International market links and volatility transmission

Valentina Corradi — University of Warwick

Walter Distaso — Imperial College London

Marcelo Fernandes — Queen Mary University of London

This version: December 10, 2009

Abstract: This paper proposes a framework to gauge the degree of volatility transmission among international stock markets by deriving tests for conditional independence among daily volatility measures. We suppose that asset prices follow a multivariate jump-diffusion process, and make no parametric assumption on the functional form of the drift, diffusive, and jump components. To check for conditional independence of asset A's daily volatility given asset B's daily volatility, we consider the integrated (relative) squared difference between two nonparametric conditional density estimates. The first estimate considers only information concerning asset A's daily volatility, whereas the second estimate also includes information about asset B's daily volatility. To proxy for the unobservable daily volatility, we employ model-free realized measures, allowing for both microstructure noise and jumps. We establish the asymptotic normality of the test statistic based on realized measures as well as the first-order validity of its bootstrap analog. In addition, we investigate volatility spillovers between the stock markets in China, Japan, and US from January 2000 to December 2005. The empirical evidence for spillovers seem stronger running from China to either Japan or US rather than vice-versa. In contrast to parametric methods, we are also able to infer whether the main channel of transmission within the quadratic variation of asset prices is either through price jumps or through volatility spillovers by varying the realized measure we employ. For instance, price jumps in China affect both the quadratic variation and the options-implied volatility of the S&P 500 index, whereas the primary channel of transmission between China and Japan is through volatility spillovers. Our findings are robust to different realized measures, conditioning sets, and sampling frequency as long as one controls for market microstructure effects.

JEL classification numbers: G12, G15, C14

Keywords: conditional independence, jump-diffusion, noncausality, quadratic variation, realized variance, stochastic volatility.

Acknowledgments: We are indebted to Giovanni Cespa, Jean-Pierre Florens, René Garcia, Liudas Giraitis, Massimo Guidolin, José Ferreira Machado, Michael McCracken, Nour Meddahi, Chris Neely, Alessio Sancetta, Pedro Santa Clara, Enrique Sentana, Ross Valkanov, and seminar participants at Banco de Portugal, Cambridge University, City University, Federal Reserve Bank of St Louis, Fundação Getulio Vargas, Queen Mary, Goethe University, Frankfurt, University of Cyprus, University of Toulouse, Far East and South Asia Meeting of the Econometric Society (Tokyo, August 2009), International Symposium on Risk Management and Derivatives (Xiamen, July 2009), QASS Conference on Financial Econometrics and Realized Volatility (London, June 2009), London-Oxbridge Time Series Workshop (Oxford, October 2008), Granger Centre Conference on Bootstrap and Numerical Methods in Time Series (Nottingham, September 2008), Meeting of the ESRC Research Seminar Series on Nonlinear Economics and Finance (Keele, February 2008), CEA@Cass Conference on Measuring Dependence in Finance (London, December 2007), Brazilian Time Series and Econometrics School (Gramado, August 2007), Meeting of the Brazilian Finance Society (São Paulo, July 2007), Finance and Econometrics Annual Conference (York, May 2007), and London-Oxford Financial Econometrics Study Group (London, November 2006) for valuable comments. We also thank Louis Mercorelli and Anthony Hall for kindly providing the data as well as João Mergulhão and Filip Zikes for excellent research assistance. We are also grateful for the financial support from the ESRC under the grant RES-062-23-0311. The usual disclaimer applies.

1 Introduction

Even though testing for the presence of international market links has a long history in asset pricing (see survey by Roll, 1989), the literature has been gaining momentum since the October 1987 crash. The main interest lies on the analysis of volatility transmission across markets. King and Wadhwani (1990) argue that the strength of international market links depends mainly on volatility. As the latter declines, the correlation between price changes in the different markets also decreases and so market links become weaker. In contrast, international market links become stronger in periods of high volatility. The idea is that, with Bayesian update of beliefs about variances, a common shock to all markets would result in an increase in the perceived variance of any common factor and hence of the correlation.

This paper develops formal statistical tools for testing conditional independence between volatility processes. We propose a nonparametric approach in stark contrast with most papers in the literature. In particular, under the assumption that asset prices follow a multivariate jump-diffusion process, we show how to test whether the conditional distribution of asset A's integrated variance (say, over a day) also depends on asset B's integrated variance. The procedure is nonparametric in that, apart from some mild regularity conditions, we impose no parametric assumption on the functional form of the drift and diffusive term as well as on the presence of leverage or jumps. Broadly speaking, our testing procedure checks whether conditioning also on asset B's integrated variance, rather than exclusively on asset A's past integrated variance, entails a different conditional distribution for asset A's integrated variance. As we do not observe daily variance, we rely on model-free realized measures based on intraday returns. We focus on the conditional distribution for two reasons. First, it allows for nonlinear channels of volatility transmission in contrast to the standard practice of carrying out pointwise analyses based on the conditional mean of the volatility. Second, the distribution of the daily integrated variance is of particular interest for pricing variance swap contracts (Carr, Geman, Madan and Yor, 2005).

The asymptotic theory we develop proceeds in three steps. First, we establish the asymptotic normality of the unfeasible statistics based on unobservable integrated variances. Second, we provide conditions on the rate of growth of intraday observations relative to the number of days under which the feasible statistic, based on realized measures, is asymptotically equivalent to its unfeasible counterpart. Our setting is general enough to allow for microstructure noise as well as for jumps. Third, we establish the first-order validity of bootstrap-based critical values based on the m out of n (henceforth, moon) bootstrap (Bretagnolle, 1983; Bickel, Götze and van Zwet, 1997). Bootstrap based inference is typically more robust to variations in the bandwidth as the latter plays roughly the same role on both the original and bootstrap statistics. Monte Carlo simulations indeed reveal that the moon bootstrap works reasonably well even in relatively small samples.

We investigate the links between international stock markets using intraday data from China, Japan, and US from January 2000 to December 2005. By employing different realized measures, we are able to infer whether the main channel of transmission within the quadratic variation of asset prices is either through price jumps or through volatility spillovers. Our results evince a link between Japan and US in both directions mainly through price jumps. Albeit the transmission between China and Japan also flows in a bidirectional manner, it is mainly about volatility spillovers rather than through jumps in the stock market indices. In addition, we find that price jumps in the Shanghai stock market affect both the quadratic variation and the options-implied volatility of the S&P 500 index, even though we observe no influence from the S&P 500 index on China. Interestingly, we show that it is sometimes crucial to account for microstructure noise for the latter typically obscures the evidence concerning the transmission mechanism. This is especially true for China given that there are several changes in the trading regulations at the Shanghai Stock Exchange within the sample period.

To ensure that the transmission running from China and Japan to US is genuine rather than spurious due to time-series persistence, we redo the analysis adding an extra control, namely, the VIX index. The latter is a market volatility index from the Chicago Board Options Exchange (CBOE) that gauges the options-implied volatility of the S&P500 index. There are no qualitative changes in our main results in that Japan and China still matter. In addition, we also find similar results if we redo the analysis replacing the S&P 500 index realized measures with the VIX index. This suggests that investors form their expectations about the future volatility of the S&P 500 index accounting for the stock market information coming from Japan and China.

Finally, we also carry out a similar analysis considering the realized measures over one-hour intervals (rather than over a day) to check for a swifter reaction time. We find that the realized measures over the last hour of trading in Japan significantly affect the realized measures over the first hour of trading in US. The same applies in the opposite direction, though transmission seems to work specifically through volatility spillovers. As for China, its realized measures remain not only conditionally independent of the realized measures of the S&P 500 index, but also exercising a stronger influence in Japan than vice-versa. In contrast, despite of the daily repercussion, we find no evidence of a China effect in the first hour of trading in US.

There are several papers in the literature that carry out similar, though mostly parametric, analyses of volatility transmission across international stock markets. Engle and Ng (1988), Hamao, Masulis and Ng (1990), Engle, Ito and Lin (1990), King, Sentana and Wadhwani (1994), Lin, Engle and Ito (1994), Karolyi (1995), and Wongswan (2006) employ multivariate GARCH models to show that volatility spillovers indeed occur across foreign exchange markets as well as international stock markets, notably, between Japan, UK and US. In contrast, Cheung and Ng (1996), Hong (2001), Pantelidis and Pittis (2004), Sensier and van Dijk (2004), and van Dijk, Osborne and Sensier (2005) propose simple tests of noncausality in variance based on the cross-correlation between leads and lags of squared GARCH-standardized residuals. More recently, Gourieroux and Jasiak (2007) address causality in variance (or even in higher-order moments) by means of approximate conditional log-Laplace transforms of compound autoregressive processes. The testing strategy of the above papers mainly differs from ours in three respects. First, they assume a discrete-time data generating mechanism in which the conditional variance is a measurable function of past asset returns. In contrast, we assume that asset prices follow a multivariate jump-diffusion process and then test for spillovers considering the quadratic variation. Second, their tests are sensitive to misspecifications in the conditional mean and variance equations, whereas the nonparametric nature of our tests minimizes any misspecification risk. Third, they do not contemplate any sort of nonlinear dependence between variances as opposed to our testing procedures, whose nontrivial power against nonlinear channels of volatility transmission results from looking at the whole volatility distribution.

The testing strategy put forth by Diebold and Yilmaz (2009, 2010) is the closest to ours. They examine volatility spillovers by means of a VAR approach using range-based measures of volatility. The latter belongs to the class of realized measures robust to market microstructure noise and hence provides a consistent nonparametric estimator of the quadratic variation of the underlying diffusion process. Their approach differs from ours in that they focus on linear spillovers in the conditional mean of the quadratic variation of asset prices, whereas our nonparametric test has nontrivial power against nonlinear spillovers in any moment of the distribution. In fact, the spillovers we evince running from China to either Japan or US are stronger in the tails.

The remainder of this paper ensues as follows. Section 2 describes the data generating process we assume for asset prices and discusses the hypotheses of interest. Section 3 establishes the asymptotic normality of the unfeasible statistic based on integrated variances, whereas Section 4 derives the asymptotic equivalence of feasible test statistics that substitute realized measures for integrated variances. Section 5 first establishes the first-order validity of bootstrap-based critical values and then examines their finite-sample properties of the resulting test through Monte Carlo simulations. Section 6 investigates whether there are significant volatility spillovers across international stock markets using data from China, Japan, and US. Section 7 offers some closing remarks, whereas the Appendix collects all technical proofs.

2 Volatility transmission: Setup and issues

In this section, we discuss how to analyze volatility transmission through nonparametric tests of conditional independence. For notational simplicity, we restrict attention to testing whether the daily variance of asset B affects the dynamic of asset A's daily variance. It is straightforward to consider more than two assets, though the usual concern with the curse of dimensionality applies.

