}$ and $J_{B,s}$ are independent.

Note that there would exist only common simultaneous jumps (or co-jumps) if only (i) fails to hold in view that a jump in either $dJ_{A,s}$ or $dJ_{B,s}$ would culminate in simultaneous jumps in both asset prices p_A and p_B . This would ultimately result in a small number of relatively large co-jumps in the data due to the

finite variation property of Poisson processes. To reconcile with Bollerslev et al.'s (2008) empirical evidence of a large number of small common simultaneous jumps among stock returns, it would suffice to replace Poisson processes with more general Lévy processes so as to allow for infinitely many small co-jumps. We consider here Poisson jumps only for ease of exposition. The realized measures we employ to estimate the jump component of the quadratic variation are actually consistent even under infinite variation.

If instead only **(iii)** fails, no simultaneous common jumps would come about, though a feedback effect would still arise given the mutual dependence between $J_{A,s}$ and $J_{B,s}$. In particular, the link is exclusively contemporaneous if both $J_{A,s}$ and $J_{B,s}$ have constant intensity in that A_t is independent of B_τ for all $\tau < t$ even if **(iii)** does not apply. Now, if the intensity processes are measurable functions of some common serial dependent process, then $\Delta J_{A,s}$ may also depend on $\Delta J_{B,\tau}$ for $\tau < t$. Examples include Duffie et al.'s (2000) affine jump diffusions, for which

$$\begin{pmatrix} \lambda_{A,s} \\ \lambda_{B,s} \end{pmatrix} = \begin{pmatrix} \lambda_A^0 \\ \lambda_B^0 \end{pmatrix} + \begin{pmatrix} \lambda_{AA}^1 & \lambda_{AB}^1 \\ \lambda_{BA}^1 & \lambda_{BB}^1 \end{pmatrix} \begin{pmatrix} p_{A,s} \\ p_{B,s} \end{pmatrix},$$

as well as Aït-Sahalia et al.'s (2015) Hawkes jump-diffusion model with intensity processes given by

$$\begin{pmatrix} \lambda_{A,s} \\ \lambda_{B,s} \end{pmatrix} = \begin{pmatrix} \lambda_{A,\infty} + \int_0^s \lambda_{AA}(s-r) dJ_{A,r} + \int_0^s \lambda_{AB}(s-r) dJ_{B,r} \\ \lambda_{B,\infty} + \int_0^s \lambda_{BA}(s-r) dJ_{A,r} + \int_0^s \lambda_{BB}(s-r) dJ_{B,r} \end{pmatrix}.$$

In principle, it is possible to test directly whether conditions **(i)** to **(iii)** hold, if one is ready to specify the functional forms of the drift, diffusive, and jump terms. The outcome would however depend heavily on the correct specification of the data generating process. To minimize the risk of misspecification, we resort to a nonparametric approach. In particular, we construct a test for the null hypothesis that **(i)** to **(iii)** hold without imposing any parametric assumption on the jump-diffusion process given by (1).

3 The infeasible statistic

We wish to test whether A_t does not depend on B_t after controlling for the past realizations of the quadratic variance components of the price of asset A . We thus define the larger information set as $\mathbf{X}_t = (\mathbf{X}_{A,t}, B_t)$, whereas the smaller information set $\mathbf{X}_{A,t}$ contains past information exclusively about asset A . We use the notation $\mathbf{X}_{A,t}$ to accommodate dependence either only on the past jump realization or, say, on both components of the quadratic variation: $\mathbf{X}_{A,t} = A_{t-1}$ or $\mathbf{X}_{A,t} = (A_{t-1}, IV_{A,t-1})$, respectively.³ We hereafter

³ We restrict attention to the past components of the quadratic variation of the price of asset A in the smaller information set, but one could certainly think of other conditioning variables. In addition, we control only for the last realization of the quadratic variation components to mitigate dimensionality issues. This is enough to accommodate both Bates's (2000) geometric jump-diffusion process with stochastic volatility and Duffie et al.'s (2000) affine jump-diffusion process, for instance.

denote by q and $q_A = q - 1$ the dimensionality of the larger and smaller information sets, respectively. The null hypothesis is that the conditional distribution of A_t given $\mathbf{X}_t$ is almost surely equal to the conditional distribution given only $\mathbf{X}_{A,t}$, i.e.,

$$\mathbb{H}_0: \Pr(A_t \geq a \mid \mathbf{X}_t = \mathbf{x}) - \Pr(A_t \geq a \mid \mathbf{X}_{A,t} = \mathbf{x}_A) = 0 \quad \text{a.s.} \quad (4)$$

We begin by assuming that we observe the *true* jump contribution to the quadratic variation and hence we may test the null hypothesis $\mathbb{H}_0$ in (4) by means of

$$S_T = h \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t) - \hat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t), \quad (5)$$

where $\pi(\mathbf{x})$ refers to an integrable weighting function that trims away observations out of the compact set $\mathcal{C}_{\mathbf{X}} \subset \{\mathbf{x} = (x_1, \dots, x_q) : x_1 \leq \bar{x}_1, \dots, x_q \leq \bar{x}_q\}$. In turn, $\hat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t)$ and $\hat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t})$ are respectively Nadaraya-Watson estimators of the conditional distributions of A_t given $\mathbf{X}_t$ and $\mathbf{X}_{A,t}$:

$$\begin{aligned} \hat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t) &= \frac{\frac{1}{T} \sum_{s=1}^T \mathbf{1}\{A_s \leq A_t\} \prod_{j=1}^q K_h(X_{j,t} - X_{j,s})}{\frac{1}{T} \sum_{s=1}^T \prod_{j=1}^q K_h(X_{j,t} - X_{j,s})} \\ \hat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_{A,t}) &= \frac{\frac{1}{T} \sum_{s=1}^T \mathbf{1}\{A_s \leq A_t\} \prod_{j=1}^{q_A} K_b(X_{j,t} - X_{j,s})}{\frac{1}{T} \sum_{s=1}^T \prod_{j=1}^{q_A} K_b(X_{j,t} - X_{j,s})}, \end{aligned}$$

with $K_\zeta(\cdot) = \frac{1}{\zeta} K(\cdot/\zeta)$ for $\zeta = b, h$.

Two remarks are in order. First, we employ different bandwidths for the estimation of the two distribution functions in order to rule out a bias term that diverges to minus infinity in Theorem 2. Second, we do not trim the estimation of the conditional distribution from below. This means that the statistic considers every zero value in the sample. This is important as in practice we observe only a noisy version of the asset prices $p_{A,t}$ and $p_{B,t}$, implying the (spurious or not) absence of zeroes. Accordingly, trimming away observations smaller than a threshold would induce unnecessary arbitrariness to the testing procedure.

Our test statistic S_T in (5) is very similar to the statistic T_3 in Aït-Sahalia, Fan and Peng's (2009b) equation (4.4), differing only in that their restricted model is parametric, while ours is also nonparametric, so that two additional bias terms arise. If $\inf_t A_t > 0$ and $\inf_t B_t > 0$ as in Lemma 1A, then our asymptotic theory would follow along similar lines to the proof of their Theorem 2. In addition, we allow for the possibility of no jumps in either assets. In fact, from a technical point of view, our main contribution is to show that, in the absence of jumps in either assets, the statistic shrinks to zero in probability, not rejecting the null hypothesis.

In the sequel we rely on the following assumptions.

Assumption A1: The kernel function K is of order 2, symmetric, nonnegative, at least twice differentiable on the interior of its bounded support, and $K(0) = C$ with $0 < \underline{c} \leq C \leq \bar{c} < \infty$.

Assumption A2: The distribution functions $F_{A|\mathbf{X}}(a|\mathbf{x})$ and $F_{\mathbf{X}}(\mathbf{x})$ are r -times continuously differentiable in $(a, \mathbf{x}) \in \mathcal{C}_{A,\mathbf{X}}$ with bounded derivatives and with $r \geq 2$. The same condition holds for the lower-dimensional distribution function $F_{A|\mathbf{X}_A}(a|\mathbf{x}_A)$. The density function $f_{\mathbf{X}}(\mathbf{x})$ is bounded away from zero for any $\mathbf{x} \in \mathcal{C}_{\mathbf{X}}$.

Assumption A3: The weighting function $\pi(\mathbf{x})$ is continuous and integrable, with second derivatives in a compact support.

Assumption A4: The stochastic process $(A_t, \mathbf{X}_t)$ is strictly stationary and β -mixing with $\beta_k = O(\rho^k)$, where $0 < \rho < 1$.

Assumption A5: The bandwidths are such that (i) $Th^{q/2+4} \rightarrow 0$, (ii) $Th^{3q/2} \rightarrow \infty$, (iii) $Th^{q/2}b^4 \rightarrow 0$, (iv) $Th^{-q}b^{5q_A/2} \rightarrow \infty$, (v) $hb^{-1} \rightarrow \infty$, (vi) $h^qb^{-q_A} \rightarrow 0$.

Assumption A1 holds for most second-order kernels, such as the Epanechnikov, Parzen, and quartic kernels. We rule out higher-order kernel to ensure the positivity of the conditional distribution estimator. Also, several high-order kernels violate the condition $K(0) = C$, which is crucial to control the behavior of statistic in the absence of jumps in at least one asset. Assumptions A2 and A3 require that the distribution and weighting functions are both well defined and smooth enough to admit functional expansions. Assumption A4 restricts the amount of data dependence by imposing absolute regularity with geometric decay rate, but allowing for both path dependence in the intensity as well as in the jump size. This for example occurs if some of the parameters governing the jump size depend on a state variable that displays time dependence. Assumption A5 states sufficient conditions for the bandwidth rates: (i) to (iv) ensure asymptotic normality in the presence of jumps in both asset, whilst (v) and (vi) guarantee that the statistic does not diverge in the absence of jumps in at least one asset. Note that, given (v), (iii) follows immediately from (i), whereas (iv) and (vi) imply $Th^{-q/2}b^{2q_A} \rightarrow \infty$. We make certain Assumption A5 holds by considering $O(T^{-1/(q/2+4)}) < h < O(T^{-2/(2q)})$ and $O(h) < b < O(h^\gamma)$, with $1 < \gamma < q/q_A$.

We are now ready to establish the limiting distribution of the test statistic in (5). Let hereafter $I_{11,t} = \mathbf{1}(A_{t-1} > 0) \mathbf{1}(B_t > 0)$, $I_{10,t} = \mathbf{1}(A_{t-1} > 0) \mathbf{1}(B_t = 0)$, $I_{01,t} = \mathbf{1}(A_{t-1} = 0) \mathbf{1}(B_t > 0)$, $I_{00,t} =$

$\mathbf{1}(A_{t-1} = 0) \mathbf{1}(B_t = 0)$, and let $T_{ij} = \sum_{t=1}^T I_{ij,t}$ for $i, j \in \{0, 1\}$.

Theorem 1: *Let Assumptions A1-A5 hold. If $T_{11}/T \xrightarrow{p} c_{11} > 0$, then*

$$S_T - h^{-q/2}\mu_1 - h^{q/2}b^{-q_A}\mu_2 + 2h^{1-q/2}\mu_3 \xrightarrow{d} N(0, \sigma^2) \quad \text{under } \mathbb{H}_0,$$

where $\sigma^2 = \frac{1}{45} \int \left(\int \mathbf{K}(\mathbf{u}) \mathbf{K}(\mathbf{u} - \mathbf{v}) d\mathbf{u} \right)^2 d\mathbf{v} \int_{\mathbf{x} > 0} [\pi(\mathbf{x})]^2 d\mathbf{x}$ and the bias terms (μ_1, μ_2, μ_3) are as in the Appendix. In addition, there exists $\varepsilon > 0$ such that $\Pr(c_{11}^{-1} T^{-1} h^{-q/2} S_T > \varepsilon) \rightarrow 1$ under $\mathbb{H}_A$.

Theorem 1 establishes that, if the fraction of days in which both assets display jumps grows at the same rate as the sample size, then the statistic has a standard normal limiting distribution under the null and diverges under the alternative. As shown in Lemmata 1A to 4A in the Appendix, the limiting distribution of the statistic depends on the subset of the sample over which both asset prices display a strictly positive jump component. On the other hand, whenever the statistic is computed over a subset of the sample in which at least one asset does not display jumps, it shrinks to zero in probability.

We next deal with the case in which at least one asset features no price jumps. In particular, we show that the statistic approaches zero in probability and hence we end up not rejecting the null. Needless to say, this situation would never arise if we could observe the true jump component. However, as it will become clearer in the next section, we observe only a realized measure of the jump contribution to the quadratic variation, which is not necessarily equal to zero in the absence of jumps.

Theorem 2: *Let Assumptions A1-A5 hold.*

- (i) *If $A_t = 0$ almost surely for all t , then $S_T = 0$ almost surely.*
- (ii) *If $B_t = 0$ almost surely for all t , then $S_T = O_p(h^{1/2} + h^{q/2}b^{-q_A/2}) + (h^{q/2}b^{-q_A} - h^{(1-q)/2})\mu_1^{(2)}$.*

The analytical expression for $\mu_1^{(2)}$ is given in the Appendix. It is worth emphasizing that, in the absence of jumps in both assets, $S_T = 0$ almost surely by Theorem 2(i).

In practice, we do not know whether $T_{11}/T \xrightarrow{p} c_{11} > 0$ as in Theorem 1 or $T_{11}/T \xrightarrow{p} 0$ as in Theorem 2. This means we cannot simply derive asymptotic critical values for S_T based on Theorem 1. We nonetheless show in Section 5 how to derive bootstrap-based critical values that give way to a consistent test with asymptotic size equal either to α if $T_{11}/T \xrightarrow{p} c_{11} > 0$ or to zero if $T_{11}/T \xrightarrow{p} 0$. One additional advantage of bootstrapping is that it automatically accounts for the bias terms without requiring their estimation as long as we use the same bandwidth for the original and bootstrap statistics.

4 The feasible statistic

The statistic S_T is infeasible as we do not observe A_t , $IV_{A,t}$ and B_t . However, in the presence of intraday observations, we can consistently estimate the components of the quadratic variation. More precisely, given a sample of M intraday observations over a time span of T days, we denote by $A_{M,t}$ and $B_{M,t}$ the realized measures for the jump contribution to the quadratic variation at day t , and by $RV_{A,t,M}$ the realized measure of the integrated variance. We next derive the conditions under which the feasible statistic resting on observable realized measures $(A_{M,t}, RV_{A,t,M}, B_{M,t})$ is asymptotically equivalent to its infeasible counterpart. We also show that the contributions of the measurement errors are still of smaller probability order even if the statistic approaches zero in probability due to the absence of jumps in at least one asset.

Given the presence of measurement error in financial transaction data due to market microstructure noise, we employ Podolskij and Vetter's (2009) realized measure of the jump contribution to the quadratic variation of the process. Their estimator measures the difference between two realized measures. The first is consistent for the total quadratic variation, whereas the second consistently estimates the integrated variance of the process. This is well in line with the literature dealing with testing for jumps and with the estimation of the degree of jump activity (see, e.g., Huang and Tauchen, 2005; Barndorff-Nielsen et al., 2006; Aït-Sahalia and Jacod, 2009b; Cont and Mancini, 2011; Todorov and Tauchen, 2011). The only difference is that Podolskij and Vetter's (2009) realized measure of the jump contribution is robust to the presence of market-microstructure noise due to a pre-averaging procedure.