Let $p_{A,t}$ and $p_{B,t}$ denote the log-prices of assets A and B with instantaneous (stochastic) volatility given by $\sigma^2_{A,t}$ and $\sigma^2_{B,t}$, respectively. To fix ideas, we consider a simple example in which asset prices follow a multivariate stochastic-volatility process with jumps:

$$\begin{pmatrix} \mathrm{d}p_{A,t} \\ \mathrm{d}p_{B,t} \\ \mathrm{d}\sigma^2_{A,t} \\ \mathrm{d}\sigma^2_{B,t} \end{pmatrix} = \begin{pmatrix} \mu_{A,t} \\ \mu_{B,t} \\ b_{1,A}(\sigma^2_{A,t}, \sigma^2_{B,t}) \\ b_{1,B}(\sigma^2_{A,t}, \sigma^2_{B,t}) \end{pmatrix} \mathrm{d}t + \begin{pmatrix} \mathrm{d}J_{1,t} \\ \mathrm{d}J_{2,t} \\ \mathrm{d}J_{3,t} \\ \mathrm{d}J_{4,t} \end{pmatrix} + \begin{pmatrix} \sqrt{1-\rho_A^2}\,\sigma_{A,t} \\ \sigma_{BA,t} \\ 0 \\ 0 \end{pmatrix} \mathrm{d}W_{1,t} + \begin{pmatrix} \sigma_{AB,t} \\ \sqrt{1-\rho_B^2}\,\sigma_{B,t} \\ 0 \\ 0 \end{pmatrix} \mathrm{d}W_{2,t}$$
$$+ \begin{pmatrix} \rho_A\,\sigma_{A,t} \\ 0 \\ b_{2,A}(\sigma^2_{A,t}, \sigma^2_{B,t}) \\ b_{2,BA}(\sigma^2_{A,t}, \sigma^2_{B,t}) \end{pmatrix} \mathrm{d}W_{3,t} + \begin{pmatrix} 0 \\ \rho_B\,\sigma_{B,t} \\ b_{2,AB}(\sigma^2_{A,t}, \sigma^2_{B,t}) \\ b_{2,B}(\sigma^2_{A,t}, \sigma^2_{B,t}) \end{pmatrix} \mathrm{d}W_{4,t}, \qquad (1)$$

where $(W_{1,t}, \ldots, W_{4,t})$ are independent standard Brownian motions and $(J_{1,t}, \ldots, J_{4,t})$ are jump processes. The latter is such that $\mathrm{d}J_{i,t} = \int_{\mathcal{A}} c_i(u)\, N_i(\mathrm{d}t, \mathrm{d}u)$, for $i = 1, \ldots, 4$, where $N_i([t_1, t_2], \mathcal{A})$ is a Poisson measure that counts the number of jumps between t_1 and t_2, whose size $c_i(u)$ is an iid random variable in $\mathcal{A}$. The jump process is such that, over a finite time span, there is only a finite number of jumps. Although we allow for leverage effects between asset prices and their own stochastic volatility through the correlation coefficients ρ_A and ρ_B, for the sake of simplicity, we assume away cross-leverage effects by imposing zero correlation between one asset price and the stochastic volatility of the other asset.[1] We also assume that the drift components $\mu_{A,t}$ and $\mu_{B,t}$ follow predictable processes. This is not stringent for their role is asymptotically negligible in the context of volatility transmission. Finally, as standard in

[1] Cross-leverage effects would add other possible channels of volatility transmission between assets A and B. Although this would bring about additional misspecification risk in any parametric approach to test for volatility transmission, it does not affect the nonparametric procedure we propose.

multivariate stochastic volatility models, we suppose for simplicity that asset prices do not directly affect volatility.

It is possible to decompose the quadratic variation process $\langle\cdot\rangle_t$ of a given asset price, say p_A, over the time interval $[0,t]$ 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, $\langle p_A\rangle_t = \langle p_A^C\rangle_t + \langle p_A^D\rangle_t$, where $\langle p_A^D\rangle_t \equiv \int_0^t \int_{\mathcal{A}} c_1^2(u)\, N_1(\mathrm{d}s, \mathrm{d}u)$ and $\langle p_A^C\rangle_t$ corresponds to the integrated variance over the time interval $[0,t]$, namely, $IV_{A,t} = \int_0^t \sigma_{A,s}^2\,\mathrm{d}s + \int_0^t \sigma_{AB,s}^2\,\mathrm{d}s$. Now, recall that

$$\begin{aligned}\int_0^t \sigma_{A,s}^2\,\mathrm{d}s = & \int_0^t \left(\int_0^s b_{1,A}(\sigma_{A,u}^2, \sigma_{B,u}^2)\,\mathrm{d}u\right)\mathrm{d}s + \int_0^t \left(\int_0^s \mathrm{d}J_{3,u}\right)\mathrm{d}s \\ & + \int_0^t \left(\int_0^s b_{2,A}(\sigma_{A,u}^2, \sigma_{B,u}^2)\,\mathrm{d}W_{3,u}\right)\mathrm{d}s + \int_0^t \left(\int_0^s b_{2,AB}(\sigma_{A,u}^2, \sigma_{B,u}^2)\,\mathrm{d}W_{4,u}\right)\mathrm{d}s\end{aligned}$$

and that $\langle p_A^C, p_B^C\rangle_t = \int_0^t (\sigma_{A,s}\sigma_{BA,s} + \sigma_{B,s}\sigma_{AB,s})\,\mathrm{d}s$ is the integrated covariation between p_A and p_B over the $[0,t]$ interval. It is easy to see that IV_B does not affect $IV_{A,t}$ if and only if

- (i) $\sigma_{AB,s} = 0$ a.s.
- (ii) $J_{3,s}$ is independent of $J_{4,t}$
- (iii) $b_{1,A}(\sigma_{A,u}^2, \sigma_{B,u}^2) = b_{1,A}(\sigma_{A,u}^2)$ a.s.
- (iv) $b_{2,A}(\sigma_{A,u}^2, \sigma_{B,u}^2) = b_{2,A}(\sigma_{A,u}^2)$ a.s.
- (v) $b_{2,AB}(\sigma_{A,u}^2, \sigma_{B,u}^2) = 0$ a.s.

If any of the above conditions fails to hold, then $IV_{A,t}$ remains dependent upon $IV_{B,t}$ even after conditioning on its own past values. Regardless of whether condition (i) holds, volatility interdependence may arise even in the case that volatility is a measurable function of past asset prices due to a violation of condition (iii), which reduces to $b_{1,A}(p_{A,u}, p_{B,u}) = b_{1,A}(p_{A,u})$ almost surely. It thus turns out that it does not suffice, nor it is necessary, to impose that the cross-variation process $\langle p_A^C, p_B^C\rangle_t$ is zero almost surely. In principle, it is possible to test directly whether conditions (i) to (v) hold if one is ready to specify the parametric functional forms of the drift, diffusive, and jump terms. The outcome would however depend heavily on the correct specification of the data generating process in (1). To minimize the risk of misspecification, we take a nonparametric route.

Our goal is to check whether the daily variance of asset B helps predict the daily variance of asset A.[2] We thus formulate a testing procedure that focuses on the density restrictions implied by conditional

[2] One could actually develop a nonparametric test to check conditions (iii) to (v) by estimating the diffusive terms using realized measures of the spot volatility. However, the latter estimators converge at a much slower rate than the realized measures of integrated variance (see, e.g., Bandi and Renò, 2008a,b). This would affect not only the power of the testing procedure, but also the ability to control for measurement error.

independence:

$$\mathbb{H}_0: \quad f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t},\boldsymbol{IV}^{(q_B)}_{B,t+k}}\left(y \,\middle|\, \boldsymbol{IV}^{(q_A)}_{A,t}, \boldsymbol{IV}^{(q_B)}_{B,t+k}\right) = f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t}}\left(y \,\middle|\, \boldsymbol{IV}^{(q_A)}_{A,t}\right) \quad \text{a.s. for all } y, \tag{2}$$

where $f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t}}$ and $f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t},\boldsymbol{IV}^{(q_B)}_{B,t+k}}$ denote the conditional density of $IV_{A,t+1}$ given $\boldsymbol{IV}^{(q_A)}_{A,t}$ and $\left(\boldsymbol{IV}^{(q_A)}_{A,t}, \boldsymbol{IV}^{(q_B)}_{B,t+k}\right)$, with $\boldsymbol{IV}^{(q_A)}_{A,t} \equiv (IV_{A,t}, \ldots, IV_{A,t-q_A+1})'$ and $\boldsymbol{IV}^{(q_B)}_{B,t+k} \equiv (IV_{B,t}, \ldots, IV_{B,t+k-q_B+1})'$ standing for vectors of dimension q_A and q_B concerning the information about the integrated variances of assets A and B, respectively. We allow for $k \in \{0, 1\}$ so as to control for time differences between the markets under consideration.[3] As usual, we define the alternative hypothesis as the negation of the null hypothesis.

In general, the integrated variance does not follow a finite-order Markov process.[4] This means that, to test for noncausality in variance, one would have to let the number of conditioning variables (q_A and q_B) to increase with the sample size. This is unfeasible due to the curse of dimensionality and hence we consider the less ambitious null of conditional independence by fixing the number of conditioning variables in (2) to a finite (and, possibly, small) figure.[5]

To implement a nonparametric test for $\mathbb{H}_0$, we propose a statistic that gauges the discrepancy between the nonparametric estimates of the density functions that appear in (2). In particular, our test statistic hinges on the sample counterpart of the following integrated square relative distance

$$\int \left[\frac{f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t},\boldsymbol{IV}^{(q_B)}_{B,t+k}}(y|\boldsymbol{x}^{(q_A)},\boldsymbol{x}^{(q_B)}) - f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t}}(y|\boldsymbol{x}^{(q_A)})}{f_{IV_{A,t+1}|\boldsymbol{IV}^{(q_A)}_{A,t}}(y|\boldsymbol{x}^{(q_A)})}\right]^2 \pi(y, \boldsymbol{x}^{(q_A)}, \boldsymbol{x}^{(q_B)})\, \mathrm{d}y\, \mathrm{d}\boldsymbol{x}^{(q_A)}\, \mathrm{d}\boldsymbol{x}^{(q_B)}, \tag{3}$$

where $\boldsymbol{IV}^{(q_i)}_{i,t} \equiv (IV_{i,t}, \ldots, IV_{i,t-q_i+1})$ with $IV_{i,t}$ denoting asset i's integrated variance over day t ($i = A, B$). We employ a weighting scheme $\pi(\cdot,\cdot,\cdot)$ so as to avoid the lack of precision that afflicts conditional density estimation in areas of low density of the conditioning variables. The integrated square distance that we adopt in (3) is convenient because it facilitates the derivation of the asymptotic theory. Bickel and Rosenblatt (1973), Aït-Sahalia (1996), Aït-Sahalia, Bickel and Stoker (2001), Amaro de Matos and Fernandes (2007), Aït-Sahalia, Fan and Jiang (2009a) and Aït-Sahalia, Fan and Peng (2009b) use similar squared distance measures, though one could also employ entropic pseudo-distance measures as in, e.g., Robinson (1991) and Hong and White (2004).

[3] For instance, as the Tokyo Stock Exchange closes before the opening of the New York Stock Exchange, one may condition on the same day information ($k = 1$) rather than on information from the previous trading day ($k = 0$).

[4] Meddahi (2003) shows, for instance, that the CIR specification entails an ARMA(1,1) process for the integrated variance.

[5] In the empirical application in Section 6, we add the VIX index as an extra control so as to accommodate for the non-Markovian nature of the data. Alternatively, one could adapt our asymptotic theory to deal with dimension reduction techniques as in, e.g., Hall and Yao (2005) and Fan, Peng, Yao and Zhang (2009).

We derive the limiting distribution of the test statistic in (3) in two steps. Section 3 first establishes the asymptotic theory for the unfeasible statistic that considers density estimators based on the unobservable series of integrated variances. We then show in Section 4 that the feasible statistic, with density estimators rooted in some observable realized measure, is asymptotically equivalent to its unfeasible counterpart. Finally, we also study how the rate at which M grows with T must vary in order to ensure that this results holds across different realized measures.

3 Asymptotic theory for the unfeasible case

As integrated volatility is a strictly positive random variable, the Nadaraya-Watson kernel estimator suffers of the well-known boundary problem. To overcome this, we shall proceed by employing local polynomial estimators, which do not suffer of boundary problems. Unfortunately, as will be outlined in detail below, this route is viable only for the case in which we have at most three conditioning variables.

To simplify notation, we denote by Y_t the integrated variance of interest $IV_{A,t+1}$, whereas we denote the conditioning vectors $\boldsymbol{IV}_{A,t}^{(q_A)}$ and $\left(\boldsymbol{IV}_{A,t}^{(q_A)}, \boldsymbol{IV}_{B,t+k}^{(q_B)}\right)$ respectively by $\boldsymbol{X}_t^{(q_A)}$ and $\boldsymbol{X}_t^{(q)}$, with $q = q_A + q_B$ standing for the higher dimension. The null hypothesis of conditional independence in (2) now reads

$$\mathbb{H}_0: \quad f_{Y|\boldsymbol{X}^{(q)}}(y \,|\, \boldsymbol{x}^{(q)}) = f_{Y|\boldsymbol{X}^{(q_A)}}(y \,|\, \boldsymbol{x}^{(q_A)}) \qquad \text{for all } (y, \boldsymbol{x}^{(q)}). \tag{4}$$

We employ local linear smoothing to estimate both the right- and left-hand sides of (4). The sample analog of (3) then is

$$\sum_{t=1}^{T} \left[\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t \,|\, \boldsymbol{X}_t^{(q)}) - \widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t \,|\, \boldsymbol{X}_t^{(q_A)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t \,|\, \boldsymbol{X}_t^{(q_A)})} \right]^2 \pi(Y_t, \boldsymbol{X}_t^{(q)}), \tag{5}$$

where the conditional density estimates $\widehat{f}_{Y|\boldsymbol{X}^{(q)}}$ and $\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}$ derive from local linear smoothing using different sets of bandwidths.

Denote by $\widehat{\boldsymbol{\beta}}_T(y, \boldsymbol{x}^{(q)}) = \left(\widehat{\beta}_{0T}(y, \boldsymbol{x}^{(q)}), \widehat{\beta}_{1T}(y, \boldsymbol{x}^{(q)}), \ldots, \widehat{\beta}_{qT}(y, \boldsymbol{x}^{(q)})\right)'$ with $\boldsymbol{x}^{(q)} = (x_1, \ldots, x_q)$ the argument that minimizes

$$\frac{1}{T} \sum_{t=1}^{T} \Big[K_b(Y_t - y) - \beta_0 - \beta_1(X_{1t} - x_1) - \ldots - \beta_q(X_{qt} - x_q) \Big]^2 \prod_{j=1}^{q} W_{h_q}(X_{jt} - x_j),$$

where $K_b(u) = b^{-1} K(u/b)$ and $W_{h_q}(u) = h_q^{-1} W(u/h)$ are symmetric kernels. The local linear estimator of the conditional density function $f_{Y|\boldsymbol{X}^{(q)}}$ is given by $\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(y \,|\, \boldsymbol{x}^{(q)}) = \widehat{\beta}_{0T}(y, \boldsymbol{x}^{(q)})$. The local linear estimator $\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}$ of the lower dimensional conditional density is analogous for $\boldsymbol{x}^{(q_A)} = (x_1, \ldots, x_{q_A})$.

To establish the limiting behavior of the test statistic in (5), we shall rely on the following assumptions.

Assumption A1: The product kernels $\boldsymbol{W}(\boldsymbol{u}) = \prod_{j=1}^{q} W(u_j)$ and $\widetilde{\boldsymbol{W}}(\boldsymbol{u}) = \prod_{j=1}^{q_A} W(u_j)$ rest on a symmetric, nonnegative, continuous univariate kernel W of second order with bounded support $[-\Delta, \Delta]$ for $1 \leq j \leq q$. The kernel W is also at least twice differentiable on the interior of its support. The symmetric kernel K is of order $s \geq 2$ (even integer) and at least twice differentiable on the interior of its bounded support $[-\Delta, \Delta]$.

Assumption A2: The density functions $f_{Y|\boldsymbol{X}^{(q)}}(y \mid \boldsymbol{x}^{(q)})$ and $f_{Y\boldsymbol{X}^{(q)}}(y, \boldsymbol{x}^{(q)})$ are r-times continuously differentiable in $(y, x^{(q)})$ with bounded derivatives and with $r \geq s$. The same condition also holds for the lower-dimensional density functions $f_{Y|\boldsymbol{X}^{(q_A)}}(y \mid \boldsymbol{x}^{(q_A)})$ and $f_{Y\boldsymbol{X}^{(q_A)}}(y, \boldsymbol{x}^{(q_A)})$.

Assumption A3: The weighting function $\pi\left(y, \boldsymbol{x}^{(q)}\right)$ is continuous and integrable, with second derivatives in a compact support.

Assumption A4: The stochastic process $\left(Y_t, \boldsymbol{X}_t^{(q)}\right)$ is strictly stationary and β-mixing with $\beta_\tau = O\left(\rho^\tau\right)$, where $0 < \rho < 1$.

Assumptions A1 to A4 are quite standard in the literature on local linear smoothing (see, e.g., Fan, Yao and Tong, 1996) and hence we only briefly discuss them in what follows. The reason why Assumption A1 rules out higher-order kernels for W is evident from, e.g., Equation (2.3) in Fan and Masry (1997). Assumptions A2 and A3 require that the weighting scheme and the density functions are both well defined and smooth enough to admit functional expansions. Assumption A4 restricts the amount of data dependence, requiring that the stochastic process is absolutely regular with geometric decay rate. Alternatively, one could assume $\alpha-$mixing conditions as in Aït-Sahalia et al. (2009b) and Gao and Hong (2008), though the conditions under which the quadratic variation of a jump-diffusion process satisfies Assumption A4 are quite weak (see discussion in Corradi, Distaso and Swanson, 2009). See also Chen, Linton and Robinson (2001) for some advantages of the $\beta-$mixing assumption relative to the $\alpha-$mixing condition in the context of nonparametric density estimation.

The scaled and centered version of (5) reads as

$$\widehat{\Lambda}_T = \widehat{\Omega}_T^{-1} \left\{ \begin{array}{l} h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left[\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - \widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} \right]^2 \pi(Y_t, \boldsymbol{X}_t^{(q)}) \\ \qquad - h_q^{-q/2}\, b^{-1/2}\, \widehat{\mu}_{1,T} - h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\, \widehat{\mu}_{2,T} + 2\, h_q^{q/2 - q_A}\, b^{-1/2}\, \widehat{\mu}_{3,T} \end{array} \right\}, \tag{6}$$

where $\widehat{\mu}_{1,T}$, $\widehat{\mu}_{2,T}$, $\widehat{\mu}_{3,T}$, and $\widehat{\Omega}_T$ are consistent estimators of the asymptotic bias terms and variance, respectively (see Appendix A for definitions).

To ensure the asymptotic standard normality of $\widehat{\Lambda}_T$ in (6), we must impose some conditions on the rates at which the bandwidths shrink to zero in Lemmata 1 to 2 (see Appendix B). It turns out that there are no bandwidth rates that jointly meet the conditions (ii) and (iii) in Lemma 1 for $q > 2$ if K is of second order ($s = 2$) and for $q > 3$ if K is of higher order ($s \geq 4$). The main snag is that one cannot increase the order of the kernel function W without incurring into problems of asymptotic singularity within local linear smoothing. For instance, restricting attention to at most two conditioning variables and to second-order kernels, the bandwidth conditions in Lemmata 1 and 2 hold if $b = O\left(T^{-1/5}\right)$ and $h_d = o\left(T^{-1/m_d}\right)$ for $d \in \{1, 2\}$, with $\frac{5}{2} < m_1 \leq 5$ and $5 < m_2 \leq \frac{50}{9}$. If there are three conditioning variables, we must then employ a higher-order kernel for K, say $s = 4$, and hence one could impose that $b = O\left(T^{-1/9}\right)$ and $h_d = o\left(T^{-1/m_d}\right)$ for $d \in \{1, 2, 3\}$, with $\frac{27}{15} < m_1 \leq \frac{81}{17}$, $\frac{54}{15} < m_2 \leq \frac{90}{17}$, and $\frac{81}{15} < m_3 \leq \frac{99}{17}$, so as to meet the conditions in Lemmata 1 and 2. It is easy to appreciate that the degree of undersmoothing increases with the dimensionality of the estimation problem.

We are now ready to state our main result concerning the asymptotic behavior of the normalized test statistic in (6) under both the null and alternative hypotheses.

Theorem 1: Let Assumptions A1 to A4 hold as well as the bandwidth conditions (i) to $(viii)$ in Lemmata 1 and 2. It follows for $q \leq 3$ and $q_A \leq 2$ that

(i) Under the null hypothesis $\mathbb{H}_0$, $\widehat{\Lambda}_T \xrightarrow{d} N(0,1)$.

(ii) Under the alternative hypothesis $\mathbb{H}_A$, $\Pr\left(T^{-1}\, h_q^{-q/2}\, b^{-1/2} \left|\widehat{\Lambda}_T\right| > \varepsilon\right) \longrightarrow 1$ for any $\varepsilon > 0$.

Part (i) of Theorem 1 provides the means to compute asymptotic critical values for the test, whereas (ii) shows consistency. If one restricts attention to the case in which $q = 1$ and $q_A = 0$, the above result follows almost immediately from Aït-Sahalia et al.'s (2009b) Corollary to Theorem 1. Yet, even in this simple case, it is necessary to account for the bias component that arises due to the nonparametric estimation of the lower-dimensional model.

We now deal with higher dimensions (i.e., $q > 3$) by employing a Nadaraya-Watson estimator based on higher-order kernels. More specifically, consider

$$\bar{f}_{Y|\boldsymbol{X}^{(q)}}(y|\boldsymbol{x}^{(q)}) = \frac{\bar{f}_{Y,\boldsymbol{X}^{(q)}}(y, \boldsymbol{x}^{(q)})}{\bar{f}_{\boldsymbol{X}^{(q)}}(\boldsymbol{x}^{(q)})} = \frac{\frac{1}{T}\sum_{t=1}^{T} \bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_t^{(q)} - \boldsymbol{x}^{(q)})\, K_b(Y_t - y)}{\frac{1}{T}\sum_{t=1}^{T} \bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_t^{(q)} - \boldsymbol{x}^{(q)})},$$

where $\bar{\boldsymbol{W}}_{h_q}(\boldsymbol{u}) = h_q^{-p} \prod_{j=1}^{q} \bar{W}(u_j/h_q)$. Define $\bar{f}_{Y|\boldsymbol{X}^{(q_A)}}$ analogously using the product kernel given by $\widetilde{\bar{\boldsymbol{W}}}_{h_{q_A}}(\boldsymbol{u}) = h_{q_A}^{-q_A} \prod_{j=1}^{q_A} \bar{W}(u_j/h_{q_A})$. We next modify Assumption A1 to accommodate higher-order kernels.