Let k_M denote a deterministic sequence such that $\frac{k_M}{\sqrt{M}} = \theta + o(M^{-1/4})$ and let g denote a continuous and piecewise differentiable function with piecewise Lipschitz derivative such that $g(0) = g(1) = 0$ and $\int_0^1 g^2(s) ds < \infty$. Typical examples are $g(u) = \min\{u, 1 - u\}$ and $g(u) = u(1 - u^2)\mathbf{1}(0 \leq u \leq 1)$. Define now the market prices of assets A and B at time $t + \ell/M$ respectively as $Z_{A,t+\ell/M} = p_{A,t+\ell/M} + \epsilon_{A,t+\ell/M}$ and $Z_{B,t+\ell/M} = p_{B,t+\ell/M} + \epsilon_{B,t+\ell/M}$, where $p_{j,t+\ell/M}$ and $\epsilon_{j,t+\ell/M}$ denote the efficient price and additive microstructure noise for asset $j \in \{A, B\}$. As in Podolskij and Vetter (2009), we proxy the jump component in the quadratic variation of asset A by means of

$$A_{t,M} = \frac{PV_{M,t}^{(A)}(2, 0) - \mu_{|\Phi|}^{-p} PV_{M,t}^{(A)}(2/p, \dots, 2/p)}{\theta \int_0^1 g^2(s) ds}, \quad (6)$$

where $\mu_{|\Phi|}$ is the first absolute moment of a standard normal distribution,

$$PV_{M,t}^{(A)}(2/p, \dots, 2/p) = \frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \left| \sum_{\ell=1}^{k_M} g(t + \ell/M) (Z_{A,t+(j+ik_M+\ell)/M} - Z_{A,t+(j+ik_M+\ell-1)/M}) \right|^{\frac{2}{p}} \quad (7)$$

is the pre-averaging multipower variation, and

$$PV_{M,t}^{(A)}(2, 0) = \frac{1}{\sqrt{M}} \sum_{j=1}^{M-2k_M+1} \left| \sum_{\ell=1}^{k_M} g(t + \ell/M) (Z_{A,t+(j+\ell)/M} - Z_{A,t+(j+\ell-1)/M}) \right|^2 \quad (8)$$

is the pre-averaging realized variance measure of Jacod, Li, Mykland, Podolskij and Vetter (2009). Finally, let $RV_{A,t,M} = PV_{M,t}^{(A)}(2/p, \dots, 2/p)$ and define $B_{t,M}$ as in (6), but substituting $PV_{M,t}^{(B)}(2, 0)$ and $PV_{M,t}^{(B)}(2/p, \dots, 2/p)$ for $PV_{M,t}^{(A)}(2, 0)$ and $PV_{M,t}^{(A)}(2/p, \dots, 2/p)$, respectively.

In the sequel, let $g(\ell/M) = \min\{\ell/M, (1 - \ell/M)\}$ and $\mathbf{X}_{t,M} = (\mathbf{X}_{A,t,M}, B_{t,M})$. Define the feasible statistic as

$$S_{T,M} = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X},M}(A_{t,M}|\mathbf{X}_{t,M}) - \widehat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_{t,M}) \quad (9)$$

where $\widehat{F}_{A|\mathbf{X},M}$ and $\widehat{F}_{A|\mathbf{X}_A,M}$ differ from $\widehat{F}_{A|\mathbf{X}}$ and $\widehat{F}_{A|\mathbf{X}_A}$ only for employing realized measures (rather than true values) of the quadratic variation components. To establish asymptotic equivalence between the infeasible and feasible statistics, we require some additional assumptions.

Assumption A6: The drift terms in (1) are continuous locally bounded processes with $\mathbb{E}|\mu_{i,t}|^{2k} < \infty$, whereas the diffusive functions are càdlàg with $\mathbb{E}(\sigma_{ij,t}^{2k}) < \infty$ for $k \geq 2$ and the jump components $\kappa_{ij,t}$ are iid with all finite moments for $i, j \in \{A, B\}$.

Assumption A7: The microstructure noises $\epsilon_{A,t}$ and $\epsilon_{B,t}$ are iid with symmetric distribution around zero and such that $\mathbb{E}(\epsilon_{A,t}^{2k}) < \infty$ and $\mathbb{E}(\epsilon_{B,t}^{2k}) < \infty$ for some $k \geq 2$.

Assumption A8: The jump components have a smaller-than-one Blumenthal-Gettoor index.⁴

The next result shows that the asymptotic equivalence between infeasible and feasible test statistics requires that the number of intraday observations M grows fast enough relative to the number of days T . This results in the usual tradeoff of whether using a nonrobust realized measure with $a_M = M$ at a frequency for which microstructure noise is negligible or a microstructure-robust realized measure with $a_M = \sqrt{M}$ at the highest available frequency. Note that, although we may observe negative values for

⁴ See, for instance, Aït-Sahalia and Jacod (2009b) for a formal definition.

$A_{t,M}$ and $B_{t,M}$, they are at most of probability order $a_M^{-1/2}$ and thus asymptotically absent. As we do not trim away zero values from the infeasible statistic, we should not trim away negative values from the feasible one. In fact, we provide below conditions ensuring that, whenever the infeasible statistic is $o_p(1)$ due to the absence of jumps in either asset, the contribution of the measurement error approaches zero at a faster rate than otherwise. The statements in Theorems 3 and 4 rely on the following result.

Lemma 1: *Given Assumptions A6 to A8, $\mathbb{E}[(A_{t,M} - A_t)^k] = a_M^{-k/2}$, $\mathbb{E}[(RV_{A,t,M} - IV_{A,t})^k] = a_M^{-k/2}$, and $\mathbb{E}[(B_{t,M} - B_t)^k] = a_M^{-k/2}$ for all $p/4 < k \leq 2(p-1)$, where p is defined in (6). In addition, $a_M = M$ for $k_M = 1$ in the absence of pre-averaging and $a_M = M^{1/2}$ for $k_M = \theta M^{1/2} + o(M^{1/2})$ in the case of pre-averaging.*

The above result extends the moment conditions on the measurement error in Corradi, Distaso and Fernandes's (2012) Lemma 1 to the case of pre-averaging realized measures. Note that the rate of decay of the measurement error moments depends not only on the moments of the drift, variance and jump sizes in Assumptions A6 and A7, but also on the order of the power variation. It turns out that, other things being equal, k increases with p . This is somewhat intuitive. In the presence of a small number of large jumps (i.e., finite activity jumps), the order of magnitude of $\mathbb{E}(\kappa_{ij,t}^{2k})$ does not decrease with k ; on the other hand, the higher is p the faster the contribution of jumps to the power variation estimator approaches zero. In other words, regardless of pre-averaging, the moments of the difference between the power-variation estimator with and without jumps approaches zero at rate getting faster as p increases. This is shown in detail in the proof of Lemma 1 in the Appendix.

Theorem 3: *Let Assumptions A1 to A8 hold. If $T^{(k+4)/k}(\ln T)a_M^{-1}h^{q/2} \rightarrow 0$ and $(\ln T)a_M^{-1}h^{-q} \rightarrow 0$ as $M, T \rightarrow \infty$, $S_T - S_{T,M} = o_p(1)$ under $\mathbb{H}_0$ and $\Pr(T^{-1}h^{-q/2}S_{T,M} > \varepsilon) \rightarrow 1$ under $\mathbb{H}_A$.*

It follows from Theorem 3 that, if the number of intraday observations grows fast enough relative to the number of days, the difference between the feasible and infeasible statistics goes to zero, regardless the presence of jumps in both assets or not.

Last but not least, it is worth mentioning that pre-averaging is not the unique approach to handle market microstructure noise. There are realized measures based on sparse sampling that are robust to some temporal dependence in the microstructure noise, e.g. realized kernel and multi-scale estimators

(Barndorff-Nielsen, Hansen, Lunde and Shephard, 2008; Aït-Sahalia, Mykland and Zhang, 2011). However, to the best of our knowledge, there are no jump-robust counterparts to these estimators.⁵ In addition, even though Hautsch and Podolskij (2013) provide a version of the pre-average estimator that allows for q -dependent microstructure noise, their jump-robust power-variation measure still requires independent noise. More recently, Li and Patton (2018) revisit forecast evaluation tests using high-frequency data, providing sufficient conditions for the asymptotic equivalence of the infeasible and feasible statistics. The primitive condition on the realized estimation error they impose in equation (3.2) is similar to our moment conditions in Lemma 1. Although they provide examples of realized measures that satisfy their high-level assumption, they do not consider the case of joint presence of jumps and microstructure noise.

5 Bootstrap critical values

It is well known that the *standard* bootstrap fails to mimic the limiting distribution of degenerate U-statistics (see, for instance, Bretagnolle, 1983; Arcones and Giné, 1992). Accordingly, we rely on the m out of n (*moon*) bootstrap, drawing $b_{\mathcal{T}}$ blocks of length $l_{\mathcal{T}}$, with $b_{\mathcal{T}}l_{\mathcal{T}} = \mathcal{T}$ and $\mathcal{T}/T \rightarrow 0$, from $\mathbf{X}_{t,M} = (\mathbf{X}_{A,t,M}, B_{t,M})$.⁶ A natural choice for the bootstrap bandwidth is to set h_* and b_* in such a way that $h_*/\mathcal{T} = h/T$ and $b_*/\mathcal{T} = b/T$. This would automatically satisfy all the rate conditions in Assumption 5 once we replace T with $\mathcal{T}$. Nevertheless, in this case, we would have to estimate the bias terms for both the original statistic and its bootstrap counterpart. This is because, $h_*/h \rightarrow 0$ and $b_*/b \rightarrow 0$ as $T, \mathcal{T} \rightarrow \infty$ and so the bias terms in the original and bootstrap statistics would not offset each other. Moreover, as we do not know whether $T_{11}/T \rightarrow c$ with $c > 0$ or $c = 0$, in general, we do not know whether it is necessary to correct for bias or not.

We instead fix the bootstrap bandwidths to $h_* = h$ and $b_* = b$. As $\mathcal{T}/T \rightarrow 0$, this choice implies a higher degree of oversmoothing for the bootstrap bandwidths, posing a constraint on how fast $\mathcal{T}/T$ shrinks to zero. We next show that, as long as $\mathcal{T} = T^\delta$ with $3q/(q+8) < \delta < 1$, the bandwidth rate conditions in Assumption 5 hold with $\mathcal{T}$ as the number of daily observations in the sample.⁷ In the presence of jumps in both assets ($T_{11}/T \rightarrow c_{11} > 0$), bootstrap-based inference results in a test with correct asymptotic size

⁵ Boudt and Zhang (2015) apply thresholding to robustify the two-scale realized measure, but the resulting estimator yields a not-fast-enough convergence rate of $a_M = M^{-1/3}$.

⁶ One obvious alternative is to employ subsampling, which also boils down to resampling m out of n observations, though without replacement. See Politis, Romano and Wolf (1999) for more details.

⁷ Note that the above condition for δ constrains the dimension of the larger information set to at most 3 conditioning variables. In a previous version of this paper, we had a less restrictive condition for at most two conditioning variables ($q = 2$): namely, $27/40 < \delta < 1$.

and unit power. However, in the absence of jumps in at least one asset, both original and bootstrap-based statistics approach zero in probability at the same speed. As a consequence, we have no longer proper control of the test size. This is the same sort of problem that arises when testing multiple slack moment inequalities and hence we may apply Andrews and Shi's (2013) uniformity correction factor to control size.

We proceed as follows. From $\mathbf{W}_{t,M} = (A_{t,M}, \mathbf{X}_{A,t,M}, B_{t,M})$, we resample $b_{\mathcal{T}}$ blocks of length $l_{\mathcal{T}}$, with $b_{\mathcal{T}}l_{\mathcal{T}} = \mathcal{T} = T^{\delta}$, with $3q/(q+8) < \delta < 1$. The *moon* bootstrap samples are then given by $(\mathbf{W}_1^*, \dots, \mathbf{W}_{\mathcal{T}}^*)$. Using the same bandwidth as in the sample statistic, the feasible bootstrap statistic reads

$$S_{\mathcal{T},M}^* = h^{q/2} \sum_{t=1}^{\mathcal{T}} \left[\hat{F}_{A|\mathbf{X},M}^*(A_{t,M}^* | \mathbf{X}_{t,M}^*) - \hat{F}_{A|\mathbf{X}_A,M}^*(A_{t,M}^* | \mathbf{X}_{A,t,M}^*) \right]^2 \pi(\mathbf{X}_{t,M}^*),$$

where the starred quantities are the bootstrap counterparts employing $(A_{t,M}^*, \mathbf{X}_{A,t,M}^*, B_{t,M}^*)_{t=1}^{\mathcal{T}}$ instead of $(A_{t,M}, \mathbf{X}_{A,t,M}, B_{t,M})_{t=1}^T$. We compute $S_{\mathcal{T},M}^{*(j)}$ for every bootstrap sample $j = 1, \dots, \mathcal{B}$ and then denote the $(1 - \alpha + \eta)$ th percentile of the empirical distribution across the $\mathcal{B}$ bootstrap samples by $c_{T,\mathcal{B},M,1-\alpha+\eta}^*$, with $0 < \eta < \alpha/2$. Let also $\tilde{c}_{T,M,1-\alpha,\eta}^* \equiv \lim_{\mathcal{B} \rightarrow \infty} c_{T,\mathcal{B},M,1-\alpha+\eta}^* + \eta$. The next result establishes the validity of the bootstrap-based critical values.

Theorem 4: *Let Assumptions A1 to A8 hold and, as $T, \mathcal{T}, M \rightarrow \infty$, let $T^{(k+4)/k}(\ln T)a_M^{-1}h^{q/2} \rightarrow 0$ and $(\ln T)a_M^{-1}h^{-q} \rightarrow 0$, as well as $l_{\mathcal{T}} \rightarrow \infty$, $l_{\mathcal{T}}/\sqrt{\mathcal{T}} \rightarrow 0$ and $\mathcal{T} = T^{\delta}$ with $3q/(q+8) < \delta < 1$. It then follows that*

- (i) Under $\mathbb{H}_0$, $\limsup_{M,T \rightarrow \infty} \Pr(S_{T,M} > \tilde{c}_{T,M,1-\alpha,\eta}^*) \leq \alpha$.
- (ii) Under $\mathbb{H}_0$, $\lim_{\eta \rightarrow 0} \limsup_{M,T \rightarrow \infty} \Pr(S_{T,M} > \tilde{c}_{T,M,1-\alpha,\eta}^*) = \alpha$ if $T_{11}/T \xrightarrow{p} c_{11} > 0$.
- (iii) Under $\mathbb{H}_A$, $\lim_{M,T \rightarrow \infty} \Pr(S_{T,M} > \tilde{c}_{T,M,1-\alpha,\eta}^*) = 1$.

It follows from Theorem 4 that the rule of rejecting the null if $S_{T,M} > \tilde{c}_{T,M,1-\alpha,\eta}^*$ provides a test of asymptotic size not larger than α and asymptotic unit power. Borrowing the terminology of Andrews and Shi (2013), the role of the uniformity factor $\eta > 0$ is the following. Suppose none of the assets has jumps. Then, in the absence of measurement error, both S_T^* and S_T are equal to zero. Without correction, we are essentially comparing zero against zero. However, by adding the uniformity factor $\eta > 0$, we ensure that we do not reject the null, regardless of how small we set η . When instead only one asset does not display jumps, then S_T^* and S_T approach zero in probability at the same rate, and the uniformity factor ensures we do not reject the null. As the contribution of the measurement error vanishes, a similar argument applies for the feasible statistics $S_{T,M}^*$ and $S_{T,M}$. On the other hand, in the presence of jumps in both assets (i.e.,

$T_{11}/T \xrightarrow{p} c_{11} > 0$), $S_{T,M}^*$ and S_T have the same limiting distribution under the null and then letting η shrink to zero yields an asymptotically nonconservative test. Finally, under the alternative, it is immediate to see that $S_{T,M}^*$ diverges at a faster rate than $S_{T,M}$ in view that $\delta < 1$.

6 Monte Carlo study

In this section, we run simulations to assess size and power of our testing procedure. For each of the 2,000 replications, we simulate intraday returns from mean-reverting square-root processes (Cox, Ingersoll and Ross, 1985) with Poisson jumps:

$$dP_{At} = \xi_A (\mu_A - P_{At}) dt + \varsigma_A \sqrt{P_{At}} dW_{At} + \kappa_{AA} dJ_{A,t} \quad (10)$$

$$dP_{Bt} = \xi_B (\mu_B - P_{Bt}) dt + \varsigma_B \sqrt{P_{Bt}} dW_{Bt} + \kappa_{BA} dJ_{A,t} + \kappa_{BB} dJ_{B,t}, \quad (11)$$

where W_{At} and W_{Bt} are independent Brownian motions, using an Euler discretization scheme with a reflection device to ensure positivity. To entail realistic asset price processes, we fix the parameter vectors to $(\xi_A, \mu_A, \varsigma_A) = (0.080, 0.150, 0.011)$ and $(\xi_B, \mu_B, \varsigma_B) = (0.120, 0.200, 0.013)$, implying returns with a first-order autocorrelation of slightly above -0.05 in daily returns and average integrated volatility of about 40% per annum.