Assumption A5: The kernel functions K and $\bar{W}$ are of order $s > 2$ (even integer), symmetric, continuous, and at least twice differentiable on the interior of their bounded support $[-\Delta, \Delta]$.

The kernel-based test statistic is essentially analogous to the one based on local linear smoothing:

$$\bar{\Lambda}_T = \bar{\Omega}_T^{-1} \left\{ \begin{array}{l} h_q^{q/2}\, b^{1/2} \sum_{t=1}^T \left[\frac{\bar{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - \bar{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{\bar{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} \right]^2 \pi(Y_t, \boldsymbol{X}_t^{(q)}) \\ \qquad - h_q^{-q/2}\, b^{-1/2}\, \bar{\mu}_{1,T} - h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\, \bar{\mu}_{2,T} + 2\, h_q^{q/2-q_A}\, b^{-1/2}\, \bar{\mu}_{3,T} \end{array} \right\}. \tag{7}$$

We provide expressions for the bias and scaling terms in Appendix A. As before, to ensure the asymptotic standard normality of $\bar{\Lambda}_T$ in (7), we must impose the bandwidth conditions in Lemmata 4 and 5 (see Appendix). For instance, in the event that $b = O\left(T^{-1/(2s+1)}\right)$ and $h_d = o\left(T^{-1/m_d}\right)$, the bandwidth conditions in Lemma 5 require that $\frac{17}{15}\, d < m_d \leq \frac{9}{17}\, (16 + d)$ for $d \in \{q, q_A\}$ if one confines attention to fourth-order kernels ($s = 4$). Once again, this implies a certain degree of undersmoothing, which increases with the dimensionality of the conditioning vector of state variables. The next result documents the asymptotic standard normality of the kernel-based test statistic in (7) under the null and consistency of the resulting test.

Theorem 2: Let the bandwidth conditions (i) to $(viii)$ in Lemmata 4 and 5 hold as well as Assumptions A2 to A5. It then follows that

(i) Under the null hypothesis $\mathbb{H}_0$, $\bar{\Lambda}_T \xrightarrow{d} N(0,1)$.

(ii) Under the alternative hypothesis $\mathbb{H}_A$, $\Pr\left(T^{-1}\, h_q^{-q/2}\, b^{-1/2}\, \left|\bar{\Lambda}_T\right| > \varepsilon\right) \longrightarrow 1$ for any $\varepsilon > 0$.

Theorems 1 and 2 form the basis for asymptotically locally strictly unbiased tests for the conditional independence null $\mathbb{H}_0$ in (4) based on local linear and kernel smoothing, respectively. The conditions under which we derive both testing procedures also clarify that the kernel-based test works in a more general environment than the local linear variant due to the more stringent limitations with respect to the dimensionality of the conditioning state vector in the latter. To alleviate the boundary bias that haunts kernel smoothing, one could always take the log of the daily variances before testing for conditional independence as an alternative to weighting down by means of $\pi(Y_t, \boldsymbol{X}_t^{(q)})$ any realized measure close to zero.

4 Accounting for the realized measure estimation

The asymptotic theory so far considers the unfeasible test statistic in (6). In this section, we show the asymptotic equivalence of the corresponding feasible test statistic that replaces integrated variances by realized measures. To discuss the impact of estimating the integrated variance, we must first establish some notation that makes explicit the dependence on the number of intraday observations that we employ to compute the realized measure. We thus denote the time series of realized measures by $Y_{t,M}$ and $\boldsymbol{X}_{t,M}^{(d)}$, where M is the number of intraday observations and $d \in \{q, q_A\}$.

Let $\widehat{\boldsymbol{\beta}}_T^{(M)}(y, \boldsymbol{x}^{(d)}) = \left(\widehat{\beta}_{0T}^{(M)}(y, \boldsymbol{x}^{(d)}), \ldots, \widehat{\beta}_{dT}^{(M)}(y, \boldsymbol{x}^{(d)})\right)'$ denote the argument that minimizes

$$\frac{1}{T}\sum_{t=1}^{T}\Big[K_b(Y_{t,M} - y) - \beta_0 - \beta_1(X_{1t,M} - x_1) - \ldots - \beta_d(X_{dt,M} - x_d)\Big]^2 \prod_{j=1}^{d} W_{h_d}(X_{jt,M} - x_j).$$

The local linear estimator of the conditional density is $\widehat{f}_{Y|\boldsymbol{X}^{(d)}}^{(M)}(y \,|\, \boldsymbol{x}^{(d)}) = \widehat{\beta}_{0T}^{(M)}(y, \boldsymbol{x}^{(d)})$, yielding the following feasible test statistic

$$\widehat{\Lambda}_{M,T} = \widehat{\Omega}_{M,T}^{-1}\left\{\begin{array}{r} h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T}\left[\dfrac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}^{(M)}(Y_{t,M} \,|\, \boldsymbol{X}_{t,M}^{(q)}) - \widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}^{(M)}(Y_{t,M} \,|\, \boldsymbol{X}_{t,M}^{(q_A)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}^{(M)}(Y_{t,M} \,|\, \boldsymbol{X}_{t,M}^{(q_A)})}\right]^2 \pi(Y_{t,M}, \boldsymbol{X}_{t,M}^{(q)}) \\ -\, h_q^{-q/2}\, b^{-1/2}\, \widehat{\mu}_{1,T}^{(M)} - h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\, \widehat{\mu}_{2,T}^{(M)} + 2\, h_q^{q/2-q_A}\, b^{-1/2}\, \widehat{\mu}_{3,T}^{(M)} \end{array}\right\}, \tag{8}$$

where $\widehat{\mu}_{\cdot,T}^{(M)}$ differs from $\widehat{\mu}_{\cdot,T}$ because it employs realized measures rather than the true integrated variance. Let $N_{0,t,M} = Y_t - Y_{t,M}$ and $N_{j,t,M} = X_{j,t} - X_{j,t,M}$ for $1 \leq j \leq d \in \{q, q_A\}$ denote the errors stemming from the estimation of the integrated variance. To ensure the asymptotic equivalence between the unfeasible and feasible test statistics, we must restrict the rate at which the moments of the estimation errors converge to zero. We do that by constraining the moments of the drift and diffusive functions as well as of the noise due to market-microstructure effects.

Assumption A6: The drift terms of (1) are continuous locally bounded processes with $\mathbb{E}\,|\mu_{i,t}|^{2k} < \infty$, whereas the diffusive functions are *càdlàg* with $\mathbb{E}\,|\sigma_{ij,t}|^{2k} < \infty$ and the jump components $c_i(t)$ are iid with $\mathbb{E}\,|c_i(t)|^{2k} < \infty$ for some $k \geq 2$ and $i, j \in \{A, B\}$. In addition, the microstructure noise is iid with symmetric distribution around zero and with finite $2k$th moment for some $k \geq 2$.

Assumption A6 ensures that the conditions in Corradi et al.'s (2009) Lemma 1 hold and hence that there exists a sequence a_M, with $a_M \to \infty$ as $M \to \infty$, such that $\mathbb{E}\,|N_{j,t,M}|^k = O\left(a_M^{-k/2}\right)$ for some $k \geq 2$ and $1 \leq j \leq d \in \{q, q_A\}$. Note that this establishes a bound to the kth moment of the absolute estimation error that depends on the realized measure we employ to estimate the integrated variance. In particular, $a_M = M$ for the realized variance (Andersen, Bollerslev, Diebold and Labys, 2001; Barndorff-Nielsen and Shephard, 2002) and tripower variation (Barndorff-Nielsen and Shephard, 2004), whereas $a_M = M^{1/3}$ for the two-scale realized variance (Zhang, Mikland and Aït-Sahalia, 2005) and $a_M = \sqrt{M}$ for the multi-scale realized variance (Zhang, 2006; Aït-Sahalia, Mykland and Zhang, 2009c) and the realized kernel estimator (Barndorff-Nielsen, Hansen, Lunde and Shephard, 2008a).

Theorem 3: Let Assumptions A1 to A4 and A6 hold as well as the bandwidth conditions (i) to $(viii)$ in Lemmata 1 and 2. Also, let $T^{\frac{k+1}{2(k-1)}} (\ln T)^{1/2} a_M^{-1/2} \to 0$ as $T, M \to \infty$ for k as defined in Assumption A6. It follows for $q \leq 3$ that

(i) Under the null hypothesis $\mathbb{H}_0$, $\widehat{\Lambda}_{M,T} \xrightarrow{d} N(0,1)$.

(ii) Under the alternative hypothesis $\mathbb{H}_A$, $\Pr\left(T^{-1} h_q^{-q/2} b^{-1/2} \,|\widehat{\Lambda}_{M,T}| > \varepsilon\right) \longrightarrow 1$ for any $\varepsilon > 0$.

As for the feasible kernel-based statistic, define $\bar{\Lambda}_{M,T}$ analogously to $\bar{\Lambda}_T$ in (7) but replacing $(Y_t, \boldsymbol{X}_t^{(q)})$ with $(Y_{t,M}, \boldsymbol{X}_{t,M}^{(q)})$. The next results documents asymptotic equivalence in the context of kernel density estimation.

Theorem 4: Let Assumptions A2 to A6 hold as well as the bandwidth conditions (i) to $(viii)$ in Lemmata 4 and 5. Also, let $T^{\frac{k+1}{2(k-1)}} (\ln T)^{1/2} a_M^{-1/2} \to 0$ as $T, M \to \infty$ for k as defined in Assumption A6. It follows that

(i) Under the null hypothesis $\mathbb{H}_0$, $\bar{\Lambda}_{M,T} \xrightarrow{d} N(0,1)$.

(ii) Under the alternative hypothesis $\mathbb{H}_A$, $\Pr\left(T^{-1} h_q^{-q/2} b^{-1/2} \,|\bar{\Lambda}_{M,T}| > \varepsilon\right) \longrightarrow 1$ for any $\varepsilon > 0$.

Theorems 3 and 4 establish that the asymptotic equivalence between unfeasible and feasible test statistics necessitates that the number of intraday observations M grows fast enough relative to the number of days T. As usual, there is a tradeoff between using a non-robust realized measure with $a_M = M$ at a frequency for which microstructure noise is negligible and employing a microstructure-robust realized measure with $a_M = \sqrt{M}$ at the highest available frequency.

5 Bootstrap critical values

It is well known that the asymptotic behavior of nonparametric tests does not always entail a reasonable approximation in finite samples and that their results may heavily depend on the bandwidth choice (see, e.g., Fan, 1995; Fan and Linton, 2003). In what follows, we aim to alleviate such concerns by employing resampling techniques. There a number of issues that one must bear in mind, though. First, given the nonparametric nature of the null hypothesis, we cannot rely on standard resampling algorithms based on either parametric or wild bootstrap methods (Härdle and Mammen, 1993; Andrews, 1997; Aït-Sahalia et al., 2009b). Second, integrated variance does not follow a Markov process, ruling out as well bootstrap algorithms for Markov sequences (Rajarshi, 1990; Paparoditis and Politis, 2002; Horowitz, 2003).[6] Third, a standard nonparametric bootstrap would also fail to mimic the limiting distribution of our test statistics for they involve degenerate U-statistics (Bretagnolle, 1983; Arcones and Giné, 1992).

To circumvent the above issues, we resort to a variation of the standard moon bootstrap by Bretagnolle (1983) and Bickel et al. (1997). We sample $\mathcal{T}$ out of T daily realized measures by blocks (rather than individually) so as to cope with time-series dependence. In addition, for each bootstrap sample, we compute the test statistics using a bandwidth vector (h_{*q}, h_{*q_A}, b_*) that shrinks to zero at a rate depending on $\mathcal{T}$ (rather than T). This implies distinct orders of magnitude for the bias terms in the original and bootstrap statistics and thence they do not cancel out. This is in stark contrast with the bias cancelation that happens within the context of parametric and wild bootstrap. It nonetheless remains unnecessary to compute the scaling term corresponding to the asymptotic variance of the test statistic.

The (unscaled) bootstrap counterparts of (6) and (7) then are respectively

$$\widehat{\Lambda}^*_{\mathcal{T}} = \left\{ \begin{array}{l} h_{*q}^{q/2}\, b_*^{1/2} \sum_{t=1}^{\mathcal{T}} \left[\frac{\widehat{f}^*_{Y|\boldsymbol{X}^{(q)}}(Y_t^*|\boldsymbol{X}_t^{*(q)}) - \widehat{f}^*_{Y|\boldsymbol{X}^{(q_A)}}(Y_t^*|\boldsymbol{X}_t^{*(q_A)})}{\widehat{f}^*_{Y|\boldsymbol{X}^{(q_A)}}(Y_t^*|\boldsymbol{X}_t^{*(q_A)})} \right]^2 \pi(Y_t^*, \boldsymbol{X}_t^{*(q)}) \\ \qquad - h_{*q}^{-q/2}\, b_*^{-1/2}\, \widehat{\mu}^*_{1,\mathcal{T}} - h_{*q}^{q/2}\, h_{*q_A}^{-q_A}\, b_*^{-1/2}\, \widehat{\mu}^*_{2,\mathcal{T}} + 2\, h_{*q}^{q/2-q_A}\, b_*^{-1/2}\, \widehat{\mu}^*_{3,\mathcal{T}} \end{array} \right\} \tag{9}$$

and

$$\bar{\Lambda}^*_{\mathcal{T}} = \left\{ \begin{array}{l} h_{*q}^{q/2}\, b_*^{1/2} \sum_{t=1}^{\mathcal{T}} \left[\frac{\bar{f}^*_{Y|\boldsymbol{X}^{(q)}}(Y_t^*|\boldsymbol{X}_t^{*(q)}) - \bar{f}^*_{Y|\boldsymbol{X}^{(q_A)}}(Y_t^*|\boldsymbol{X}_t^{*(q_A)})}{\bar{f}^*_{Y|\boldsymbol{X}^{(q_A)}}(Y_t^*|\boldsymbol{X}_t^{*(q_A)})} \right]^2 \pi(Y_t^*, \boldsymbol{X}_t^{*(q)}) \\ \qquad - h_{*q}^{-q/2}\, b_*^{-1/2}\, \bar{\mu}^*_{1,\mathcal{T}} - h_{*q}^{q/2}\, h_{*q_A}^{-q_A}\, b_*^{-1/2}\, \bar{\mu}^*_{2,\mathcal{T}} + 2\, h_{*q}^{q/2-q_A}\, b_*^{-1/2}\, \bar{\mu}^*_{3,\mathcal{T}} \end{array} \right\}. \tag{10}$$

As before, we provide in Appendix A the expressions for the bias terms in (9) and (10).

We next establish the first-order validity of the moon bootstrap only for the unfeasible test statistic given that the asymptotic equivalence between the feasible and unfeasible bootstrap statistics ensues along the same lines as in Theorem 3 provided that Assumption A6 holds.

Theorem 5: Let Assumptions A1 to A4 hold and let the bandwidth conditions (i) to $(viii)$ in Lemmata

[6] Restricting attention to the class of eigenfunction stochastic volatility models would actually yield integrated variances with a finite ARMA representation (Meddahi, 2003) and hence approximately Markov (even if of higher order).

1 and 2 hold for (h_{*q}, h_{*q_A}, b_*) and $\mathcal{T}$ in lieu of (h_q, h_{q_A}, b) and T, respectively. Letting $T, \mathcal{T}, T/\mathcal{T} \to \infty$ yields, for $q \leq 3$ and for any $\varepsilon > 0$, $\Pr\left(\sup_{v\in\mathbb{R}} \left|\Pr_*(\widehat{\Lambda}^*_{\mathcal{T}} \leq v) - \Pr(\widehat{\Omega}_T \widehat{\Lambda}_T \leq v)\right| > \varepsilon\right) \longrightarrow 0$ under the null, whereas $\Pr\left(\Pr_*\left(\left|\mathcal{T}^{-1} h_{*q}^{-q/2} b_*^{-1/2} \widehat{\Lambda}^*_{\mathcal{T}}\right| > \varepsilon\right)\right) \longrightarrow 1$ under the alternative. In addition, replacing Assumption A1 with A5 and letting bandwidth conditions (i) to $(viii)$ in Lemmata 4 and 5 hold ensure the first-order validity of the moon bootstrap for the kernel-based test statistic in (10) as well.

It is immediate to see that the sample and bootstrap statistics have the same limiting distribution under the null, whereas they diverge at different rates under the alternative. In particular, (9) and (10) diverge at a slower rate $\mathcal{T} h_{*q}^{q/2} b_*^{1/2}$ relative to their sample counterparts. In practice, one must deal with the feasible bootstrap test statistics $\widehat{\Lambda}^*_{M,\mathcal{T}}$ and $\bar{\Lambda}^*_{M,\mathcal{T}}$ that replace $(Y_t^*, \boldsymbol{X}_t^{*(q)})$ with the corresponding realized measures $(Y_{M,t}^*, \boldsymbol{X}_{M,t}^{*(q)})$. Assumption A6 ensures that the statement in Theorem 5 also applies if one substitutes $\widehat{\Lambda}^*_{M,\mathcal{T}}$, $\bar{\Lambda}^*_{M,\mathcal{T}}$ and $\widehat{\Omega}_{M,T} \widehat{\Lambda}_{M,T}$ for $\widehat{\Lambda}^*_{\mathcal{T}}$, $\bar{\Lambda}^*_{\mathcal{T}}$ and $\widehat{\Omega}_T \widehat{\Lambda}_T$, respectively. The bootstrap critical values for $\widehat{\Omega}_{M,T} \widehat{\Lambda}_{M,T}$ are readily available from the empirical distribution of $\widehat{\Lambda}^*_{M,\mathcal{T}}$ using a large number, say B, of bootstrap statistics.

To check whether the moon block-bootstrap indeed entails accurate critical values, we run a limited Monte Carlo study. In particular, we simulate intraday returns from two independent mean-reverting CIR processes (Cox, Ingersoll and Ross, 1985) and then examine how the empirical size of our two-step testing procedure varies according to the bandwidth choice. We employ the CIR process not only because it is a standard model in finance, but also because it implies a simple ARMA(1,1) process for the integrated variance (Meddahi, 2003). For each of the 500 Monte Carlo replications, we simulate intraday data from

$$\mathrm{d}P_{At} = \kappa_A\,(\mu_A - P_{At})\,\mathrm{d}t + \varsigma_A \sqrt{P_{At}}\,\mathrm{d}W_{At}$$
$$\mathrm{d}P_{Bt} = \kappa_B\,(\mu_B - P_{Bt})\,\mathrm{d}t + \varsigma_B \sqrt{P_{Bt}}\,\mathrm{d}W_{Bt},$$

where W_{At} and W_{Bt} are two 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 $(\kappa_A, \mu_A, \varsigma_A) = (0.080, 0.150, 0.011)$ and $(\kappa_B, \mu_B, \varsigma_B) = (0.120, 0.200, 0.013)$. After burning the first 200 observations of the sample, we employ the last $M\,T$ 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 $M = 144$ and $T = 200$ so as to assess how important is the condition in Theorem 3 that calls for M to grow at a faster rate than T.

From the intraday log-returns $r_{At} = \ln P_{At} - \ln P_{At-1}$ and $r_{Bt} = \ln P_{Bt} - \ln P_{Bt-1}$, we retrieve the daily realized variances RV_{Ad} and RV_{Bd} for each day $d = 1, \ldots, T$ and then test for conditional independence of asset A's daily variance with respect to asset B's daily variance by looking at the conditional density of $X = \ln RV_{Ad}$ given $Y = \ln RV_{Ad-1}$ and $Z = \ln RV_{Bd}$. We first standardize the data by their mean and standard deviation and then estimate the conditional densities using local-linear smoothing. We employ standard second-order Gaussian kernels, for which $C_1(K) = C_1(W) = 1/(2\sqrt{\pi})$ and $C_2(K) = C_2(W) = 1/(2\sqrt{2\pi})$. As for the bandwidths, we consider $b = 0.776\,\kappa_b\,T^{-1/5}$, $h_1 = 0.776\,\kappa_h\,T^{-1/5}/\ln\ln T$, and $h_2 = 0.776\,\kappa_h\,T^{-9/50}/\ln\ln T$ with scaling factors $(\kappa_b, \kappa_h) \in \{1, 2.5\} \times \{1, 2.5, 5\}$. For simplicity, we employ a weighting scheme relying on a standard multivariate normal density, i.e., $\pi_{XYZ}(x, y, z) = \phi(x, y, z) = \phi(x)\phi(y)\phi(z)$. Given that the distribution of the realized variance logarithm is typically close to normal (Andersen et al., 2001, 2003), such a weighting function keeps the focus on the bulk of the data rather than on extreme levels of volatility.

To ensure a reasonable number of daily observations in the bootstrap artificial samples, we consider $\mathcal{T} = \lfloor T^{.87} \rfloor$, though further simulations how that the results are quite robust to changes in $\mathcal{T}$. Table 1 reports the results for $B = 199$ bootstrap samples using a block length of $\lfloor T^{1/4} \rfloor$ daily observations and scaling factors set to $\kappa_b \in \{1, 2.5\}$ and $\kappa_h \in \{1, 2.5, 5\}$. Despite the fact that $M < T$, we find that empirical size is close to nominal as long as κ_h is not too high relative to κ_b. All in all, fixing $\kappa_b = \kappa_h$ yields very encouraging results, thereby providing some guidance for the bandwidth choice in practice.

6 Spillovers across international stock markets

We examine whether there are volatility spillovers between China, Japan, and US using data from their main stock market indices. In particular, we collect ultra-high-frequency data for the SSE B share index, the Topix 100 index, and the S&P 500 index from Reuters, available at the Securities Industry Research Centre of Asia-Pacific (www.sirca.org.au).