We assume a constant intensity for $J_{A,t}$ and $J_{B,t}$, with one jump every 5 trading days, on average. As in Barndorff-Nielsen and Shephard (2006), we calibrate the jump sizes to have an order of magnitude similar to the daily implied volatility. In particular, we assume that, under the null hypothesis of no jump spillover, $J_{A,t}$ and $J_{B,t}$ are independent Poisson processes, whereas κ_{jj} has a normal distribution with mean $-\sqrt{\Sigma_{j,t}/2}$ and variance $\Sigma_{j,t}/2$, with $\Sigma_{j,t} = [p_j^{(c)}, p_j^{(c)}]_t$ denoting the daily integrated variance of asset $j = A, B$. Under the alternative, the jump realization in asset B also reflects the jump in asset A , with $\kappa_{BA} \sim \frac{1}{2} N(-\sqrt{\Sigma_{A,t}/2}, \Sigma_{A,t}/2)$ and $\kappa_{BB} \sim \frac{1}{2} N(-\sqrt{\Sigma_{B,t}/2}, \Sigma_{B,t}/2)$.

After burning the first 2,000 observations of the sample, we employ the last MT intraday observations, where M and T correspond respectively to the number of intraday observations within a day and to the number of days. We focus on the relatively small sample sizes of $M \in \{78, 390, 780\}$ and $T \in \{250, 500, 750\}$ so as to assess how important is the condition that requires M to grow at a faster rate than T . If one considers the trading hours of the New York Stock Exchange (NYSE), for instance, our choice of M values corresponds to intraday data at the 5-minute, 1-minute, and 30-second frequencies, whereas our choice of T values represent approximately 1, 2 and 3 years of daily data. To each of the MT intraday observations,

we then add a market microstructure noise to the intraday prices of asset j by randomly sampling from a normal distribution with mean zero and variance $c\Sigma_{j,t}$. We entertain two values for the noise-to-signal ratio: $c = 0.001$ in line with the empirical evidence in Hansen and Lunde (2006) and $c = 0$ so as to address the impact of the market microstructure noise.

From the intraday log-returns, we retrieve the daily components of the quadratic variation $A_{t,M}$, $RV_{A,t,M}$ and $B_{t,M}$ for each day $t = 1, \dots, T$ using a standard pre-averaging procedure, with a flat top kernel and a square root window: i.e., $g(\ell/M) = \min\{\ell/M, (1 - \ell/M)\}$ and $k_M = \sqrt{M}$. We then test for conditional independence between jumps by looking at the squared difference between the conditional distribution of $A_{t,M}$ given $\mathbf{X}_{t,M} = (\mathbf{X}_{A,t,M}, B_{t,M})$ and the conditional distribution of $A_{t,M}$ given $\mathbf{X}_{A,t,M}$ only. We employ a standard Gaussian kernel and rule-of-thumb bandwidths with a Gaussian reference that we adjust to comply with the conditions in Assumption A5. In particular, we set $h = O(T^{-1/(q/2+4)} / (\ln \ln T))$ and $b = O(h^{(q+1)/q})$.⁸ For simplicity, we do not weigh the data and hence $\pi(\cdot) = 1$. To ensure a reasonable number of daily observations in the bootstrap artificial samples, we consider $\mathcal{T} = \max\{100, \lfloor T^{0.87} \rfloor\}$, though further simulations show that the results are quite robust to variations in the bootstrap sample size. In what follows, we discuss results based on $\mathcal{B} = 2,000$ bootstrap samples using overlapping blocks of length equal to the nearest integer to $\sqrt{T}$.

Tables 1 and 2 report the empirical size of the test for jump spillovers, respectively conditioning either only on the past realized jumps or on both past components of the quadratic variation. In each table, Panel A considers independent Poisson jumps by setting $\kappa_{jj} \sim N(-\sqrt{\Sigma_{j,t}/2}, \Sigma_{j,t}/2)$ for $j = A, B$ and $\kappa_{BA} = 0$ in (10) and (11), whereas Panel B examines the test behavior in the absence of jumps (i.e., $\kappa_{AA} = \kappa_{BA} = \kappa_{BB} = 0$).

There are several interesting findings. First, it is reassuring to observe that pre-averaging controls very well for market microstructure contamination in that results barely change once we increase the noise-to-signal ratio c from zero to 0.001. Second, there is a very good matching between empirical and nominal sizes, even for the smallest sample with $M = 78$ and $T = 250$, regardless of whether we additionally control or not for the past realized variance. Third, size distortions typically decrease as the number of intraday observations M grows, though not necessarily as the number of daily observations increases. This reflects

⁸To avoid too much undersmoothing in the higher dimensional case in which we also control for the past realized variance, we multiply these adjusted rule-of-thumb bandwidths by 10 similarly to Corradi et al. (2012).

to some extent the conditions in Theorem 5 that require M to grow at a faster rate than T . Fourth, in the absence of jumps, there is no single instance across the 2,000 Monte Carlos replications in which we reject the null of no jump spillover in Table 1, regardless of whether we add noise or not, whereas we find only a very few rejections at the usual significance levels in Table 2. This is very reassuring in that it coincides exactly with what our asymptotic theory predicts.

Table 3 documents the behavior of the test under the alternative. Our test for jump spillovers exhibits reasonable power even for small sample sizes of $T = 250$, with power rapidly converging to one as the number of daily observations grows. In contrast, the number of intraday observations that we use to compute the realized measures of jump contribution to the quadratic variation seems to matter very little given that power remains very stable across the different values of M . As before, the presence of noise does not affect much the properties of the test, confirming the effectiveness of pre-averaging. Finally, unnecessarily conditioning on the past realized variance harms power substantially only at the 1% significance level given that power figures are still strong at the 5% and 10% levels.

7 Jump spillovers between ETF and futures markets

We collect data at the 1-minute interval for the E-mini S&P 500 futures roll-over front contracts and the State Street Global Advisers SPDR S&P 500 exchange-traded fund (ETF) between July 1st, 1998 and February 02, 2017. The E-mini trades electronically at the Chicago Mercantile Exchange (CME) under the symbol ES, with one-fifth the size of the standard S&P 500 futures contract. Unlike standard-sized futures contracts, E-mini futures have only five contracts listed for trading, with expirations on the third Friday of the month on a March quarterly cycle. The vast majority of trades are for the nearest-to-expire contract, with volume shifting to the second-nearest-to-expiry contract exactly one week before expiration. The E-mini market runs almost continuously from Sunday 18:00 to Friday 17:00 (Eastern Time). It features daily trading halts from 16:15 to 16:30 and maintenance periods from 17:00 to 18:00, though. The SPDR S&P 500 ETF trades on US equity markets, under the NYSE Arca symbol SPY, Monday through Friday, from 9:30 to 16:00. SPY corresponds to one-tenth the value of the S&P 500 index, whereas E-mini contracts are for $\$50 \times$ the index. Accordingly, one could well think of the latter as a futures contract on approximately 500 shares of SPY.

We compute log-returns on SPY at the 1-minute interval for every trading day from 9:30 to 16:00.

We ignore any record outside the normal NYSE trading hours. For the E-mini futures, we consider the nearest-to-expiry contract up to the second Friday of the expiry month, when we shift attention to the next-to-expiry contract. We compute 1-minute log-returns both for the NYSE trading hours (9:30 to 16:00) and for the overnight period (18:00 to 9:29 in the next day). Although trading activity in the intraday period is typically much heavier than in the overnight period, jumps in the overnight period could well convey relevant information. After excluding any day in which we do not have at least 78 nonzero intraday returns for both assets, we end up with a sample of 3,867 trading days. Finally, we retrieve the daily jump contributions to the quadratic variation for both ES and SPY using the same pre-averaging procedure as in the previous section.⁹ Figure 1 plots the time series of the relevant realized measures. Interestingly, the integrated variance is of the same order of magnitude of the quadratic variation, whereas the jump contribution is typically of lower order, even if it increases considerably at times of market turmoil. This is in line with Aït-Sahalia and Xiu’s (2016) findings based on intraday data, but in contrast to the results of Santa-Clara and Yan’s (2010) analysis using daily data.

Table 4 reports the results of the bootstrap-based tests using the same choices for kernel, bandwidth, weighting scheme and bootstrap specifications (block length, bootstrap sample size, and number of bootstrap samples) as in the previous section. We carry out not only tests that control exclusively for past jump realizations, but also for tests that control additionally for the realized variance in the previous day. Apart from the full sample, we also contemplate different subsamples. The first ranges from July 1998 to December 2003, corresponding to a period of high (continuous and discontinuous) variation in Figure 1. In contrast, tranquility definitely characterizes the second subsample from January 2004 to January 2007. The third period is between February 2007 and December 2013, exhibiting extremely high levels of uncertainty due to the subprime crisis. Lastly, we observe between January 2014 to February 2017 a different sort of regime, with reasonably high quadratic variation, but little jump activity.¹⁰

We first test whether jumps in SPY yesterday affect the jumps in ES today given the realized jumps in ES yesterday during NYSE trading hours ($\mathbb{H}_0^{(1)}: ES_t \perp\!\!\!\perp SPY_{t-1} | ES_{t-1}$). To take advantage of the overnight trading in the futures market, we not only try conditioning either on the past after-hours realized jump in ES ($\mathbb{H}_0^{(2)}: ES_t \perp\!\!\!\perp SPY_{t-1} | ES_{t-1}^O$) or on the realized jump in ES over the entire previous day

⁹ The qualitative results do not change if we do not control for market microstructure noise and estimate the realized jump component by the difference between the realized variance and the bipower variation.

¹⁰ Our choice for subsamples is arbitrary, essentially resting on visual inspection of the realized measures in Figure 1.

$(\mathbb{H}_0^{(3)}: ES_t \perp\!\!\!\perp SPY_{t-1} | (ES_{t-1} + ES_{t-1}^O))$, but also test for jump spillovers from SPY to the overnight jump realization in ES given the realized jump in ES during NYSE trading hours ($\mathbb{H}_0^{(4)}: ES_t^O \perp\!\!\!\perp SPY_t | ES_t$).

Using the full sample and controlling only for own past jumps, we reject the null hypotheses $\mathbb{H}_0^{(1)}$ at the 5% level of significance, $\mathbb{H}_0^{(2)}$ and $\mathbb{H}_0^{(3)}$ at the 1% level, and $\mathbb{H}_0^{(4)}$ at the 10% level. However, the picture changes dramatically once we run the tests on the different subsamples. Interestingly, we cannot reject the absence of jump spillovers from SPY to ES in cases of $\mathbb{H}_0^{(1)}$ and $\mathbb{H}_0^{(4)}$. We take this finding as preliminary evidence that ES_t and SPY_t convey very similar information. By contrast, we keep rejecting $\mathbb{H}_0^{(2)}$ and $\mathbb{H}_0^{(3)}$ at the usual levels of significance, at least in times of high uncertainty (July 1998 to December 2003 and February 2007 to December 2013). This indicates that conditioning on the overnight integrated jumps in the E-mini futures contracts is indeed informative, despite Aït-Sahalia and Xiu's (2016) evidence of few (co-)jumps during after-market hours.¹¹

We now turn our attention to contemporaneous jump spillovers.¹² In particular, we first test whether the conditional distribution of ES_t given (SPY_t, ES_{t-1}) is equal to the conditional distribution given only ES_{t-1} , namely, $\mathbb{H}_0^{(5)}: ES_t \perp\!\!\!\perp SPY_t | ES_{t-1}$. We then substitute ES_{t-1}^O for ES_{t-1} in order to check the null $\mathbb{H}_0^{(6)}: ES_t \perp\!\!\!\perp SPY_t | ES_{t-1}^O$. Next, we also try conditioning on the realized jump component for ES over the entire previous day and on yesterday's realized jump in SPY, giving way respectively to $\mathbb{H}_0^{(7)}: ES_t \perp\!\!\!\perp SPY_t | (ES_{t-1} + ES_{t-1}^O)$ and $\mathbb{H}_0^{(8)}: SPY_t \perp\!\!\!\perp ES_t | SPY_{t-1}$. We reject at the 1% significance level the absence of contemporaneous jumps, regardless of the null hypothesis, sample period, and conditioning set we consider. The presence of contemporaneous jumps is not at all surprising given that both ES and SPY are liquid enough to reflect the daily information flow on the S&P 500 index. Moreover, it confirms that ES_t and SPY_t carry very similar information content, despite the fact that futures contract also depends on interest rates.

The last set of tests check whether jumps in ES yesterday affect jumps in SPY today given the realized jumps in SPY yesterday. Interestingly, we cannot reject the absence of jump spillovers at the usual levels of significance if focusing on ES_{t-1} (namely, $\mathbb{H}_0^{(9)}: SPY_t \perp\!\!\!\perp ES_{t-1} | SPY_{t-1}$), irrespective of the subsample we appraise. However, the result changes if we consider conditional independence with respect to the past after-hours realized jump in the E-mini futures. In particular, we reject the null hypothesis

¹¹ Further analysis nonetheless rejects at the 1% significance level the null hypothesis that today's jump component for ES during NYSE trading hours does not depend on yesterday's given their overnight counterpart.

¹² We deliberately avoid saying co-jumps because it may give the impression that jumps are necessarily at the same time, rather than just over the same trading day.

$\mathbb{H}_0^{(10)}$: $\text{SPY}_t \perp\!\!\!\perp \text{ES}_{t-1}^O | \text{SPY}_{t-1}^O$ at the 1% level of significance not only for the full sample, but also for the most recent subsamples. This indicates anew that ES_{t-1} and SPY_{t-1} carry similar information.

Finally, the results become more robust across subsamples if we further control for the past realized variance of the E-mini futures, corroborating the evidence that the volatility regime matters. In particular, once we add the realized variance, we cannot reject $\mathbb{H}_0^{(1)}$ anymore at the usual levels of significance for any (sub)sample, whereas we start rejecting $\mathbb{H}_0^{(2)}$ and $\mathbb{H}_0^{(10)}$ always at the 1% level of significance. There are essentially no qualitative changes for testing the null hypotheses $\mathbb{H}_0^{(3)}$ to $\mathbb{H}_0^{(9)}$.

All in all, we find overwhelming evidence of jump spillovers between ES and SPY. Apart from very strong evidence of contemporaneous jumps, we also find that transmission runs in both directions, indicating that price discovery is not exclusive to one particular market. This is to some extent in line with Aldrich and Lee (2016), who argue that market A may affect market B if conditions favor cross-market arbitrage opportunities for market makers in A (and vice-versa).

8 Conclusion

This paper develops formal statistical tools for nonparametric tests of conditional independence between jumps. In particular, we show how to test whether the conditional distribution of asset A 's jump contribution to quadratic variation also depends on information concerning asset B 's jump contribution. Our testing procedure involves two steps. The first stage estimates realized measures of jump contributions using intraday returns data, whereas the second step tests for conditional independence between the resulting realized measures. We show how to construct more accurate critical values by means of a simple bootstrap algorithm. Our theoretical contribution to the literature is twofold. First, the asymptotic theory we put forth specifically accounts for the impact of the estimation error in the first step of the testing procedure. Second, we test for jump spillovers without testing for jumps and hence there is no misclassification issue.

Using our nonparametric tests for conditional independence, we study jump transmission in the US equity index markets. In particular, we unveil strong evidence of contemporaneous jump spillovers between E-mini S&P 500 futures contract and the SPDR exchange-traded fund. We also find that jumps in the latter today help excite jumps in E-mini futures tomorrow. In addition, we show that jumps in ES during the overnight period matter in view that the jump components in ES and SPY during NYSE trading hours seem to have very similar information content. For instance, we find evidence of jump spillovers running

from yesterday's realized jump in ES to today's jump component in SPY only if we focus on the after-hours market for E-mini futures.