Before describing the data, it is important to justify our index selection by establishing some background. We adopt the S&P 500 index to measure the movements in the US stock market because it is one of the main bellwethers for the US economy. It is also quite straightforward to trade on the performance of the S&P 500 index by means of a wide array of derivatives (e.g., futures and options on the Chicago Mercantile Exchange, and variance swaps in the over-the-counter market) as well as of exchange-traded funds on the American Stock Exchange. In addition, the CBOE also publishes a volatility index (VIX) that measures market expectations of the near-term volatility implied by the S&P 500 index options. This is convenient because it provides an extra control variable to cope with the time-series persistence in the daily volatility of the S&P 500 index.

As for the Topix 100 index, it is a weighted index gauging the performance of the 100 most liquid stocks with the largest market capitalization on the Tokyo Stock Exchange (TSE). There are two continuous trading sessions on the TSE, with a call auction-procedure determining their opening prices. The morning session runs from 9:00 to 11:00, whereas the afternoon session is from 12:30 to 15:00. In view of the time difference, there is no overlapping trading hours between Tokyo and the US stock markets. The same applies to the Shanghai Stock Exchange (SSE), whose morning and afternoon consecutive bidding sessions run from 9:30 to 11:30 and from 13:00 to 15:00. One of the particular features of the Chinese stock market is the relative importance of individual investors despite the fact they face substantial trading restrictions, e.g., a very stringent short-sale constraint (Hertz, 1998; Feng and Seasholes, 2008). In addition, local investors could not own B shares before March 2001 and, even though they may now purchase them using foreign currency, capital controls still restrict their ability to do so. See Allen, Qian and Qian (2007) and Mei, Scheinkman and Xiong (2009) for more details on the institutional background. Our motivation to include the SSE B share index in the analysis is twofold. First, because pricing of trading for B shares is in US dollars, there is no room for exchange rate movements to blur (or to cause spurious) stock market links. Second, albeit its stock market is relatively young, dating back only to November 1990, China is becoming a major player in the world economy and hence it is interesting to study the role it plays within the context of volatility transmission. The fact that B shares are not as liquid as A shares in the Shanghai Stock Exchange means that controlling for market-microstructure noise is essential for China.

The sample runs from January 3, 2000 to December 30, 2005 with 1,301 common trading days. In fact, the sample size slightly changes according to the stock market of reference in view that China, Japan and US do not share all holidays in common and that we always condition on the most recent pre-determined observation of the other stock market. To compute the realized measures of daily integrated variance, we first compute continuously compounded returns over regular time intervals of 1, 5, 10, 15, and 30 minutes. The sample does not include overnight returns in that the first intraday return refers to the opening price that ensues from the, if any, pre-sessional auction. Similarly, we also exclude returns over the lunch break for China and Japan, though they do not affect at all the results (see Section 6.1).

Table 2 reports the descriptive statistics for the 1-minute and 30-minute returns. The average intraday return is slightly negative for every stock market, though relative lower for Japan and China. This is to some extent surprising in view that the Topix 100 and, especially, the Shanghai B share indices exhibit larger standard deviations than the S&P500 index. As usual, index returns exhibit substantial excess kurtosis, which rapidly increases with the sampling frequency. As for skewness, it is strongly negative for the S&P 500 index at the 1-minute frequency, though slightly positive at the 30-minute frequency. The opposite applies to the Topix 100 index, namely, skewness is marginally positive at the 1-minute frequency, whereas strongly negative at the 30-minute frequency. As for the SSE B share index, skewness increases with the sampling frequency from slightly above zero to almost one.

A possible explanation for the differences between the skewness and kurtosis of the 1-minute and 30-minute returns rests on market microstructure effects due to liquidity issues. The proportion of zero returns are indeed higher for the Topix 100 index and much higher for the SSE B share index than for the S&P 500 index. It is thus not surprising that we observe strong first-order autocorrelation for every stock market index at the 1-minute frequency as well as for the 30-minute SSE B share index returns. Further analysis shows that the liquidity of the SSE B share index, as measured by the proportion of nonzero returns, increases over time, especially after March 2001.

In what follows, we carry out our empirical analysis of the volatility transmission using realized measures of daily variance. We consider the realized variance, the tripower variation, the two-scale realized variance, and the realized kernel based on 1-minute and 5-minute returns. In addition, we also compute the realized variance and tripower variation using 15-minute and 30-minute returns. The realized variance based on 1-minute and 5-minute returns essentially gauges the overall quadratic variation, including not only information about the daily variance but also about price jumps and microstructure noise. As the sampling frequency decreases, reducing the market microstructure effects, the realized variance starts reflecting more the diffusive and jump contributions to the quadratic variation. The tripower variation excludes the contribution of price jumps to the quadratic variation and hence provide a reasonable estimator for the daily variance if based on 15-minute and 30-minute returns. Otherwise, at the 1-minute and 5-minute frequencies, it estimates the quadratic variation of the microstructure noise plus the integrated variance over the day. Finally, the two-scale and realized kernel approaches eliminate the contribution of the microstructure noise to the quadratic variation, capturing only its jump and diffusive components.

Figure 1 plots the time series of the realized variance and tripower variation based on the 1-minute and 30-minute index returns as well as the realized kernel estimate of the daily variance based on 1-minute index returns. It is interesting to observe that controlling for market microstructure noise affects in a substantial manner the estimates of the daily variance. In particular, the realized variance based on the 1-minute returns is very small relative not only to the corresponding realized kernel estimate, but also to the realized variance based on 30-minute returns. This is particularly true for the SSE B share index in the first half of the sample and hence well in line with our preliminary findings concerning illiquidity. In contrast, accounting for jumps seems to matter only for the Topix 100 index, especially if we also control for market microstructure noise. To appreciate that, it suffices to observe that the little discrepancy between the realized variance and the tripower variation based on 30-minute returns on the SSE B share index and on the S&P 500 index.

Tables 3 to 5 report the test results for the null hypothesis of conditional independence using bootstrap critical values. As the null of conditional independence is invariant to monotonic transformations, we first standardize the logarithms of the realized measures by their mean and standard deviation and then estimate the conditional densities using local linear smoothing with Gaussian-type kernels.[7] In accordance with the Monte Carlo results, we focus on bandwidth scaling factors set to $\kappa_b = \kappa_h = 2.5$, though the results remain qualitatively the same for $(\kappa_b, \kappa_h) \in \{(1,1),(2.5,5)\}$. As before, we employ a weighting scheme based on the standard multivariate normal density. To obtain bootstrap critical values, we construct $B = 499$ bootstrap artificial samples of size $\mathcal{T} = \lfloor T^{.65} \rfloor$ by resampling blocks of $\lfloor \mathcal{T}^{1/4} \rfloor$ daily observations.

Table 3 documents that there is some evidence of volatility spillovers at the 5% significance level running from Japan to US regardless of the realized measure we employ at the 1-minute frequency. It is interesting to note that controlling either for jumps through tripower variation or for microstructure noise by means of a realized kernel approach yields stronger evidence. This is in line with the presence of both jumps and market microstructure effects in the Topix 100 index as seen in Figure 1. Apart from a somewhat borderline result for the realized kernel (p-value of 0.124), we cannot reject conditional independence at the 10% significance level using realized measures based on 5-minute returns. In contrast, tests using realized variance based on 15-minute and 30-minute returns detect strong spillovers in the quadratic variation process running from the Topix 100 index to the S&P 500 index. This indicates that market microstructure noise to some extent blurs the evidence of spillovers in the quadratic variation. In addition, the strength of the statistical evidence reduces once we employ tripower variation, with p-values increasing to about 10%. We thus conjecture that the primary channel of transmission from Japan to US is due mostly to jumps even if there is some weak evidence of volatility spillovers.

The results for the SSE B share index are similar concerning market microstructure effects in that they seem to mask in a substantial manner the evidence of spillovers. In particular, we reject the null of conditional independence only for the more microstructure-robust realized measures, namely, two-scale realized variance and realized kernel based on 1-minute and 5-minute returns as well as realized variance

[7] More specifically, we employ the standard Gaussian kernel for $q = 2$ and the fourth-order kernel derived from the Gaussian density for $q = 3$ (Schucany and Sommers, 1977).

using 15-minute and 30-minute returns. To identify whether the transmission is through the diffusive part or through the discontinuous part of the quadratic variation, we turn our attention to the tests using tripower variation. We view their failure to reject the null as a clue that the transmission is essentially through jumps, without much action taking place in the volatility.

Table 4 then asks whether the VIX index conditionally depends on the quadratic variation processes of the Topix 100 index and of the SSE B share index given its own previous realization. The motivation is to investigate whether the diffusive and jump components of the quadratic variation in Japan and/or China affect the risk-neutral expectation of the S&P 500 index variance as measured by the options-implied variance. A positive answer would mean that investors price these spillovers in the VIX index. As per the Japan effect, we reject the null for every realized measure based on 1-minute returns as well as for the realized variance and tripower variation at the 5-minute frequency. We also observe a borderline rejection at the 10% level of significance for the realized variance based on 15-minute returns (p-value of 0.100). Incorporating the realized measure of the S&P 500 index over the previous day to the conditioning set debilitates in general the evidence of transmission from the Topix 100 index to the VIX index. We now reject the null of conditional independence only for the realized variance and tripower variation at the 1-minute frequency and for the realized variance using 15-minute returns. There is also a borderline result for the realized variance at the 30-minute frequency (p-value of 0.114). Altogether, the statistical evidence of spillovers running from the Topix 100 index to the the VIX index is well in line, though weaker, with the one for the quadratic variation of the S&P 500 index.

The quadratic variation of the SSE B share index appears to have a much more significative impact in the VIX index than in the quadratic variation of the S&P 500 index. However, adding the quadratic variation of the S&P 500 index as an extra control weakens the evidence against conditional independence to a large extent, even if we still reject the null for the realized variance at the 15-minute frequency, for the realized variance and tripower variation based on 30-minute returns, and for the realized kernel at the 5-minute frequency. As before, market microstructure noise appears to eclipse the dependence of the VIX index on both the diffusive and jump components of the quadratic variation of the SSE B share index.

Panel A in Table 5 evinces that the dependence structure between the S&P 500 index and the Topix 100 index is sort of symmetric in that the transmission seems to eventuate in both directions mainly through jumps. We indeed reject the null of conditional independence using the tripower variation only at the 1-minute frequency, whereas we always reject if we employ any realized measure that does not exclude the contribution of jumps to the quadratic variation. Also, we find some evidence of volatility spillover running from China to Japan for the realized kernel estimates, for the realized variance at the 1-minute frequency, and for the tripower variation based on 1-minute and 30-minute returns. The presence of jumps and market microstructure effects seems to somehow conceal the volatility transmission from the SSE B share index to the Topix 100 index.

Panel B in Table 5 reveals a somewhat different pattern for (volatility) spillovers running from either Japan or US to China. At the usual levels of significance, we fail to reject the null of conditional independence with respect to the S&P 500 index for every realized measure at any sampling frequency. As for the Topix 100 index, we find evidence of spillovers at the 5% significance level only if we do not control for market microstructure noise, whereas at the 10% level we also reject the null for the tripower variation using 15-minute returns. These results suggest that the main channel of transmission from Japan is through the diffusive component of the quadratic variation. All in all, it surprisingly seems that the volatility transmission runs mostly from China to either Japan or US than vice-versa.

Finally, we also compare in Figures 2 and 3 the conditional density of the daily quadratic variation of the Topix 100 and S&P 500 indices given past realizations at their median values with the conditional density both given their past realization and given the quadratic variation of the SSE B share index at different quartiles. Figure 2 illustrates that the China effect in Japan is apparent in every quartile, though it slightly increases with the value of the Shanghai quadratic variation. This is in contrast with the way it affects the quadratic variation of the S&P 500 index in Figure 3. In particular, the China effect seems more strong as we move towards both tails of the distribution. This is interesting for we would most likely fail to capture this sort of nonlinear effect within the extant parametric analyses in the literature. It is only because we look at the whole distribution, rather than restricting attention to its conditional mean, that we are able to detect these spillovers coming from China.

6.1 Robustness analysis

We revisit our empirical findings so as to evaluate their sensitiveness to the testing setup. To begin with, we check whether there is any qualitative change in the results if we use different realized measures. We initially observe that including or excluding the return on the Topix 100 index and on the SSE B share index over their lunch breaks makes almost no difference in the quantitative results. In addition, we also perform our tests using the realized semivariance (Barndorff-Nielsen, Kinnebrock and Shephard, 2008b), the median realized variance (Andersen, Dobrev and Schaumburg, 2009), and the multi-scale realized variance (Zhang, 2006; Aït-Sahalia et al., 2009c). The first is a better measure of risk than realized variance for asymmetric distributions, whereas the second and third are alternative volatility estimators robust to jumps and to market microstructure noise, respectively. Apart from experiencing some numerical problems with the multi-scale approach, we find that the results are generally in line with what we report in Tables 3 to 5. We omit them not only to conserve on space, but also because there are no characterization results yet in the literature for the estimation errors of the semi-variance and of the median realized variance estimators.

In what follows, we complement our robustness analysis by conducting two additional inspections. First, we assess how pivotal is the assumption that, under the null hypothesis, the past integrated variance suffices to control for the persistence in the data by also conditioning on the past implied volatility. Second, we examine how fast the volatility transmission occurs by looking at the integrated variance over shorter periods of time. In particular, instead of confining attention to daily volatility, we focus on the conditional distribution of the integrated variance over the first hour of the trading day given the last hour of previous trading day and the last hour of trading on the other stock market.

6.1.1 Data persistence

The conditional independence restriction we test does not exactly correspond to a null of noncausality in variance given the non-Markovian character of the daily integrated variance. In particular, the empirical analysis in Section 6 controls only for the own integrated variance in the previous day rather than on the whole history of daily variances. This raises a concern on whether our findings are in fact genuine or just an artifact due to higher-order dependence in the data. To appreciate the latter, suppose for instance that one must condition on the integrated variance of the last two days to approximate well the transition density of the integrated variance of asset B. Even if causality in variance runs only from asset B to asset A, we could well reject the null of conditional independence because the past integrated variance of asset A at day t would proxy for the integrated variance of asset B at day $t-2$.

To assess robustness against such concerns, we redo our empirical analysis of volatility transmission to US including the past VIX index as an additional control. The latter measures the options-implied volatility of the S&P500 index and hence should provide information about the future integrated variance.[8] In addition, Bandi and Perron (2006) find strong evidence of fractional cointegration between implied and realized variances and so conditioning on the VIX index should effectively control for any high-order dependence implied by the non-Markovian nature of the integrated variance, regardless of whether the null of conditional independence is true or not.

[8] Note that it is not necessary to assume that the VIX index is the best forecast for future realizations of the integrated variance nor that it is unbiased or efficient (see, among others, Christensen and Prabhala, 1998).

Table 3 shows that adding the VIX index to the conditioning set does not alter much the qualitative results concerning the Topix 100 index. The p-values for the tests based on realized variance and tripower variation using 15-minute and 30-minute returns are indeed comparable in magnitude. The main difference is that we are now unable to reject the null using the microstructure-robust realized measures rooted in 1-minute and 5-minute returns. In contrast, accounting for the VIX index somewhat affects the results for the SSE B share index. Although we still evince quadratic variation spillovers using the realized variance based on 15-minute returns, we no longer reject the null using the realized variance at the 30-minute frequency. In addition, we now find borderline results for the tripower variation based on 15-minute and 30-minute returns as well as for the two-scale estimator using 5-minutes returns.

6.1.2 Reaction time

King and Wadhwani (1990) derive an imperfectly revealing equilibrium model to explain contemporaneous transmission of volatility between stock markets. Their framework posits that price jumps will take place as soon as a market reopens so as to reflect changes in both idiosyncratic and common factors since last trade. Given that other stock market indices also depend on the common factors, contagion will result in immediate spillovers from one market to another as the latter reopens for trading. This is in stark contrast with our empirical study focusing on daily quadratic variation (rather than over a shorter interval of time) in that we could well miss the almost instantaneous reaction that King and Wadhwani (1990) predict.

We thus investigate in this section whether reaction time is indeed an issue. For the transmission from either China or Japan to US, we test whether the realized measures based on 1-minute returns over the first hour of trading in US depends on the corresponding realized measures over the last hour of trading in either China or Japan in that same day given the realized measures over the last hour of trading in US in the previous day. To examine spillovers running in the opposite direction, we then check whether the realized measures over the last hour of trading in US in the previous day affects the realized measures over the first hour of trading in Asia even after controlling for the latter's realized measures over the last hour of trading in the previous day. The same applies if testing for transmission from China to Japan given that the former is shut when the latter opens in the morning. Finally, given the one-hour difference between Shanghai and Tokyo, we test for spillovers from Japan to China by looking at whether the realized measures of the SSE B share index over the first hour of trading depends on the realized measures of the Topix 100 index over the one-hour period immediately before the opening of the Shanghai Stock Exchange (i.e., from 9:30 to 10:30 Tokyo time) given the realized measures of the SSE B share index over the last hour of trading in the previous day.

As before, asymptotic equivalence between the feasible and unfeasible statistics allows us to interpret realized variance results as concerning the total quadratic variation of the process, including the contributions of the integrated variance, of jumps in the asset prices, and of the market microstructure noise. Tripower variation purges the influence of price jumps and hence estimates the quadratic variation of the market microstructure noise plus the integrated variance of the process, whereas the two-scale and realized kernel estimators measure the contributions of the jump and diffusive components to the quadratic variation.

Panel A in Table 6 documents that there are spillovers from Japan to the S&P 500 index regardless of the realized measure we employ. In contrast, we find no such evidence for the SSE B share index at the usual levels of significance. At first glance, this is a bit surprising given that the primary channel of transmission over the day is through price jumps (see above discussion). There is nevertheless no contradiction in view that failure to reject the null at the hourly frequency does not exclude a response of the daily quadratic variation of the S&P 500 index to price jumps in China before its last hour of trading. Panel B evinces mainly volatility spillovers from either China or US to Japan given that we reject the null only if we control for either jumps by means of the tripower variation or market microstructure noise using the two-scale realized variance. This is somewhat consistent with volatility spillovers through jumps in the volatility given that the reaction time story is mainly about jumps either in prices or volatility. As for the transmission running from China to Japan, apart from jumps in prices, it is also important to control for the market microstructure imprint. We indeed reject the null at the 5% level for the test based on two-scale realized variance and at the 10% level also for tests using tripower variation, whereas we find only a borderline result for the realized variance (p-value of 0.114). This again suggests that the main channel of transmission is through jumps in the volatility process. Finally, as for spillovers to the SSE B share index, Panel C reveals that the S&P 500 index exerts no impact, whereas there is some weak evidence of a Japan effect at the 10% significance level for the two-scale realized variance.

7 Conclusion

This paper develops formal statistical tools for nonparametric tests of conditional independence between integrated variances. Under the assumption that asset prices follow a multivariate jump-diffusion processes with stochastic volatility, we show how to test whether the conditional distribution of asset A's integrated variance also depends on information concerning asset B's integrated variance. Our testing procedure involves two steps. In the first stage, we estimate the integrated variances using intraday returns data by means of realized measures so as to avoid misspecification risks. In the second step, we then test for conditional independence between the resulting realized measures. Although asymptotic critical values are not very reliable in finite samples, we show how to construct more accurate critical values by means of a simple bootstrap algorithm.

Our contribution to the literature on nonparametric density-based tests 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 also consider a more general setup in which the transition distribution may depend on a state vector of any dimension. It turns out that such a generalization is not so straightforward as it seems at first glance. In particular, one must employ kernel-based methods rather than local linear smoothing if the dimension of the conditioning set is large enough.

We also contribute to the literature on international market links by investigating the quadratic variation transmission between China, Japan, and US. Our empirical findings evince that the quadratic variation of the Topix 100 index has a bidirectional link with the quadratic variation of the stock markets in US and China. The primary channels of transmission are through price jumps in the case of US and through volatility spillovers as to what concerns China. Further, we document that price jumps in the latter also affect both the quadratic variation and the options-implied volatility of the S&P 500 index, though we find no evidence of transmission in the opposite direction. Finally, by focusing on realized measures over one-hour intervals (rather than over a day), we are also able to uncover some interesting transmission patterns that are likely to result from dependence between jumps in the volatility.

Appendix

A Bias and scaling terms

Let $C_1(K) \equiv \int K(u)^2 \, \mathrm{d}u$ and $C_2(K) \equiv \int \left(\int K(u)K(u+v) \, \mathrm{d}u \right)^2 \, \mathrm{d}v$. Define $C_1(\boldsymbol{W})$, $C_2(\boldsymbol{W})$, $C_1(\widetilde{\boldsymbol{W}})$, and $C_2(\widetilde{\boldsymbol{W}})$ analogously. The bias and scaling terms that appear in (6) are given by

$$\widehat{\mu}_{1,T} = C_1(K)\, C_1(\boldsymbol{W})\, \frac{1}{T} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t^{(q)})}{\bar{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)})} - b\, C_1(\boldsymbol{W})\, \frac{1}{T} \sum_{t=1}^{T} \frac{\frac{1}{T}\sum_{s=1}^{T} \boldsymbol{W}_{h_q}(\boldsymbol{X}_s^{(q)} - \boldsymbol{X}_t^{(q)})\, \pi(Y_s, \boldsymbol{X}_s^{(q)})}{\bar{f}_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)}) \frac{1}{T}\sum_{s=1}^{T} \boldsymbol{W}_{h_q}(\boldsymbol{X}_s^{(q)} - \boldsymbol{X}_t^{(q)})}$$