Appendix: Technical Proofs

We start with the bias terms in Theorem 1. Let $\mathbf{x} = (\mathbf{x}_A, x_q)$ and $\mu_1 = \mu_1^{(1)} + (h^q b^{-q_A} - h^{1/2}) \mu_1^{(2)} + h^{1/2} \mu_1^{(3)}$, with

$$\mu_1^{(1)} = \frac{c_{11}}{6} C_1(\mathbf{K}) \int_0^\infty \pi(\mathbf{x}) d\mathbf{x} \quad (12)$$

$$\mu_1^{(2)} = \frac{c_{10}^{(A)}}{6} C_1(\mathbf{K}) \int_0^\infty \pi(\mathbf{x}_A, 0) d\mathbf{x}_A \quad (13)$$

$$\mu_1^{(3)} = \frac{c_{01}}{6} C_1(\mathbf{K}) \int_0^\infty \pi(\mathbf{0}, x_q) dx_q, \quad (14)$$

where $c_{10}^{(A)} = \text{plim}_{T \rightarrow \infty} \frac{T_{10}}{T_{1A}}$ and $T_{1A} = \sum_{t=1}^T \mathbf{1}(A_{t-1} > 0)$. Let also

$$\mu_2 = \frac{c_{11}}{6} C_1(\mathbf{K}) \int_0^\infty \mathbb{E}[\pi(\mathbf{x}) | \mathbf{x}_A] d\mathbf{x}_A \quad (15)$$

$$\mu_3 = \frac{c_{11}}{6} K(0) \int_0^\infty \mathbb{E}[\pi(\mathbf{x}) | \mathbf{x}_A] d\mathbf{x}_A. \quad (16)$$

We now rewrite the infeasible statistic as $S_T = \sum_{j=1}^4 S_{j,T}$, where

$$S_{1,T} = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t) - \widehat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{11,t}$$

$$S_{2,T} = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t) - \widehat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{10,t}$$

$$S_{3,T} = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t) - \widehat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{01,t}$$

$$S_{4,T} = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t | \mathbf{X}_t) - \widehat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{00,t}.$$

For notational simplicity and without loss of generality, we assume from now on that $K(0) = C = 1$ in Assumption A1. The proof of Theorem 1 follows directly from Lemmata 1A to 4A, which we first state and then prove in what follows.

Lemma 1A: *Let Assumptions A1 to A5 hold and $T_{11}/T \xrightarrow{p} c_{11}$, with $0 < c_{11} \leq 1$.*

(i) $S_{1,T} - h^{-q/2} \mu_1^{(1)} - h^{q/2} b^{-q_A} \mu_2 + 2h^{1-q/2} \mu_3 \xrightarrow{d} N(0, \sigma^2)$ under $\mathbb{H}_0$.

(ii) $Pr(T_{11}^{-1} h^{-q/2} S_{1,T} > \varepsilon) \rightarrow 1$ under $\mathbb{H}_A$.

Lemma 2A: *Under Assumptions A1 to A5, $S_{2,T} - (h^{q/2} b^{-q_A} - h^{1-q/2}) \mu_1^{(2)} = O_p(h^{1/2} + h^{q/2} b^{-q_A/2})$.*

Lemma 3A: *Under Assumptions A1 to A5, $S_{3,T} - h^{1-q/2} \mu_1^{(3)} = O_p(h^{1/2})$.*

Lemma 4A: *Under Assumptions A1 to A5, $S_{4,T} = O_p(h^{q/2} b^{-q_A/2})$.*

Proof of Lemma 1A:

(i) We first rewrite $S_{1,T}$ as

$$\begin{aligned} S_{1,T} &= h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_t) \right]^2 \pi(\mathbf{X}_t) I_{11,t} + h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) - F_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{11,t} \\ &\quad - 2h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_t) \right] \left[\widehat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) - F_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right] \pi(\mathbf{X}_t) I_{11,t} \\ &= S_{11,T} + S_{12,T} + S_{13,T}. \end{aligned}$$

The proof then follows by showing that (a) $S_{11,T} - h^{-q/2} \mu_1^{(1)} \xrightarrow{d} N(0, \sigma^2)$, (b) $S_{12,T} = h^{q/2} b^{-q_A} \mu_2 + o_p(1)$, and (c) $S_{13,T} = -2h^{1-q/2} \mu_3 + o_p(1)$.

(a) Recalling that $\widehat{f}_{\mathbf{X}}(\mathbf{X}_t) = \frac{1}{T_{11}} \sum_{s=1}^T \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t)$, with $\mathbf{K}_h(\mathbf{u}) = \prod_{j=1}^d K_h(u_j)$ for $\mathbf{u} = (u_1, \dots, u_d)$, it follows that

$$\begin{aligned} S_{11,T} &= h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_t) \right]^2 \pi(\mathbf{X}_t) I_{11,t} \\ &= h^{q/2} \sum_{t=1}^T \frac{\pi(\mathbf{X}_t) I_{11,t}}{\widehat{f}_{\mathbf{X}}^2(\mathbf{X}_t)} \left[\frac{1}{T_{11}} \sum_{s=1}^T \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t) (\mathbf{1}(A_s \leq A_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_s)) \right]^2 \\ &\quad + h^{q/2} \sum_{t=1}^T \frac{\pi(\mathbf{X}_t) I_{11,t}}{\widehat{f}_{\mathbf{X}}^2(\mathbf{X}_t)} \left\{ \frac{1}{T_{11}} \sum_{s=1}^T \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t) [F_{A|\mathbf{X}}(A_t|\mathbf{X}_s) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_t)] \right\}^2 \\ &\quad + h^{q/2} \sum_{t=1}^T \frac{\pi(\mathbf{X}_t) I_{11,t}}{\widehat{f}_{\mathbf{X}}^2(\mathbf{X}_t)} \left\{ \frac{1}{T_{11}} \sum_{s=1}^T \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t) [\mathbf{1}(A_s \leq A_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_s)] \right. \\ &\quad \left. \times \frac{1}{T_{11}} \sum_{s=1}^T \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t) [F_{A|\mathbf{X}}(A_t|\mathbf{X}_s) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_s)] \right\} \\ &= S_{11,T}^{(1)} + S_{11,T}^{(2)} + S_{11,T}^{(3)}, \end{aligned}$$

Note that the reason why we rescale by T_{11} , rather than by T , is that the number of observations in a neighborhood interval $\mathbf{x} \pm h$ is almost surely of order $O(T_{11}h^q)$ for both $\mathbf{x}_A$ and $\mathbf{x}_q$ greater than zero. Under Assumptions A1 and A2, it follows that $S_{11,T}^{(2)} = O_p(T_{11}h^{4+q/2}) = o_p(1)$ by Assumption A5(i). By the same argument as in the proof of Theorem 2 in Aït-Sahalia et al. (2009b), $S_{11,T}^{(3)} = O_p(T_{11}h^{4+q/2}) = o_p(1)$ as well. By a similar argument as in the proof of Lemma 1 in Corradi et al. (2012), this yields

$$\begin{aligned} S_{11,T} &= h^{q/2} \sum_{t=1}^T \frac{\pi(\mathbf{X}_t) I_{11,t}}{T_{11}^2 \widehat{f}_{\mathbf{X}}(\mathbf{X}_t)} \left\{ \sum_{s=1}^T \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t) [\mathbf{1}(A_s \leq A_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_s)] \right\}^2 + O_p\left(T_{11}^{-1/2} \sqrt{\ln T_{11}} h^{-q/2}\right) + o_p(1) \\ &= \sum_{t < s < k} \left[\phi(t, s, k) + \phi(t, k, s) + \phi(s, k, t) + \phi(s, t, k) + \phi(k, s, t) + \phi(k, t, s) \right] \\ &\quad + \sum_{t < s} \left[\phi(t, t, s) + \phi(t, s, t) + \phi(s, t, t) + \phi(s, s, t) + \phi(s, t, s) + \phi(t, s, s) \right] + \sum_t \phi(t, t, t) + o_p(1), \end{aligned} \tag{17}$$

where

$$\phi(t, s, k) = h^{q/2} \frac{\pi(\mathbf{X}_t) I_{11,t}}{T_{11}^2 f_{\mathbf{X}}(\mathbf{X}_t)} \mathbf{K}_h(\mathbf{X}_s - \mathbf{X}_t) \left[\mathbf{1}(A_s \leq A_t) - F_{A|\mathbf{X}}(A_t | \mathbf{X}_s) \right] \mathbf{K}_h(\mathbf{X}_k - \mathbf{X}_t) \left[\mathbf{1}(A_k \leq A_t) - F_{A|\mathbf{X}}(A_t | \mathbf{X}_k) \right].$$

It is immediate to see that $\sum_t \phi(t, t, t) = o_p(1)$. By a similar argument as in Aït-Sahalia et al.'s (2009b) proof of Theorem 1, in view that Assumption A5(ii) implies that $T_{11} h^{3q/2} \rightarrow \infty$, the first term on the last equality in (17) reads $(T - 2) \sum_{t < s} \phi^\dagger(t, s) + o_p(1)$, where $\phi^\dagger(t, s) = \int \phi^\dagger(t, s, k) dF_{A, \mathbf{X}}(a_k, \mathbf{x}_k)$ and

$$\phi^\dagger(t, s, k) = \phi(t, s, k) + \phi(t, k, s) + \phi(s, k, t) + \phi(s, t, k) + \phi(k, s, t) + \phi(k, t, s). \quad (18)$$

In addition, the second term on the right-hand side of the last equality in (17) equals $\frac{T_{11}(T_{11}-1)}{2} \phi(0) + o_p(1)$, where $\phi(0) = \mathbb{E}[\phi(t)]$ and $\phi(t) = \int \phi(t, s) dF_{A, \mathbf{X}}(a_s, \mathbf{x}_s)$. The expressions for μ_1 and σ^2 follow by the same argument as in the proof of Theorem 2 in Aït-Sahalia et al. (2009b).

(b) By Assumptions A5(iii) and A5(iv), $Th^{q/2} b^4 \rightarrow 0$ and $Th^{-q} b^{5q_A/2} \rightarrow \infty$, and hence

$$S_{12,T} = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) - F_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{11,t} = (T_{11} - 2) \sum_{t < s} \widetilde{\phi}^\dagger(t, s) + \frac{T_{11}(T_{11}-1)}{2} \widetilde{\phi}(0) + o_p(1),$$

where $T_{11} \leq T_{1A} = \sum_{t=1}^T \mathbf{1}(A_{t-1} > 0) \leq T$, and $\widetilde{\phi}^\dagger(t, s)$ and $\widetilde{\phi}(0)$ are analogous to $\phi^\dagger(t, s)$ and $\phi(0)$, but using

$$\begin{aligned} \widetilde{\phi}(t, s, k) = h^{q/2} \frac{\pi(\mathbf{X}_t) I_{11,t}}{T_{1A}^2 f_{\mathbf{X}_A}(\mathbf{X}_{A,t})} \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t}) \left[\mathbf{1}(A_s \leq A_t) - F_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,s}) \right] \\ \times \mathbf{K}_b(\mathbf{X}_{A,k} - \mathbf{X}_{A,t}) \left[\mathbf{1}(A_k \leq A_t) - F_{A|\mathbf{X}_A}(A_t | \mathbf{X}_{A,k}) \right] \end{aligned}$$

instead of $\phi(k, t, s)$. As before, $(T_{11} - 2) \sum_{t < s} \widetilde{\phi}^\dagger(t, s) = O_p(h^{1/2})$. Also,

$$\begin{aligned} \frac{T_{11}^2}{c_{11}^{(A)}} \widetilde{\phi}(0) &= 2 h^{q/2} \int_{\mathbf{x}_i > 0} \int_{a_j, \mathbf{x}_{A,j}} \frac{\pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}^2(\mathbf{x}_{A,i})} \left\{ \mathbf{K}_b(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) \left[\mathbf{1}(a_j \leq a_i) - F_{A|\mathbf{X}_A}(a_i | \mathbf{x}_{A,j}) \right] \right\}^2 dF_{A, \mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A, \mathbf{X}}(a_i, \mathbf{x}_i) \\ &= 2 h^{q/2} \int_{\mathbf{x}_i > 0} \int_{a_j, \mathbf{x}_{A,j}} \frac{\pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}^2(\mathbf{x}_{A,i})} \mathbf{K}_b^2(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) \mathbf{1}(a_j \leq a_i) dF_{A, \mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A, \mathbf{X}}(a_i, \mathbf{x}_i) \\ &\quad + 2 h^{q/2} \int_{\mathbf{x}_i > 0} \int_{a_j, \mathbf{x}_{A,j}} \frac{\pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}^2(\mathbf{x}_{A,i})} \mathbf{K}_b^2(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) F_{A|\mathbf{X}_A}^2(a_i | \mathbf{x}_{A,j}) dF_{A, \mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A, \mathbf{X}}(a_i, \mathbf{x}_i) \\ &\quad - 4 h^{q/2} \int_{\mathbf{x}_i > 0} \int_{a_j, \mathbf{x}_{1,j}} \frac{\pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}^2(\mathbf{x}_{A,i})} \mathbf{K}_b^2(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) \mathbf{1}(a_j \leq a_i) F_{A|\mathbf{X}_A}(a_i | \mathbf{x}_{A,j}) dF_{A, \mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A, \mathbf{X}}(a_i, \mathbf{x}_i). \quad (19) \end{aligned}$$

In view that $F_{A|\mathbf{X}_A}$ is uniform over the unit interval and given Assumption A5(vi), the first term on the right-hand side of the second equality in (19) is equal to

$$h^{q/2} b^{-q_A} C_1(K) \int_{\mathbf{x}_i > 0} \pi(\mathbf{x}_i) f_{B|\mathbf{X}_A}(x_{q,i} | \mathbf{x}_{A,i}) d\mathbf{x}_i = h^{q/2} b^{-q_A} C_1(K) \int_{\mathbf{x}_{A,i} > 0} \mathbb{E}[\pi(\mathbf{x}_i | \mathbf{x}_{A,i}, x_{q,i} > 0)] d\mathbf{x}_{A,i}.$$

Treating analogously the second and third terms of the second equality on the right-hand side of (19) and then following the same reasoning as in Aït-Sahalia et al.'s (2009b) proof of Theorem 2 yield the result.

(c) Define $\bar{\phi}^\dagger(t, s)$ and $\bar{\phi}(0)$ analogously to $\phi^\dagger(t, s)$ and $\phi(0)$, but using

$$\bar{\phi}(t, s, k) = h^{q/2} \frac{\pi(\mathbf{X}_t) I_{11,t}}{T_{11} T_{1A} f_{\mathbf{X}_A}(\mathbf{X}_{A,t})} K_b(A_{s-1} - A_{t-1}) [\mathbf{1}(A_s \leq A_t) - F_{A|\mathbf{X}_A}(A_t|A_{s-1})] \mathbf{K}_h(\mathbf{X}_k - \mathbf{X}_t) [\mathbf{1}(A_k \leq A_t) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_k)]$$

instead of $\phi(k, t, s)$. It then holds that

$$S_{13,T} = (T_{11} - 2) \sum_{t < s} \bar{\phi}^\dagger(t, s) + \frac{T_{11}(T_{1A} - 1)}{2} \bar{\phi}(0) + o_p(h),$$

with $(T_{11} - 2) \sum_{t < s} \bar{\phi}^\dagger(t, s) = o_p(h^{1/2})$. Given the bandwidth rate conditions, by a similar argument as in the proof (a) and (b),

$$\begin{aligned} \frac{T_{11}^2}{c_{11}^{(A)}} \bar{\phi}(0) &= 2 \int_{\substack{a_i > 0 \\ \mathbf{x}_i > 0}} \int_{a_j, \mathbf{x}_{A,j}} \frac{h^{q/2} \pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}(\mathbf{x}_{A,i}) f_{\mathbf{X}}(\mathbf{x}_i)} \mathbf{K}_b(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) \mathbf{K}_h(\mathbf{x}_j - \mathbf{x}_i) \mathbf{1}(a_j \leq a_i) dF_{A,\mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A,\mathbf{X}}(a_i, \mathbf{x}_i) \\ &+ 2 \int_{\substack{a_i > 0 \\ \mathbf{x}_i > 0}} \int_{a_j, \mathbf{x}_{A,j}} \frac{h^{q/2} \pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}^2(\mathbf{x}_{A,i}) f_{\mathbf{X}}(\mathbf{x}_i)} \mathbf{K}_b(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) \mathbf{K}_h(\mathbf{x}_j - \mathbf{x}_i) F_{A|\mathbf{X}_A}^2(a_i|\mathbf{x}_{A,j}) dF_{A,\mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A,\mathbf{X}}(a_i, \mathbf{x}_i) \\ &- 4 \int_{\substack{a_i > 0 \\ \mathbf{x}_i > 0}} \int_{a_j, \mathbf{x}_{A,j}} \frac{h^{q/2} \pi(\mathbf{x}_i)}{f_{\mathbf{X}_A}(\mathbf{x}_{A,i}) f_{\mathbf{X}}(\mathbf{x}_i)} \mathbf{K}_b(\mathbf{x}_{A,j} - \mathbf{x}_{A,i}) \mathbf{K}_h(\mathbf{x}_j - \mathbf{x}_i) \mathbf{1}(a_j \leq a_i) F_{A|\mathbf{X}_A}(a_i|\mathbf{x}_{A,j}) \\ &\quad \times dF_{A,\mathbf{X}_A}(a_j, \mathbf{x}_{A,j}) dF_{A,\mathbf{X}}(a_i, \mathbf{x}_i). \end{aligned} \quad (20)$$

As before, we note that it is possible to rewrite the first term on the right-hand side of (20) as

$$K(0) \int_{\mathbf{x}_i > 0} \pi(\mathbf{x}_i) f_{B|\mathbf{X}_A}(x_{q,i}|\mathbf{x}_{A,i}) d\mathbf{x}_i = K(0) \int_{\mathbf{x}_{A,i} > 0} \mathbb{E}[\pi(\mathbf{x}_i) | \mathbf{x}_{A,i}, x_{q,i} > 0] d\mathbf{x}_{A,i}.$$

The same treatment applies to the second and third terms on the right-hand side of (20), yielding the desired result by the same argument as in Aït-Sahalia et al.'s (2009b) proof of Theorem 2. $\blacksquare$

Proof of Lemma 2A: Note that $h^{q/2} \sum_{t=1}^T [F_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0)] \pi(\mathbf{X}_t) = 0$. As only the positive realizations of A_{s-1} have a contribution,

$$\hat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) = \frac{\frac{1}{T_{1A}} \sum_{s=1}^T \mathbf{1}(A_s \leq A_t) \mathbf{K}_b(\mathbf{X}_{A,t} - \mathbf{X}_{A,s})}{\frac{1}{T_{1A}} \sum_{s=1}^T \mathbf{K}_b(\mathbf{X}_{A,t} - \mathbf{X}_{A,s})}.$$

By the same argument as in the proof of Lemma 1A,

$$h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) - F_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) I_{10,t} = O_p(h^{q/2} b^{-qA/2}) + h^{q/2} b^{-qA} \mu_1^{(2)}.$$

Similarly,

$$\begin{aligned}\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) &= \frac{\frac{1}{T_{10}} \sum_{s \in T_{10}} \mathbf{1}(A_s \leq A_t) \mathbf{K}_h(\mathbf{X}_{A,t} - \mathbf{X}_{A,s}) + \frac{1}{T_{10}} \sum_{s \notin T_{10}} \mathbf{1}(A_s \leq A_t) \mathbf{K}_h(\mathbf{X}_{A,t} - \mathbf{X}_{A,s}) K_h(B_s)}{\frac{1}{T_{10}} \sum_{s \in T_{10}} \mathbf{K}_h(\mathbf{X}_{A,t} - \mathbf{X}_{A,s}) + \frac{1}{T_{10}} \sum_{s \notin T_{10}} \mathbf{K}_h(\mathbf{X}_{A,t} - \mathbf{X}_{A,s}) K_h(B_s)} \\ &= \frac{\frac{1}{T_{10}} \sum_{s \in T_{10}} \mathbf{1}(A_s \leq A_t) \mathbf{K}_h(\mathbf{X}_{A,t} - \mathbf{X}_{A,s}) (1 + o_p(1))}{\frac{1}{T_{10}} \sum_{s \in T_{10}} \mathbf{K}_h(\mathbf{X}_{A,t} - \mathbf{X}_{A,s}) (1 + o_p(1))}\end{aligned}$$

with the $o_p(1)$ terms holding uniformly for $A_{t-1} > 0$. By the same argument as in the proof of Lemma 1A,

$$h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) \right]^2 \pi(\mathbf{X}_t) I_{10,t} = O_p(h^{1/2}) + h^{1-q/2} \mu_1^{(2)}.$$

In turn, by a similar argument as in the proof of step (c) in Lemma 1A, it follows that

$$h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) - F_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right] \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) \right] \pi(\mathbf{X}_t) I_{10,t} = O_p(h^{q/2-1}).$$

To complete the proof, it now suffices to follow the same argument as in the proof of Theorem 1 in Corradi et al. (2012). ■

Proof of Lemma 3A: Recall that $\mathbf{X}_{A,t} = (A_{t-1}, X_{qA,t})$, where $X_{qA,t}$ is either $IV_{A,t}$ or nothing. Under $\mathbb{H}_0$, further conditioning on B_t does not make any difference once we control for $\mathbf{X}_{A,t}$ and hence $F_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) = F_{A|\mathbf{X}}(A_t|0, X_{qA,t}, B_t)$. The kernel estimator for the former is given by

$$\widehat{F}_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) = \frac{\frac{1}{T_{0A}} \sum_{s=1}^T \mathbf{1}(A_s \leq A_t) K_b(A_{s-1}) K_b(X_{qA,t} - X_{qA,s})}{\frac{1}{T_{0A}} \sum_{s=1}^T K_b(A_{s-1}) K_b(X_{qA,t} - X_{qA,s})}, \quad (21)$$

where $T_{0A} = \sum_{t=1}^T \mathbf{1}(A_{t-1} = 0)$. By the same argument as in the proof of Lemma 2A, for $t \in T_{01}$,

$$\begin{aligned}\widehat{F}_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) &= \frac{\frac{1}{T_{01}b} \sum_{s \in T_{01}} \mathbf{1}(A_s \leq A_t) K_b(X_{qA,t} - X_{qA,s}) + \frac{1}{T_{01}} \sum_{s \notin T_{01}} \mathbf{1}(A_s \leq A_t) K_b(A_{s-1}) K_b(X_{qA,t} - X_{qA,s})}{\frac{1}{b} + \frac{1}{T_{01}} \sum_{s \notin T_{01}} K_b(A_{s-1}) K_b(X_{qA,t} - X_{qA,s})} \\ &= \frac{\frac{1}{T_{01}} \sum_{s \in T_{01}} \mathbf{1}(A_s \leq A_t) K_b(X_{qA,t} - X_{qA,s}) + \frac{1}{T_{01}} \sum_{s \notin T_{01}} \mathbf{1}(A_s \leq A_t) K(A_{s-1}/b) K_b(X_{qA,t} - X_{qA,s})}{1 + \frac{1}{T_{01}} \sum_{s \notin T_{01}} K(A_{s-1}/b) K_b(X_{qA,t} - X_{qA,s})},\end{aligned}$$

so that

$$\widehat{F}_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) - F_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) = \left[\frac{1}{T_{01}} \sum_{s \in T_{01}} \mathbf{1}(A_s \leq A_t) K_b(X_{qA,t} - X_{qA,s}) - F_{\mathbf{X}_A}(0, X_{qA,t}) \right] (1 + o_p(1)).$$

This means that $h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) - F_{A|\mathbf{X}_A}(A_t|0, X_{qA,t}) \right]^2 \pi(\mathbf{X}_t) I_{01,t} = O_p(h^{q/2})$. Also, for $t \in T_{01}$,

$$\begin{aligned}\widehat{F}_{A|\mathbf{X}}(A_t|0, X_{qA,t}, B_t) &= \frac{\frac{1}{T_{01}h^qA} \sum_{s \in T_{01}} \mathbf{1}(A_s \leq A_t) K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s) + \frac{1}{T_{01}} \sum_{s \notin T_{01}} \mathbf{1}(A_s \leq A_t) K_h(A_{s-1}) K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s)}{\frac{1}{T_{01}h^qA} \sum_{s \in T_{01}} K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s) + \frac{1}{T_{01}} \sum_{s \notin T_{01}} K_h(A_{s-1}) K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s)} \\ &= \frac{\frac{1}{T_{01}} \sum_{s \in T_{01}} \mathbf{1}(A_s \leq A_t) K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s) + \frac{h^qA}{T_{01}} \sum_{s \notin T_{01}} \mathbf{1}(A_s \leq A_t) K_h(A_{s-1}) K_b(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s)}{\frac{1}{T_{01}} \sum_{s \in T_{01}} K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s) + \frac{h^qA}{T_{01}} \sum_{s \notin T_{01}} K_h(A_{s-1}) K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s)} \\ &= \frac{\frac{1}{T_{01}} \sum_{s \in T_{01}} \mathbf{1}(A_s \leq A_t) K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s) (1 + o_p(1))}{\frac{1}{T_{01}} \sum_{s \in T_{01}} K_h(X_{qA,t} - X_{qA,s}) K_h(B_t - B_s) (1 + o_p(1))}\end{aligned}$$

and so $h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|0, X_{q_A,t}, B_t) - F_{A|\mathbf{X}}(A_t|0, X_{q_A,t}, B_t) \right]^2 \pi(\mathbf{X}_t) I_{01,t} = O_p(h^{1/2}) + h^{1-q/2} \mu_1^{(3)}$. $\blacksquare$

Proof of Lemma 4A: By the same argument as in the proof of Lemma 3A,

$$h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}_A}(A_t|0, X_{q_A,t}) - F_{A|\mathbf{X}_A}(A_t|0, X_{q_A,t}) \right]^2 \pi(\mathbf{X}_t) I_{00,t} = O_p(h^{q/2} b^{-q_A/2}),$$

whereas, for $t \in T_{00}$,

$$\widehat{F}_{A|\mathbf{X}}(A_t|0, X_{q_A,t}, 0) = \frac{\frac{1}{T_{00}} \sum_{s \in T_{00}} \mathbf{1}(A_s \leq A_t) K_h(X_{q_A,t} - X_{q_A,s}) (1 + o_p(1))}{\frac{1}{T_{00}} \sum_{s \in T_{00}} K_h(X_{q_A,t} - X_{q_A,s}) (1 + o_p(1))}.$$

It then follows that $h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|0, X_{q_A,t}, 0) - F_{A|\mathbf{X}}(A_t|0, X_{q_A,t}, 0) \right]^2 \pi(\mathbf{X}_t) I_{00,t} = O_p(h^{q/2-1})$. $\blacksquare$

Proof of Theorem 1: It readily follows from Lemmata 1A to 4A. $\blacksquare$

Proof of Theorem 2: The proof is very similar to that in Lemma 3A and hence we provide only a sketch in the sequel.

(i) Without jumps in A , $\mathbf{1}(A_s \leq A_t) = 1$ almost surely and $F_{A|\mathbf{X}_A}(0|0, X_{q_A,t}) = F_{A|\mathbf{X}}(0|0, X_{q_A,t}, B_t) = 1$ as well. It then follows that $h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|0, X_{q_A,t}, B_t) - \widehat{F}_{A|\mathbf{X}_A}(A_t|0, X_{q_A,t}) \right]^2 \pi(\mathbf{X}_t) = 0$.

(ii) In the absence of jumps in B , $h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) - F_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) = O_p(h^{q/2} b^{-q_A/2}) + h^{q/2} b^{-q_A} \mu_1^{(2)}$ as in the proofs of Lemmata 1A and 2A, whereas by the same argument as in the proof of Lemma 2A $h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) - F_{A|\mathbf{X}}(A_t|\mathbf{X}_{A,t}, 0) \right]^2 = O_p(h^{1/2}) - h^{(1-q)/2} \mu_1^{(2)}$ given that $K(0) = 1$.

Proof of Lemma 1: For notational simplicity, we suppress any subscript index referring to the specific asset. Recall that we observe only the noisy version $Z_t = p_t + \epsilon_t$ of the efficient asset price. By decomposing the latter into continuous and discontinuous components, viz. $p_t = p_t^{(c)} + p_t^{(d)}$, the pre-averaging realized variance in (8) becomes

$$\begin{aligned} PV_{M,t}(2, 0) &= \frac{1}{\sqrt{M}} \sum_{j=1}^{M-2k_M+1} \left\{ \left[\sum_{\ell=1}^{k_M} g(\ell/M) \left(p_{t+(j+\ell)/M}^{(c)} - p_{t+(j+\ell+1)/M}^{(c)} + \epsilon_{t+(j+\ell)/M} - \epsilon_{t+(j+\ell+1)/M} \right) \right]^2 \right. \\ &\quad \left. + \sum_{\ell=1}^{k_M} g(\ell/M) \left(p_{t+(j+\ell)/M}^{(d)} - p_{t+(j+\ell+1)/M}^{(d)} \right)^2 + \text{cross-terms} \right\} \\ &= V_{M,t}(p_t^{(c)} + \epsilon_t) + V_{M,t}(p_t^{(d)}) + \text{cross-terms}. \end{aligned}$$

Recall that $a_M = \sqrt{M}$ as the pre-averaging realized variance is robust to microstructure noise and also that $\int_0^1 g^2(s) ds = 1/12$ for $g(u) = \min\{u, 1-u\}$. The proof follows in four steps.

(a) We begin showing that

$$\mathbb{E} \left| \frac{1}{12\theta} V_{M,t}(p_t^{(d)}) - \sum_{t-1 \leq s \leq t} |\Delta p_s|^2 \right|^k = a_M^{-k/2}.$$

Given Assumption A8, it follows from Aït-Sahalia and Jacod's (2012) Lemma 1 that $p_t^{(d)}$ is a process of finite variation for all t and hence, with probability one, $\sum_{t-1 \leq s \leq t} |\Delta p_s| < \infty$. This means that, on any unit interval, we have with probability one at most M^ψ jumps of size $M^{-\psi}$, with $\psi \in [0, 1/2)$. For $\psi = 1/4$ and $\varepsilon = O(M^{1/4})$, independence between jumps within each day ensures that

$$\mathbb{E} \left[\sum_{t-1 \leq s \leq t} |\Delta p_s|^2 \mathbf{1}(|\Delta p_s| \leq \varepsilon) \right]^k = O(M^{-\psi k}) = O(a_M^{-k/2}).$$

Now, let $\Omega_{M,t}(\varepsilon)$ denote the set of realizations ω such that, at day t , jumps of size larger than ε are far apart by at least $M^{-1/2}$ price changes. It turns out that, by steps 1 to 4 in the proof of Jacod, Podolskij and Vetter's (2010) Theorem 1, $\frac{1}{12\theta} V_{M,t}(p_t^{(d)}) - \sum_{t-1 \leq s \leq t} |\Delta p_s|^2 \mathbf{1}(|\Delta p_s| > \varepsilon) = o(M^{-1/4})$ for every $\omega \in \Omega_{M,t}(\varepsilon)$.

This means that

$$\mathbb{E} \left[\left| \frac{1}{12\theta} V_{M,t}(p_t^{(d)}) - \sum_{t-1 \leq s \leq t} |\Delta p_s|^2 \mathbf{1}(|\Delta p_s| > \varepsilon) \right|^k \right] = o(M^{-k/4}) = o(a_M^{-k/2}).$$

Given that ε is of order $O(M^{1/4})$, $\Pr(\Omega_{M,t}(\varepsilon)) \rightarrow 1$ as $M \rightarrow \infty$, completing the first step of the proof.

(b) We next show that

$$\mathbb{E} \left[\left| V_{M,t}(p_t^{(c)} + \epsilon_t) - \frac{1}{M} \frac{1}{24\theta^2} RV_t - \Sigma_t \right|^k \right] = O(a_M^{-k/2}),$$

with $RV_t = \sum_{j=0}^{M-1} \left(p_{t+(j+\ell)/M}^{(c)} - p_{t+(j+\ell+1)/M}^{(c)} + \epsilon_{t+(j+\ell)/M} - \epsilon_{t+(j+\ell+1)/M} \right)^2$ corresponding to the standard realized variance measure (i.e., without any pre-averaging). Remark 1 in Jacod et al. (2009) clarifies that $V_{M,t}(p_t^{(c)} + \epsilon_t) - \frac{1}{M} \frac{1}{24\theta^2} RV_t$ is equivalent, up to some border terms, to the realized kernel estimator of Barndorff-Nielsen et al. (2008), with a kernel given by $\frac{1}{12} \int_s^1 g(u)g(u-s) du$. In addition, the border terms have mean zero and are of the same order as the difference between the realized kernel estimator and the integrated volatility. The statement then readily ensues from Lemma 1 in Corradi, Distaso and Swanson (2011).

(c) We now show that, as long as $p \geq (k+2)/2$,

$$\mathbb{E} \left[\left| PV_{M,t}(2/p, \dots, 2/p) - PV_{M,t}^{(c)}(2/p, \dots, 2/p) \right|^k \right] = a_M^{-k/2},$$

where $PV_{M,t}^{(c)}(2/p, \dots, 2/p) = \frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \left| \sum_{\ell=1}^{k_M} g(t + \ell/M) \left(\Delta Z_{t+(j+ik_M+\ell)/M} - \Delta p_{t+(j+ik_M+\ell)/M}^{(d)} \right) \right|^{2/p}$. Let $V_{t,(j+ik_M)/M}^{(z)} = \left| \sum_{\ell=1}^{k_M} g(t + \ell/M) \Delta Z_{t+(j+ik_M+\ell)/M} \right|^{2/p}$ and define $V_{t,(j+ik_M)/M}^{(c)}$ analogously, but using only the continuous part of Z_t , that is to say,

$$V_{t,(j+ik_M)/M}^{(c)} = \left| \sum_{\ell=1}^{k_M} g(t + \ell/M) \left(\Delta Z_{t+(j+ik_M+\ell)/M} - \Delta p_{t+(j+ik_M+\ell)/M}^{(d)} \right) \right|^{2/p}.$$

By the same argument used in Section 3 of Barndorff-Nielsen et al. (2006),

$$\begin{aligned} \left| \frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \left(V_{t,(j+ik_M)/M}^{(z)} - V_{t,(j+ik_M)/M}^{(c)} \right) \right| &\leq \frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} V_{t,(j+ik_M)/M}^{(z)} \\ &\quad + \frac{\binom{p}{1}}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-2} V_{t,(j+ik_M)/M}^{(z)} V_{t,(j+(p-1)k_M)/M}^{(c)} \\ &\quad + \frac{\binom{p}{2}}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-3} V_{t,(j+ik_M)/M}^{(z)} \prod_{i=p-2}^{p-1} V_{t,(j+ik_M)/M}^{(c)} + \dots \\ &\quad + \frac{\binom{p}{p-1}}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} V_{t,j/M}^{(z)} \prod_{i=1}^{p-1} V_{t,(j+ik_M)/M}^{(c)}. \end{aligned} \quad (22)$$

Let now $\bar{V}_{t,(j+ik_M)/M}^{(z)} = V_{t,(j+ik_M)/M}^{(z)} - \mu_{V^{(z)}}$ and $\bar{V}_{t,(j+ik_M)/M}^{(c)} = V_{t,(j+ik_M)/M}^{(c)} - \mu_{V^{(c)}}$, with $\mu_{V^{(z)}} = \mathbb{E}[V_{t,(j+ik_M)/M}^{(z)}]$ and $\mu_{V^{(c)}} = \mathbb{E}[V_{t,(j+ik_M)/M}^{(c)}]$. We first deal with the case of finite activity jumps for which there is at most a finite number of jumps over a day:

$$\mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} - \sqrt{M} \mu_{V^{(z)}}^p \right]^k \leq \mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \right]^k + \left(\sqrt{M} \mu_{V^{(z)}}^p \right)^k. \quad (23)$$

Given that the probability of having a jump in each interval of length M^{-1} is of order M^{-1} and that jumps are bounded in size, $\sqrt{M} \mu_{V^{(z)}}^p = O(M^{(1-p)/2}) = O(a_M^{1-p})$, and thus $M^{k/2} \mu_{V^{(z)}}^{pk} = O(a_M^{-k/2})$ for all k and $p \geq 3/2$.

We now turn our attention to the first term on the right-hand side of (23), but setting $k = 4$ for the sake of simplicity. It follows from $\mathbb{E} \left[\prod_{i=0}^{p-1} \bar{V}_{t,(j_1+ik_M)/M}^{(z)} \prod_{i=0}^{p-1} \bar{V}_{t,(j_2+ik_M)/M}^{(z)} \right] = 0$ for $|j_1 - j_2| > M$ that

$$\begin{aligned} V_4 &\equiv \mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \right]^4 \\ &= \frac{1}{M^2} \sum_{1 \leq j, j_1, j_2, j_3 \leq M-pk_M+1} \sum \sum \sum \sum \mathbb{E} \left[\prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \prod_{i=0}^{p-1} \bar{V}_{t,(j_1+ik_M)/M}^{(z)} \prod_{i=0}^{p-1} \bar{V}_{t,(j_2+ik_M)/M}^{(z)} \prod_{i=0}^{p-1} \bar{V}_{t,(j_3+ik_M)/M}^{(z)} \right] \\ &\leq \sqrt{M} \sqrt{\mathbb{E} \left[\left(\prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \right)^2 \left(\prod_{i=0}^{p-1} \bar{V}_{t,(j_1+ik_M)/M}^{(z)} \right)^2 \right]} \sqrt{\mathbb{E} \left[\left(\prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \right)^2 \left(\prod_{i=0}^{p-1} \bar{V}_{t,(j_1+ik_M)/M}^{(z)} \right)^2 \right]} \\ &\leq \sqrt{M} \sqrt[4]{\mathbb{E} \left(\prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \right)^4} \sqrt[4]{\mathbb{E} \left(\prod_{i=0}^{p-1} \bar{V}_{t,(j_1+ik_M)/M}^{(z)} \right)^4} \sqrt[4]{\mathbb{E} \left(\prod_{i=0}^{p-1} \bar{V}_{t,(j_2+ik_M)/M}^{(z)} \right)^4} \sqrt[4]{\mathbb{E} \left(\prod_{i=0}^{p-1} \bar{V}_{t,(j_3+ik_M)/M}^{(z)} \right)^4} \\ &= O(M^{(1-p)/2}) = O(a_M^{1-p}), \end{aligned}$$

which is of order $O(a_M^{-k/2})$ provided that $p \geq (k+2)/2$. It is easy to see that this holds for a generic k , for the order of magnitude depends only on the number p of terms in the product (rather than on k). As for the last term on the right-hand side of (22),

$$\begin{aligned} \mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} V_{t,j/M}^{(z)} \prod_{i=1}^{p-1} V_{t,(j+ik_M)/M}^{(c)} \right]^k &\leq \mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \bar{V}_{t,j/M}^{(z)} \prod_{i=1}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(c)} \right]^k + M^{k/2} \mu_{V^{(z)}}^k \mu_{V^{(c)}}^{k(p-1)} \\ &\quad + \mathbb{E} \left[\frac{\mu_{V^{(z)}}}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=1}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(c)} \right]^k \\ &\quad + \mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \bar{V}_{t,j/M}^{(z)} \mu_{V^{(c)}}^{p-1} \right]^k. \end{aligned}$$

Lemma 1 in Podolskij and Vetter (2009) ensures that $\mu_{V^{(c)}} = O(M^{-1/4})$ and, as a result,

$$M^{k/2} \mu_{V^{(z)}}^k \mu_{V^{(c)}}^{k(p-1)} = O(M^{-k(1-p)/4}) = O(a_M^{k(1-p)/2}),$$

which is of order $O(a_M^{-k/2})$ for $p \geq 2$. Also, Assumption A7 implies that

$$\mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \bar{V}_{t,j/M}^{(z)} \prod_{i=1}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(c)} \right]^k = O(M^{1/2} M^{-1/2} M^{-pk/4}) = O(a_M^{-pk/2}).$$

We now move to the case of infinitely many small jumps. Assumption A8 ensures that, over a day, there are at most M^ψ jumps of size $M^{-\psi}$, with $0 < \psi < 1/2$, over a day. The case of $\psi = 0$ corresponds to the aforementioned case of a finite number of large jumps. As the probability of having p consecutive jumps is $M^{-(1-\psi)p/2}$, $\sqrt{M} \mu_{V^{(z)}}^p = O(M^{1/2} M^{-(1-\psi)p/2} M^{-2\psi}) = o(a_M^{1-p})$, and hence $M^{k/2} \mu_{V^{(z)}}^{kp} = O(a_M^{-k/2})$ for every k as long as $p \geq \frac{3/2-4\psi}{1-\psi}$. In addition,

$$\mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \prod_{i=0}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(z)} \right]^k = O(M^{1/2} M^{-(1-\psi)p/2} M^{-2\psi k}) = O(a_M^{-k/2})$$

for $p \geq \frac{(1/2-4\psi)k+1}{1-\psi}$, whereas Podolskij and Vetter's (2009) Lemma 1 ensures that

$$\sqrt{M} \mu_{V^{(z)}} \mu_{V^{(c)}}^{p-1} = O(M^{1/2} M^{(\psi-1)/2} M^{-2\psi/p} M^{(1-p)/4}).$$

Altogether, this results in $M^{k/2} \mu_{V^{(z)}}^k \mu_{V^{(c)}}^{(p-1)k}$ of order $O(a_M^{-k/2})$ for all k provided that $p \geq 2(1+\psi)$. Finally,

$$\mathbb{E} \left[\frac{1}{\sqrt{M}} \sum_{j=1}^{M-pk_M+1} \bar{V}_{t,j/M}^{(z)} \prod_{i=1}^{p-1} \bar{V}_{t,(j+ik_M)/M}^{(c)} \right]^k = O(M^{1/2} M^{(\psi-1)/2} M^{-2\psi k/p} M^{(1-p)k/4}) = O(a_M^{(1-p)k/2}),$$

which is once more of order $O(a_M^{-k/2})$ for any $p \geq 2$.

(d) Along the same lines as in (b), it is straightforward to show that

$$\mathbb{E} \left| \mu_{2/p}^{-p} PV_{M,t}^{(c)}(2/p, \dots, 2/p) - \frac{1}{M} \frac{1}{24\theta^2} RV_t - \Sigma_t \right|^k = O(a_M^{-k/2}). \quad \blacksquare$$

Proof of Theorem 3: Lemma 1 ensures that the result follows by the same argument regardless of whether there are jumps in both assets or not. Let $N_{A,t,M} = A_{t,M} - A_t$, $N_{V,t,M} = RV_{A,t,M} - IV_{A,t}$, and $N_{B,t,M} = B_{t,M} - B_t$, and then define $\hat{F}_{A|\mathbf{X},M}(a|\mathbf{x})$ analogously to $\hat{F}_{A|\mathbf{X}}(a|\mathbf{x})$, but using $(A_{t,M}, \mathbf{X}_{t,M})$ instead of $(A_t, \mathbf{X}_t)$. Note that

$$\begin{aligned} S_{T,M} &= h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X},M}(A_{t,M}|\mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_{t,M}) \\ &= h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X},M}(A_{t,M}|\mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X},M}(A_{t,M}|\mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) \right]^2 \left[\pi(\mathbf{X}_{t,M}) - \pi(\mathbf{X}_t) \right], \end{aligned}$$

which leads to

$$\begin{aligned} S_{T,M} &= S_T + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X},M}(A_{t,M}|\mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}}(A_{t,M}|\mathbf{X}_{t,M}) \right]^2 \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X}}(A_{t,M}, \mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_t) \right]^2 \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X},M}(A_{t,M}, \mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}}(A_{t,M}, \mathbf{X}_{t,M}) \right] \left[\hat{F}_{A|\mathbf{X}}(A_{t,M}, \mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_t) \right] \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) - \hat{F}_{A|\mathbf{X}_A}(A_{t,M}|\mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X}_A}(A_{t,M}|\mathbf{X}_{A,t,M}) - \hat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right]^2 \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) - \hat{F}_{A|\mathbf{X}_A}(A_{t,M}|\mathbf{X}_{A,t,M}) \right] \left[\hat{F}_{A|\mathbf{X}_A}(A_{t,M}|\mathbf{X}_{A,t,M}) - \hat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right] \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\hat{F}_{A|\mathbf{X},M}(A_{t,M}|\mathbf{X}_{t,M}) - \hat{F}_{A|\mathbf{X}_A,M}(A_{t,M}|\mathbf{X}_{A,t,M}) \right]^2 \left[\pi(\mathbf{X}_{t,M}) - \pi(\mathbf{X}_t) \right] \\ &= S_T + \Delta_{1,T,M} + \Delta_{2,T,M} + \Delta_{3,T,M} + \Delta_{4,T,M} + \Delta_{5,T,M} + \Delta_{6,T,M} + \Delta_{7,T,M}. \end{aligned} \quad (24)$$

We next show that $\Delta_{j,T,M} = o_p(1)$ for every $j = 1, \dots, 7$. Letting $\tilde{f}_{\mathbf{X}_A}(\mathbf{X}_{A,t,M}) = \frac{1}{T} \sum_{s=1}^T \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M})$ yields

$$\Delta_{4,T,M} = h^{q/2} \sum_{t=1}^T \left[\frac{1}{T} \sum_{s=1}^T \frac{\mathbf{1}(A_{s,M} \leq A_{t,M}) \mathbf{K}_b(\mathbf{X}_{A,s,M} - \mathbf{X}_{A,t,M}) - \mathbf{1}(A_s \leq A_t) \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M})}{\tilde{f}_{\mathbf{X}_A}(\mathbf{X}_{A,t,M})} \right]^2 \pi(\mathbf{X}_t) \{1 + o_p(1)\}.$$

Given that $\tilde{f}_{\mathbf{X}_A}(\mathbf{X}_{A,t,M}) > 0$, we ignore the denominator in $\Delta_{4,T,M}$. The leading term in $\Delta_{4,T,M}$ is given by

$$\begin{aligned}\tilde{\Delta}_{4,T,M} &= h^{q/2} \sum_{t=1}^T \left\{ \frac{1}{T} \sum_{s=1}^T \mathbf{1}(A_s \leq A_{t,M}) [\mathbf{K}_b(\mathbf{X}_{A,s,M} - \mathbf{X}_{A,t,M}) - \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M})] \right\}^2 \pi(\mathbf{X}_t) \\ &\quad + h^{q/2} \sum_{t=1}^T \left[\frac{1}{T} \sum_{s=1}^T [\mathbf{1}(A_{s,M} \leq A_{t,M}) - \mathbf{1}(A_s \leq A_{t,M})] \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_t) + \text{cross term} \\ &= \tilde{\Delta}_{4,T,M}^{(1)} + \tilde{\Delta}_{4,T,M}^{(2)} + \text{cross term}.\end{aligned}\tag{25}$$

By noting that $\mathbf{X}_{A,t,M} \in C_{\mathbf{X}_{A,M}} \equiv C_{\mathbf{X}_A} \pm c a_M^{-1/2}$,

$$\begin{aligned}|\tilde{\Delta}_{4,T,M}| &\leq T h^{q/2} \sup_{\mathbf{X}_{A,t,M} \in C_{\mathbf{X}_{A,M}}} \left\{ \frac{1}{T} \sum_{s=1}^T \mathbf{1}(A_s \leq A_{t,M}) [\mathbf{K}_b(\mathbf{X}_{A,s,M} - \mathbf{X}_{A,t,M}) - \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M})] \right\}^2 \pi(\mathbf{X}_t) \\ &= O_p \left(T h^{q/2} a_M^{-1} (1 + b^4 + T^{-1} \ln T b^{-q_A}) \right)\end{aligned}$$

by the same argument as in the proof of Theorem 1 in Corradi et al. (2011). In turn,

$$\tilde{\Delta}_{4,T,M}^{(2)} \leq h^{q/2} \sum_{t=1}^T \left[\frac{1}{T} \sum_{s=1}^T \mathbf{1} \left(A_t - \sup_t |N_{A,t,M}| \leq A_s \leq A_t + \sup_t |N_{A,t,M}| \right) \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_t).$$

Let $\Omega_{T,M} = \left\{ \omega : T^{2/k} a_M^{-1/2} \sup_t |N_{A,t,M}| > c \right\}$. Given Lemma 1,

$$T h^{q/2} \Pr(\Omega_{T,M}) = T h^{q/2} \Pr \left(T^{2/k} a_M^{-1/2} \sup_t |N_{A,t,M}| > c \right) \leq T^2 h^{q/2} T^{-\frac{2}{k}} c^{-k} a_M^{k/2} \mathbb{E} |N_{t,M}|^k = o(1),$$

so that we may proceed conditioning on $\Omega_{T,M}^c$. By the same argument as in the proof of Theorem 1 in

Corradi et al. (2011), letting $d_{T,M} = c T^{2/k} a_M^{-1/2}$ yields

$$\begin{aligned}\tilde{\Delta}_{4,T,M}^{(2)} &\leq h^{q/2} \sum_{t=1}^T \left[\frac{1}{T} \sum_{s=1}^T \mathbf{1}(A_t - d_{T,M} \leq A_s \leq A_t + d_{T,M}) \mathbf{K}_b(\mathbf{X}_{A,s} - \mathbf{X}_{A,t,M}) \right]^2 \pi(\mathbf{X}_t) \\ &= O_p \left(T h^{q/2} d_{T,M}^2 + \ln T h^{q/2} b^{-q_A} d_{T,M} \right) = O_p \left(T^{(k+4)/k} h^{q/2} a_M^{-1} + T^{2/k} \ln T h^{q/2} b^{-q_A} a_M^{-1/2} \right)\end{aligned}$$

for all $\omega \in \Omega_{T,M}^c$. Note that $T^{(k+4)/k} (\ln T) h^{q/2} a_M^{-1}$ is of larger order than both $T^{2/k} (\ln T) h^{q/2} b^{-q_A} a_M^{-1/2}$ and $T h^{q/2} a_M^{-1}$, whereas $h^{q/2} b^{-q_A} a_M^{-1}$ is of larger order than $b^{-q_A} a_M^{-1}$ by Assumption A5(v). This means that $\Delta_{4,T,M} = O_p(T^{(k+4)/k} (\ln T) h^{q/2} a_M^{-1})$. It is also immediate to see that $\Delta_{1,T,M} = O_p(T^{(k+4)/k} h^{q/2} a_M^{-1} + (\ln T) h^{-q} a_M^{-1})$.

As for $\Delta_{2,T,M}$ and $\Delta_{5,T,M}$, they are of smaller probability order than $\Delta_{1,T,M}$ and $\Delta_{4,T,M}$, respectively.

The same applies to the cross terms in $\Delta_{3,T,M}$ and $\Delta_{6,T,M}$. It also follows from Assumption A3 that

$\Delta_{7,T,M} = O_p(h^{-q/2} a_M^{-1/2})$ under $\mathbb{H}_0$. Finally, under the alternative hypothesis $\mathbb{H}_A$, $\Delta_{j,T,M}$ ($j=1, \dots, 6$) are

all of the same probability order as under $\mathbb{H}_0$, whereas $\Delta_{7,T,M} = O_p(T h^{q/2} a_M^{-1/2})$ and $S_T = O_p(T h^{q/2})$. This

ensures the appropriate rate of divergence. ■

Proof of Theorem 4: Denote the infeasible counterparts of $S_{\mathcal{T},M}^*$ and $S_{T,M}$ respectively as

$$S_{\mathcal{T}}^* = h^{q/2} \sum_{t=1}^{\mathcal{T}} \left[\widehat{F}_{A|\mathbf{X}}^*(A_t^*|\mathbf{X}_t^*) - \widehat{F}_{A|\mathbf{X}_A}^*(A_t^*|\mathbf{X}_{A,t}^*) \right]^2 \pi(\mathbf{X}_{t,M}^*)$$

$$S_T = h^{q/2} \sum_{t=1}^T \left[\widehat{F}_{A|\mathbf{X}}(A_t|\mathbf{X}_t) - \widehat{F}_{A|\mathbf{X}_A}(A_t|\mathbf{X}_{A,t}) \right]^2,$$

and let $\tilde{c}_{T,1-\alpha,\eta}^* = \lim_{\mathcal{B} \rightarrow \infty} c_{T,\mathcal{B},1-\alpha+\eta}^* + \eta$, with $c_{T,\mathcal{B},1-\alpha+\eta}^*$ denoting the $(1-\alpha+\eta)$ th percentile of the empirical distribution of $S_{\mathcal{T}}^*$. We must show that $\limsup_{T \rightarrow \infty} \Pr(S_T > \tilde{c}_{T,1-\alpha,\eta}^*) \leq \alpha$ under $\mathbb{H}_0$ and that, if $T_{11}/T \xrightarrow{p} c_{11} > 0$, $\lim_{\eta \rightarrow 0} \limsup_{T \rightarrow \infty} \Pr(S_T > \tilde{c}_{T,1-\alpha,\eta}^*) = \alpha$ under $\mathbb{H}_0$ and $\lim_{T \rightarrow \infty} \Pr(S_T > \tilde{c}_{T,1-\alpha,\eta}^*) = 1$ under $\mathbb{H}_A$. Denote by $\mathbb{E}_*$ and $\mathbb{V}_*$ the mean and variance operators under the bootstrap probability law, respectively. We also let o_{p*} and O_{p*} denote terms respectively converging to zero and bounded under the bootstrap probability law conditionally on the sample. If $T^{-2/(q+8)} < h < T^{-2/(3q)}$ and $b = h^\gamma$, with

$$1 < \gamma < \min \left\{ \frac{9}{10} \left(\frac{3q}{q+8} + \frac{2}{3} \right), \frac{q}{q_A} \right\},$$

then h and b satisfy every rate condition in Assumption A5. In addition, if $3q/(q+8) < \delta < 1$, then Assumption A5 also holds for $\mathcal{T} = T^\delta$. We next prove (ii) before establishing (i) and then (iii).

(ii) Let $I_{11,t}^* = \mathbf{1}(A_{t-1}^* > 0) \mathbf{1}(B_t^* > 0)$ and $\mathcal{T}_{11}^* = \sum_{t=1}^{\mathcal{T}} I_{11,t}^*$. Now, $\mathcal{T}_{11}^*/\mathcal{T} - \mathbb{E}^*(\mathcal{T}_{11}/\mathcal{T}) = o_{p*}(1)$ given that $\mathbb{E}^*(\mathcal{T}_{11}^*/\mathcal{T}) = \frac{1}{\mathcal{T}} \sum_{t=1}^{\mathcal{T}} \mathbf{1}(A_{t-1}^* > 0) \mathbf{1}(B_t^* > 0) + O_p(l_{\mathcal{T}}/\mathcal{T})$. This means that $\mathcal{T}_{11}^*/\mathcal{T} = c_{11} + o_{p*}(1) + o_p(1)$. By the same argument, $\mathcal{T}_{ij}^*/\mathcal{T} = c_{ij} + o_{p*}(1) + o_p(1)$ for all $i, j = \{0, 1\}$. We begin with the case of $\mathcal{T}_{11}^*/\mathcal{T} \rightarrow c_{11} > 0$:

$$\begin{aligned} S_{1,\mathcal{T}}^* &= h^{q/2} \sum_{t=1}^{\mathcal{T}} \left[\widehat{F}_{A|\mathbf{X}}^*(A_t^*|\mathbf{X}_t^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_t^*) \right]^2 \pi(\mathbf{X}_t^*) I_{11,t}^* \\ &\quad + h^{q/2} \sum_{t=1}^{\mathcal{T}} \left[\widehat{F}_{A|\mathbf{X}_A}^*(A_t^*|\mathbf{X}_{A,t}^*) - F_{A|\mathbf{X}_A}(A_t^*|\mathbf{X}_{A,t}^*) \right]^2 \pi(\mathbf{X}_t^*) I_{11,t}^* \\ &\quad - 2h^{q/2} \sum_{t=1}^{\mathcal{T}} \left[\widehat{F}_{A|\mathbf{X}}^*(A_t^*|\mathbf{X}_t^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_t^*) \right] \left[\widehat{F}_{A|\mathbf{X}_A}^*(A_t^*|\mathbf{X}_{A,t}^*) - F_{A|\mathbf{X}_A}(A_t^*|\mathbf{X}_{A,t}^*) \right] \pi(\mathbf{X}_t^*) I_{11,t}^* \\ &= S_{11,T}^* + S_{12,T}^* + S_{13,T}^*, \end{aligned}$$

with

$$\begin{aligned}
S_{11,\mathcal{T}}^* &= h^{q/2} \sum_{t=1}^{\mathcal{T}} \frac{\pi(\mathbf{X}_t^*) I_{11,t}^*}{[\widehat{f}_{\mathbf{X}^*}(\mathbf{X}_t^*)]^2} \left(\frac{1}{\mathcal{T}_{11}^*} \sum_{s=1}^{\mathcal{T}} \mathbf{K}_h(\mathbf{X}_s^* - \mathbf{X}_t^*) [\mathbf{1}(A_s^* \leq A_t^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_s^*)] \right)^2 \\
&\quad + h^{q/2} \sum_{t=1}^{\mathcal{T}} \frac{\pi(\mathbf{X}_t^*) I_{11,t}^*}{[\widehat{f}_{\mathbf{X}^*}(\mathbf{X}_t^*)]^2} \left(\frac{1}{\mathcal{T}_{11}^*} \sum_{s=1}^{\mathcal{T}} \mathbf{K}_h(\mathbf{X}_s^* - \mathbf{X}_t^*) [F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_s^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_t^*)] \right)^2 \\
&\quad + h^{q/2} \sum_{t=1}^{\mathcal{T}} \frac{\pi(\mathbf{X}_t^*) I_{11,t}^*}{[\widehat{f}_{\mathbf{X}^*}(\mathbf{X}_t^*)]^2} \left(\frac{1}{\mathcal{T}_{11}^*} \sum_{s=1}^{\mathcal{T}} \mathbf{K}_h(\mathbf{X}_s^* - \mathbf{X}_t^*) [\mathbf{1}(A_s^* \leq A_t^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_s^*)] \right. \\
&\quad \quad \quad \left. \times \frac{1}{\mathcal{T}_{11}^*} \sum_{s=1}^{\mathcal{T}} \mathbf{K}_h(\mathbf{X}_s^* - \mathbf{X}_t^*) [F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_s^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_s^*)] \right) \\
&= S_{11,\mathcal{T}}^{*(1)} + S_{11,\mathcal{T}}^{*(2)} + S_{11,\mathcal{T}}^{*(3)},
\end{aligned}$$

where

$$S_{11,\mathcal{T}}^{*(1)} = \sum_{t < s < k} \phi^{\dagger*}(t, s, k) + \sum_{s < k} \phi^{\dagger*}(s, k) + o_{p_*}(1),$$

$\phi^{\dagger*}(t, s, k) = \phi^*(t, s, k) + \phi^*(t, k, s) + \phi^*(s, k, t) + \phi^*(s, t, k) + \phi^*(k, s, t) + \phi^*(k, t, s)$, and

$$\begin{aligned}
\phi^*(t, s, k) &= h^{q/2} \frac{\pi(\mathbf{X}_t^*) I_{11,t}^*}{\mathcal{T}_{11}^2 f_{\mathbf{X}^*}(\mathbf{X}_t^*)} \left[\mathbf{K}_h(\mathbf{X}_s^* - \mathbf{X}_t^*) \left(\mathbf{1}(A_s^* \leq A_t^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_s^*) \right) \right. \\
&\quad \quad \quad \left. \times \mathbf{K}_h(\mathbf{X}_k^* - \mathbf{X}_t^*) \left(\mathbf{1}(A_k^* \leq A_t^*) - F_{A|\mathbf{X}}(A_t^*|\mathbf{X}_k^*) \right) \right].
\end{aligned}$$

By the same argument as in the proof of Theorem 5 in Corradi et al. (2012), it turns that $\phi^*(t, s) = \mathbb{E}^*\left(\phi^*(t, s, k)|A_k, \mathbf{X}_k\right) = \phi(t, s) + o_{p_*}(1) + o_p(1)$ and $\phi^*(0) = \mathbb{E}^*\left(\phi^*(t, s, s) + \phi^*(s, t, t)\right) = \phi(0) + o_{p_*}(1) + o_p(1)$. It then follows that $S_{11,\mathcal{T}}^{*(1)} = S_{11,\mathcal{T}}^{(1)} + o_{p_*}(1) + o_p(1)$. Similarly, $S_{11,\mathcal{T}}^{*(j)} = S_{11,\mathcal{T}}^{(j)} + o_{p_*}(1) + o_p(1)$ for $j = 2, 3$ and $S_{1j,\mathcal{T}}^* = S_{1j,\mathcal{T}} + o_{p_*}(1) + o_p(1)$. Along similar lines to the proof of Lemma 1 and Theorem 3, $S_{\mathcal{T},M}^* - S_{\mathcal{T}}^* = o_{p_*}(1)$, establishing the result.

(i) In the absence of jumps, $S_{\mathcal{T},M}^*$ and $S_{T,M}$ approach zero at the same rate given that they use the same bandwidth. Accordingly, it is the uniformity factor $\eta > 0$ that ensures a probability of rejecting the null shrinking to zero, and hence smaller than α .

(iii) The result follows immediately from the fact that, as $\mathcal{T}/T \rightarrow 0$, $S_{\mathcal{T},M}^*$ diverges at a slower rate than $S_{T,M}$ under $\mathbb{H}_A$. ■

Figure 1: Realized measures of quadratic variation, integrated variance and jump contribution
The first plot displays the realized quadratic variation as measured by the realized variance estimators. The second chart exhibits the realized bipower variation estimates of the integrated variances, whereas the third plot depicts the realized jump contribution. To control for market microstructure effects, we compute pre-averaging versions of every realized measure.

quadratic variation

integrated variance

jump contribution

Table 1
Empirical size of the test for jump spillovers, conditioning only on past realized jumps

For each of the 2,000 Monte Carlo replications, we simulate intraday returns from the bivariate jump-diffusion process in (10) and (11) with $\kappa_{BA} = 0$. We then add a Gaussian microstructure noise with variance $c\Sigma_{j,t}$ to the intraday prices before testing for jump spillovers. Panel A considers independent Poisson jumps by setting $\kappa_{jj} \sim N(-\sqrt{\Sigma_{j,t}/2}, \Sigma_{j,t}/2)$, whereas Panel B examines the test behavior in the absence of jumps by fixing κ_{jj} to zero ($j = A, B$). We consider tests at the 1%, 5% and 10% levels of significance for sample sizes of $T \in \{250, 500, 750\}$ daily realized jumps based on $M \in \{78, 390, 780\}$ intraday observations. Critical values rest on 2,000 bootstrap samples of size $\mathcal{T} = \max\{100, \lfloor T^{0.87} \rfloor\}$ using overlapping blocks of length equal to the nearest integer to $\sqrt{T}$.

noise	sample size	1%			5%			10%		
		$M = 78$	$M = 390$	$M = 780$	$M = 78$	$M = 390$	$M = 780$	$M = 78$	$M = 390$	$M = 780$
		Panel A: $\kappa_{jj} \sim N(-\sqrt{\Sigma_{j,t}/2}, \Sigma_{j,t}/2)$ for $j = A, B$								
$c = 0$	$T = 250$	0.0050	0.0055	0.0055	0.0350	0.0370	0.0375	0.0870	0.1150	0.1000
	$T = 500$	0.0060	0.0065	0.0080	0.0375	0.0415	0.0490	0.1155	0.1005	0.1035
	$T = 750$	0.0080	0.0085	0.0085	0.0620	0.0655	0.0575	0.1170	0.1120	0.1155
$c = 0.001$	$T = 250$	0.0060	0.0050	0.0060	0.0320	0.0400	0.0340	0.0940	0.0880	0.1060
	$T = 500$	0.0060	0.0060	0.0085	0.0500	0.0420	0.0460	0.1040	0.0860	0.1140
	$T = 750$	0.0090	0.0080	0.0100	0.0460	0.0460	0.0540	0.1200	0.1160	0.1160
		Panel B: $\kappa_{jj} = 0$ for $j = A, B$								
$c = 0$	$T = 250$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	$T = 500$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	$T = 750$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$c = 0.001$	$T = 250$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	$T = 500$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	$T = 750$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 2
Empirical size of the test for jump spillovers, conditioning on past realized jumps and variance

For each of the 2,000 Monte Carlo replications, we simulate intraday returns from the bivariate jump-diffusion process in (10) and (11) with $\kappa_{BA} = 0$. We then add a Gaussian microstructure noise with variance $c\Sigma_{j,t}$ to the intraday prices before testing for jump spillovers. Panel A considers independent Poisson jumps by setting $\kappa_{jj} \sim N(-\sqrt{\Sigma_{j,t}/2}, \Sigma_{j,t}/2)$, whereas Panel B examines the test behavior in the absence of jumps by fixing κ_{jj} to zero ($j = A, B$). We consider tests at the 1%, 5% and 10% levels of significance for sample sizes of $T \in \{250, 500, 750\}$ daily realized jumps based on $M \in \{78, 390, 780\}$ intraday observations. Critical values rest on 2,000 bootstrap samples of size $\mathcal{T} = \max\{100, \lfloor T^{0.87} \rfloor\}$ using overlapping blocks of length equal to the nearest integer to $\sqrt{T}$.

noise	sample size	1%			5%			10%		
		$M = 78$	$M = 390$	$M = 780$	$M = 78$	$M = 390$	$M = 780$	$M = 78$	$M = 390$	$M = 780$
		Panel A: $\kappa_{jj} \sim N(-\sqrt{\Sigma_{j,t}/2}, \Sigma_{j,t}/2)$ for $j = A, B$								
$c = 0$	$T = 250$	0.0155	0.0140	0.0105	0.0405	0.0480	0.0470	0.0880	0.0775	0.0960
	$T = 500$	0.0115	0.0160	0.0130	0.0565	0.0580	0.0590	0.1165	0.1135	0.1165
	$T = 750$	0.0145	0.0130	0.0155	0.0590	0.0595	0.0890	0.1125	0.1125	0.1115
$c = 0.001$	$T = 250$	0.0080	0.0090	0.0075	0.0475	0.0475	0.0410	0.0825	0.0845	0.0885
	$T = 500$	0.0135	0.0110	0.0150	0.0535	0.0560	0.0570	0.1155	0.0980	0.1045
	$T = 750$	0.0140	0.0150	0.0150	0.0670	0.0678	0.0655	0.1175	0.1150	0.1050
		Panel B: $\kappa_{jj} = 0$ for $j = A, B$								
$c = 0$	$T = 250$	0.0000	0.0000	0.0010	0.0000	0.0010	0.0010	0.0000	0.0015	0.0045
	$T = 500$	0.0000	0.0010	0.0005	0.0000	0.0020	0.0005	0.0010	0.0040	0.0015
	$T = 750$	0.0000	0.0010	0.0000	0.0000	0.0010	0.0005	0.0010	0.0015	0.0005
$c = 0.001$	$T = 250$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	$T = 500$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	$T = 750$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Table 3
Empirical power of the test for jump spillovers

For each of the 2,000 replications, we simulate intraday returns from the bivariate jump-diffusion process in (10) and (11), with $\kappa_{AA} \sim N(-\sqrt{\Sigma_{A,t}/2}, \Sigma_{A,t}/2)$, $\kappa_{BA} \sim \frac{1}{2} N(-\sqrt{\Sigma_{A,t}/2}, \Sigma_{A,t}/2)$ and $\kappa_{BB} \sim \frac{1}{2} N(-\sqrt{\Sigma_{B,t}/2}, \Sigma_{B,t}/2)$. We then add a Gaussian microstructure noise with variance $c\Sigma_{j,t}$ to the intraday prices before testing for jump spillovers. We consider tests at the 1%, 5% and 10% levels of significance for sample sizes of $T \in \{250, 500, 750\}$ daily realized jumps based on $M \in \{78, 390, 780\}$ intraday observations. Critical values rest on 2,000 bootstrap samples of size $T = \max\{100, \lfloor T^{0.87} \rfloor\}$ using overlapping blocks of length equal to the nearest integer to $\sqrt{T}$.

noise	sample size	1%			5%			10%		
		$M = 78$	$M = 390$	$M = 780$	$M = 78$	$M = 390$	$M = 780$	$M = 78$	$M = 390$	$M = 780$
conditioning set includes only past realized jumps										
$c = 0$	$T = 250$	0.2505	0.3175	0.3185	0.5870	0.6085	0.6085	0.7305	0.7310	0.7325
	$T = 500$	0.9490	0.9655	0.9730	0.9980	0.9950	0.9930	0.9725	0.9745	0.9745
	$T = 750$	0.9995	0.9995	1.0000	0.9995	1.0000	1.0000	0.9965	0.9990	0.9985
$c = 0.001$	$T = 250$	0.2690	0.3050	0.3105	0.5690	0.6100	0.6000	0.7305	0.7310	0.7325
	$T = 500$	0.9500	0.9600	0.9730	0.9950	0.9960	0.9970	0.9725	0.9745	0.9745
	$T = 750$	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9965	0.9990	0.9985
conditioning set includes past realized jumps and variance										
$c = 0$	$T = 250$	0.0405	0.1545	0.3240	0.9880	0.9945	0.9900	0.9990	0.9995	1.0000
	$T = 500$	0.0300	0.1205	0.3470	0.9985	1.0000	1.0000	1.0000	1.0000	1.0000
	$T = 750$	0.0200	0.1004	0.3050	0.9985	1.0000	1.0000	1.0000	1.0000	1.0000
$c = 0.001$	$T = 250$	0.0290	0.0255	0.0250	0.8505	0.8120	0.8340	0.9745	0.9360	0.9585
	$T = 500$	0.0870	0.0790	0.0950	0.9985	0.9785	0.9985	0.9995	0.9975	0.9990
	$T = 750$	0.2990	0.2785	0.2823	0.9985	0.9985	0.9985	0.9995	1.0000	0.9995

Table 4
Testing for jump spillovers between ES and SPY

We report rejections at the 1%, 5% and 10% levels of significance respectively by $***$, $**$ and $*$, whereas 0 indicates a nonrejection. Critical values rest on 2,000 bootstrap samples of size $\mathcal{T} = \max\{100, \lfloor T^{0.87} \rfloor\}$ using overlapping blocks of length equal to the nearest integer to $\sqrt{T}$. Apart from the full sample from July 1998 to February 2017, we also run tests for 4 subsamples: July 1998 to December 2003, January 2004 to January 2007, February 2007 to December 2013, January 2014 to February 2017. Finally, we also run the tests controlling additionally for the realized variance in the previous day. We report the latter results only if different after a slash.

null hypothesis	full sample	subsamples			
	1998–2017	1998–2003	2004–2007	2007–2013	2014–2017
jump spillovers from SPY to ES					
$\mathbb{H}_0^{(1)}: \text{ES}_t \perp\!\!\!\perp \text{SPY}_{t-1} \text{ES}_{t-1}$	$**/0$	0	0	0	0
$\mathbb{H}_0^{(2)}: \text{ES}_t \perp\!\!\!\perp \text{SPY}_{t-1} \text{ES}_{t-1}^O$	$***$	$***$	$*/***$	$***$	$***$
$\mathbb{H}_0^{(3)}: \text{ES}_t \perp\!\!\!\perp \text{SPY}_{t-1} (\text{ES}_{t-1}^O + \text{ES}_{t-1})$	$***$	$**$	0	$***/*$	$*/***$
$\mathbb{H}_0^{(4)}: \text{ES}_t^O \perp\!\!\!\perp \text{SPY}_t \text{ES}_t$	$*$	0	0	0	0
contemporaneous jump spillovers between SPY and ES					
$\mathbb{H}_0^{(5)}: \text{ES}_t \perp\!\!\!\perp \text{SPY}_t \text{ES}_{t-1}$	$***$	$***$	$***$	$***$	$***$
$\mathbb{H}_0^{(6)}: \text{ES}_t \perp\!\!\!\perp \text{SPY}_t \text{ES}_{t-1}^O$	$***$	$***$	$***$	$***$	$***$
$\mathbb{H}_0^{(7)}: \text{ES}_t \perp\!\!\!\perp \text{SPY}_t (\text{ES}_{t-1}^O + \text{ES}_{t-1})$	$***$	$***$	$***$	$***$	$***$
$\mathbb{H}_0^{(8)}: \text{SPY}_t \perp\!\!\!\perp \text{ES}_t \text{SPY}_{t-1}$	$***$	$***$	$***$	$***$	$***$
jump spillovers from ES to SPY					
$\mathbb{H}_0^{(9)}: \text{SPY}_t \perp\!\!\!\perp \text{ES}_{t-1} \text{SPY}_{t-1}$	0	0	0	0	0
$\mathbb{H}_0^{(10)}: \text{SPY}_t \perp\!\!\!\perp \text{ES}_{t-1}^O \text{SPY}_{t-1}$	$***$	$0/***$	$**/*$	$***$	$***$
sample size	3,867	1,237	601	1,103	826

References

- Aït-Sahalia, Y., Cacho-Diaz, J., Hurd, T., 2009a, Portfolio choice with jumps: A closed-form solution, *Annals of Applied Probability* 19, 556–584.
- Aït-Sahalia, Y., Cacho-Diaz, J., Laeven, R. J. A., 2015, Modelling financial contagion: Using mutually exciting jump processes, *Journal of Financial Economics* 117, 585–606.
- Aït-Sahalia, Y., Fan, J., Peng, H., 2009b, Nonparametric transition-based tests for diffusions, *Journal of the American Statistical Association* 104, 1102–1116.
- Aït-Sahalia, Y., Jacod, J., 2009b, Estimating the degree of activity of jumps in high frequency financial data, *Annals of Statistics* 37, 2202–2244.
- Aït-Sahalia, Y., Jacod, J., 2012, Identifying the successive Blumenthal-Gettoor indices of a discretely observed process, *Annals of Statistics* 40, 1430–1464.
- Aït-Sahalia, Y., Mykland, P., Zhang, L., 2011, Ultra high frequency volatility estimation with dependent microstructure noise, *Journal of Econometrics* 160, 160–175.
- Aït-Sahalia, Y., Xiu, D., 2016, Increased correlation among asset classes: Are volatility or jumps to blame, or both?, *Journal of Econometrics* 194, 205–219.
- Aldrich, E. M., Lee, S., 2016, Relative spread and price discovery, Discussion Paper 33, Center for Analytical Finance, University of California, Santa Cruz.
- Andrews, D. W., Shi, X., 2013, Inference based on conditional moment inequalities, *Econometrica* 81, 609–666.
- Arcones, M. A., Giné, E., 1992, On the bootstrap of U and V statistics, *Annals of Statistics* 20, 655–674.
- Bajgrowicz, P., Scaillet, O., Treccani, A., 2016, Jumps in high-frequency data: Spurious detections, dynamics and news, *Management Science* 62, 2198–2217.
- Barndorff-Nielsen, O. E., Hansen, P. H., Lunde, A., Shephard, N., 2008, Designing realized kernels to measure the ex-post variation of equity prices in the presence of noise, *Econometrica* 76, 1481–1536.
- Barndorff-Nielsen, O. E., Shephard, N., 2006, Econometrics of testing for jumps in financial economics using bipower variation, *Journal of Financial Econometrics* 4, 1–30.
- Barndorff-Nielsen, O. E., Shephard, N., Winkel, M., 2006, Limit theorems for multipower variation in the presence of jumps, *Stochastic Processes and Their Applications* 116, 796–806.
- Bates, D. S., 2000, Post-'87 crash fears in the S&P 500 futures option market, *Journal of Econometrics* 94, 181–238.
- Ben-David, I., Franzoni, F., Moussawi, R., 2017, Exchange traded funds (ETFs), *Annual Review of Financial Economics*.

- Bollerslev, T., Law, T. H., Tauchen, G., 2008, Risk, jump and diversification, *Journal of Econometrics* 144, 234–256.
- Boswijk, H. P., Laeven, R. J. A., Yang, X., 2018, Testing for self-excitation in jumps, *Journal of Econometrics* 203, 256–266.
- Boudt, K., Zhang, J., 2015, Jump robust two-time scale covariance estimation and realized volatility budgets, *Quantitative Finance* 15, 1041–1054.
- Bowsher, C. G., 2007, Modelling security market events in continuous time: Intensity based, multivariate point process models, *Journal of Econometrics* 141, 876–912.
- Bretagnolle, J., 1983, Lois limites du bootstrap de certaines fonctionnelles, *Annales de l’Institut Henri Poincaré, Section B, Probabilités et Statistiques* 19, 281–296.
- Chakravarty, S., Gulen, H., Mayhew, S., 2004, Informed trading in stock and option markets, *Journal of Finance* 59, 1235–1257.
- Chen, S. N., Dahl, G. B., Khan, S., 2005, Nonparametric estimation and identification of a censored location-scale regression model, *Journal of the American Statistical Association* 100, 212–221.
- Cont, R., Mancini, C., 2011, Nonparametric tests for pathwise properties of semimartingales, *Bernoulli* 17, 781–813.
- Cont, R., Tankov, P., 2004, *Financial Modelling with Jump Processes*, Chapman and Hall, Boca Raton.
- Corradi, V., Distaso, W., Fernandes, M., 2012, International market links and volatility transmission, *Journal of Econometrics* 170, 117–141.
- Corradi, V., Distaso, W., Swanson, N. R., 2011, Predictive inference for integrated volatility, *Journal of the American Statistical Association* 106, 1496–1512.
- Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985, A theory of the term structure of interest rates, *Econometrica* 53, 385–407.
- Das, S. R., Uppal, R., 2004, Systemic risk and international portfolio choice, *Journal of Finance* 59, 2809–2834.
- Duffie, D., Pan, J., Singleton, K., 2000, Transform analysis and asset pricing for affine jump diffusion, *Econometrica* 68, 1343–1376.
- Dungey, M., Erdemlioglu, D., Mateia, M., Yang, X., 2017, Testing for mutually exciting jumps and financial flights in high frequency data, working paper, University of Tasmania, IESEG School of Management and Rutgers University.
- Easley, D., O’Hara, M., Srinivas, P. S., 1998, Volume and stock prices: Evidence on where informed traders trade, *Journal of Finance* 53, 431–465.
- Fleming, J., Ostdiek, B., Whaley, R. E., 1996, Trading costs and the relative rates of price discovery in stock, futures, and option markets, *Journal of Futures Markets* 16, 353–387.

- Hansen, P., Lunde, A., 2006, Realized variance and market microstructure noise, *Journal of Business and Economic Statistics* 24, 127–218.
- Hasbrouck, J., 2003, Intraday price formation in U.S. equity index markets, *Journal of Finance* 58, 2375–2399.
- Hautsch, N., Podolskij, M., 2013, Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence, *Journal of Business and Economic Statistics* 31, 165–183.
- Huang, X., Tauchen, G., 2005, The relative contribution of jumps to total price variance, *Journal of Financial Econometrics* 3, 456–499.
- Jacod, J., Li, Y., Mykland, P. A., Podolskij, M., Vetter, M., 2009, Microstructure noise in the continuous time case: The pre-averaging approach, *Stochastic Processes and Their Application* 119, 2249–2276.
- Jacod, J., Podolskij, M., Vetter, M., 2010, Limit theorems for moving averages of discretized processes plus noise, *Annals of Statistics* 38, 1478–1545.
- Jacod, J., Todorov, V., 2009, Testing for common arrival of jumps in discretely-observed multidimensional processes, *Annals of Statistics* 37, 1792–1838.
- Jacod, J., Todorov, V., 2010, Do price and volatility jump together?, *Annals of Applied Probability* 20, 1425–1469.
- Kawaller, I. G., Koch, P. D., Kock, T. W., 1987, The temporal price relationship between S&P 500 futures and the S&P 500 index, *Journal of Finance* 42, 1309–1329.
- Lee, S. S., Mykland, P. A., 2008, Jumps in financial markets: A new nonparametric test and jump dynamics, *Review of Financial Studies* 21, 2535–2563.
- Li, J., Patton, A., 2018, Asymptotic inference about predictive accuracy using high frequency data, *Journal of Econometrics* 203, 223–240.
- Li, J., Todorov, V., Tauchen, G., 2017a, Jump regressions, *Econometrica* 85, 173–195.
- Li, J., Todorov, V., Tauchen, G., 2017b, Robust jump regressions, *Journal of the American Statistical Association* 112, 332–341.
- Li, J., Todorov, V., Tauchen, G., Chen, R., 2017c, Mixed-scale jump regressions with bootstrap inference, *Journal of Econometrics* 201, 417–432.
- Li, Q., Racine, J. S., 2008, Nonparametric estimation of conditional CDF and quantile functions with mixed categorical and continuous data, *Journal of Business and Economic Statistics* 26, 423–434.
- Maehu, J. M., McCurdy, T. H., 2004, News arrival, jumps dynamics and volatility components for individual stock return, *Journal of Finance* 59, 755–793.
- Mancini, C., Gobbi, F., 2012, Identifying the Brownian covariation from the co-jumps given discrete observations, *Econometric Theory* 28, 249–273.

- Podolskij, M., Vetter, M., 2009, Bipower-type estimation in a noisy diffusion setting, *Stochastic Processes and Their Applications* 119, 2803–2831.
- Politis, D. N., Romano, J. P., Wolf, M., 1999, *Subsampling*, Springer, New York.
- Santa-Clara, P., Yan, S., 2010, Crashes, volatility, and the equity premium: Lessons from S&P 500 options, *Review of Economics and Statistics* 92, 435–451.
- Stephan, J. A., Whaley, R. E., 1990, Intraday price change and trading volume relations in the stock and stock option markets, *Journal of Finance* 45, 191–220.
- Stoll, H. R., Whaley, R. E., 1990, The dynamics of stock index and stock index futures returns, *Journal of Financial and Quantitative Analysis* 25, 441–468.
- Todorov, V., Tauchen, G., 2011, Volatility jumps, *Journal of Business and Economic Statistics* 29, 356–371.
- Wallace, D., Kalev, P. S., Lian, G., 2014, The evolution of price discovery in US equity and derivatives markets, working paper, University of South Australia.
- Wermers, R., Xue, J., 2015, Intraday ETF trading and the volatility of the underlying, working paper, University of Maryland (sponsored by Lyxor Asset Management).
- Yang, J., 2009, Do futures lead price discovery in electronic foreign exchange markets, *Journal of Futures Markets* 29, 137–156.