$$\widehat{\mu}_{2,T} = C_1(K)\, C_1(\widetilde{\boldsymbol{W}})\, \frac{1}{T} \sum_{t=1}^{T} \frac{\frac{1}{T}\sum_{s=1}^{T} K_b(Y_s - Y_t) \widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)} - \boldsymbol{X}_t^{(q_A)})\, \pi(Y_s, \boldsymbol{X}_s^{(q)})}{\bar{f}_{Y,\boldsymbol{X}^{(q_A)}}(Y, \boldsymbol{X}_t^{(q_A)}) \frac{1}{T}\sum_{s=1}^{T} K_b(Y_s - Y_t) \widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)} - \boldsymbol{X}_t^{(q_A)})}$$
$$- b\, C_1(\widetilde{\boldsymbol{W}})\, \frac{1}{T} \sum_{t=1}^{T} \frac{\frac{1}{T}\sum_{s=1}^{T} \widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)} - \boldsymbol{X}_t^{(q_A)})\, \pi(Y_s, \boldsymbol{X}_s^{(q)})}{\bar{f}_{\boldsymbol{X}^{(q_A)}}(\boldsymbol{X}_t^{(q_A)}) \frac{1}{T}\sum_{s=1}^{T} \widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)} - \boldsymbol{X}_t^{(q_A)})}$$

$$\widehat{\mu}_{3,T} = C_1(K)\, \widetilde{\boldsymbol{W}}(0)\, \frac{1}{T} \sum_{t=1}^{T} \frac{\frac{1}{T}\sum_{s=1}^{T} K_b(Y_s - Y_t) \widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)} - \boldsymbol{X}_t^{(q_A)})\, \pi(Y_s, \boldsymbol{X}_s^{(q)})}{\bar{f}_{Y,\boldsymbol{X}^{(q_A)}}(Y_t, \boldsymbol{X}_t^{(q_A)}) \frac{1}{T}\sum_{s=1}^{T} K_b(Y_s - Y_t) \widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)} - \boldsymbol{X}_t^{(q_A)})}$$

$$- b\,\widetilde{\boldsymbol{W}}(0)\,\frac{1}{T}\sum_{t=1}^{T}\frac{\frac{1}{T}\sum_{s=1}^{T}\widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)}-\boldsymbol{X}_t^{(q_A)})\,\pi(Y_s,\boldsymbol{X}_s^{(q)})}{\bar{f}_{\boldsymbol{X}^{(q_A)}}(\boldsymbol{X}_t^{(q_A)})\frac{1}{T}\sum_{s=1}^{T}\widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_s^{(q_A)}-\boldsymbol{X}_t^{(q_A)})}$$

$$\widehat{\Omega}_T^2 = 2\,C_2(K)\,C_2(\boldsymbol{W})\,\frac{1}{T}\sum_{t=1}^{T}\frac{\pi^2(Y_t,\boldsymbol{X}_t^{(q)})}{\bar{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})}$$

with $\boldsymbol{W}_{h_q}(\boldsymbol{u}) = h_q^{-q}\prod_{i=1}^{q} W(u_i/h_q)$ and $\widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{u}) = h_{q_A}^{-q_A}\prod_{i=1}^{q_A} W(u_i/h_{q_A})$.

To estimate the asymptotic bias and variance of the integrated squared relative difference statistic based on kernel smoothing, we employ similar bias and scaling terms in (7). The only difference is that we replace the second-order univariate kernel W with the s-order kernel function $\bar{W}$. For instance,

$$\bar{\mu}_{1,T} = C_1(K)\,C_1(\bar{\boldsymbol{W}})\,\frac{1}{T}\sum_{t=1}^{T}\frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{\bar{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})} - b\,C_1(\bar{\boldsymbol{W}})\,\frac{1}{T}\sum_{t=1}^{T}\frac{\frac{1}{T}\sum_{s=1}^{T}\bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_s^{(q)}-\boldsymbol{X}_t^{(q)})\,\pi(Y_s,\boldsymbol{X}_s^{(q)})}{\bar{f}_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)})\frac{1}{T}\sum_{s=1}^{T}\bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_s^{(q)}-\boldsymbol{X}_t^{(q)})}.$$

Finally, for the bootstrap test statistics in (9) and (10), we obtain $(\widehat{\mu}_{1,\mathcal{T}}^*, \widehat{\mu}_{2,\mathcal{T}}^*, \widehat{\mu}_{3,\mathcal{T}}^*)$ and $(\bar{\mu}_{1,\mathcal{T}}^*, \bar{\mu}_{2,\mathcal{T}}^*, \bar{\mu}_{3,\mathcal{T}}^*)$ by replacing the sample quantities in $(\widehat{\mu}_{1,T}, \widehat{\mu}_{2,T}, \widehat{\mu}_{3,T})$ and in $(\bar{\mu}_{1,T}, \bar{\mu}_{2,T}, \bar{\mu}_{3,T})$ with their bootstrap counterparts, that is to say, we substitute $(Y_t^*, X_t^{*(q)}, \mathcal{T}, b_*, h_{*q}, h_{*q_A})$ for $(Y_t, X_t^{(q)}, T, b, h_q, h_{q_A})$. For instance,

$$\widehat{\mu}_{1,\mathcal{T}}^* = C_1(K)\,C_1(\boldsymbol{W})\,\frac{1}{\mathcal{T}}\sum_{t=1}^{\mathcal{T}}\frac{\pi(Y_t^*,\boldsymbol{X}_t^{*(q)})}{\bar{f}_{Y,\boldsymbol{X}^{(q)}}^*(Y_t^*,\boldsymbol{X}_t^{*(q)})} - b_*\,C_1(\boldsymbol{W})\,\frac{1}{\mathcal{T}}\sum_{t=1}^{\mathcal{T}}\frac{\frac{1}{\mathcal{T}}\sum_{s=1}^{\mathcal{T}}\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_s^{*(q)}-\boldsymbol{X}_t^{*(q)})\pi(Y_s^*,\boldsymbol{X}_s^{*(q)})}{\bar{f}_{\boldsymbol{X}^{(q)}}^*(\boldsymbol{X}_t^{*(q)})\frac{1}{\mathcal{T}}\sum_{s=1}^{\mathcal{T}}\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_s^{*(q)}-\boldsymbol{X}_t^{*(q)})}.$$

B Proofs

B.1 Lemmata

The proof of Theorem 1 relies heavily on Lemmata 1 to 3, whereas we employ the results in Lemmata 4 to 6 in the proof of Theorem 2.

Lemma 1: Assume that there are at most three conditioning variables in the higher dimensional density ($q \leq 3$) and that the bandwidths satisfy the following conditions: (i) $Th_q^{q+4} \to 0$, (ii) $Th_q^{3q/2}\,b^{3/2} \to \infty$, (iii) $Th_q^{4+q/2}\,b^{1/2} \to 0$, (iv) $Th_q^{q/2}\,b^{2s+1/2} \to 0$, (v) $h_q^{4-q}\,b^{-1} \to 0$, (vi) $h_q^{-q}\,b^{2s-1} \to 0$, (vii) $Th_{q_A}^{q_A}\,b\left(h_{q_A}^4 + b^{2s}\right) \to 0$, and $(viii)$ $T(\ln T)^{-1}h_q^q\,h_{q_A}^{q_A}\,b^2 \to \infty$. It then follows from Assumptions A1 to A4 that, under the null $\mathbb{H}_0$,

$$\Omega^{-1}\left\{h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\pi(Y_t,\boldsymbol{X}_t^{(q)})\left[\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - f_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}\right]^2 - h_q^{-q/2}\,b^{-1/2}\,\mu_1\right\} \xrightarrow{d} N(0,1)$$

where $\Omega^2 \equiv 2\,C_2(K)\,C_2(\boldsymbol{W})\int\pi^2(y,\boldsymbol{x})\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x}$ and

$$\mu_1 = C_1(K)\,C_1(\boldsymbol{W})\int\pi(y,\boldsymbol{x}^{(q)})\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x}^{(q)} - b\,C_1(\boldsymbol{W})\int\mathbb{E}\Big[\pi(Y,\boldsymbol{X}^{(q)})\big|\boldsymbol{X}^{(q)} = \boldsymbol{x}^{(q)}\Big]\,\mathrm{d}\boldsymbol{x}^{(q)}.$$

Lemma 2: Assume that there are at most three conditioning variables in the higher dimensional density ($q \leq 3$) and that the bandwidths are such that: (i) $Th_d^{d+4} \to 0$, (ii) $Th_d^{3d/2}\,b^{3/2} \to \infty$, (iii) $Th_d^{4+d/2}\,b^{1/2} \to 0$, (iv) $Th_d^{d/2}\,b^{2s+1/2} \to 0$, (v) $h_d^{4-d}\,b^{-1} \to 0$, and (vi) $h_d^{-d}\,b^{2s-1} \to 0$ for $d \in \{q, q_A\}$ as well as (vii) $Th_{q_A}^{q_A}\,b\left(h_{q_A}^4 + b^{2s}\right) \to 0$ and $(viii)$ $T(\ln T)^{-1}h_q^{-q}\,h_{q_A}^{3\,q_A}\,b^2 \to \infty$. Assumptions A1 to A4 then ensures that

$$\Omega^{-1}\left\{h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\pi(Y_t,\boldsymbol{X}_t^{(q)})\left[\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) - f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}\right]^2 - h_q^{q/2}\,h_{q_A}^{-q_A}\,b^{-1/2}\,\mu_2\right\} = o_p(1)$$

where

$$\mu_2 = C_1(K)\,C_1(\widetilde{\boldsymbol{W}})\int\mathbb{E}\Big[\pi(Y,\boldsymbol{X}^{(q)})\big|Y = y, \boldsymbol{X}^{(q_A)} = \boldsymbol{x}^{(q_A)}\Big]\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x}^{(q_A)}$$
$$- b\,C_1(\widetilde{\boldsymbol{W}})\int\mathbb{E}\Big[\pi(Y,\boldsymbol{X}^{(q)})\big|\boldsymbol{X}^{(q_A)} = \boldsymbol{x}^{(q_A)}\Big]\,\mathrm{d}\boldsymbol{x}^{(q_A)}.$$

Lemma 3: Let the bandwidth conditions (i) to $(viii)$ in Lemma 2 hold. Assumptions A1 to A4 ensure that, under the null $\mathbb{H}_0$,

$$\Omega^{-1}\left[h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\pi(Y_t,\boldsymbol{X}_t^{(q)})\,\frac{\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})\,\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} - h_q^{q/2-q_A}\,b^{-1/2}\,\mu_3\right] = o_p(1),$$

where $\widehat{\epsilon}_{Y|\boldsymbol{X}^{(\cdot)}}(Y_t|\boldsymbol{X}_t^{(\cdot)}) \equiv \widehat{f}_{Y|\boldsymbol{X}^{(\cdot)}}(Y_t|\boldsymbol{X}_t^{(\cdot)}) - f_{Y|\boldsymbol{X}^{(\cdot)}}(Y_t|\boldsymbol{X}_t^{(\cdot)})$ and

$$\mu_3 = C_1(K)\, \widetilde{\boldsymbol{W}}(0) \int \mathbb{E}\Big[\pi(Y, \boldsymbol{X}^{(q)})\big|Y = y, \boldsymbol{X}^{(q_A)} = \boldsymbol{x}^{(q_A)}\Big]\, \mathrm{d}y\, \mathrm{d}\boldsymbol{x}^{(q_A)} - b\, \widetilde{\boldsymbol{W}}(0) \int \mathbb{E}\Big[\pi(Y, \boldsymbol{X}^{(q)})\big|\boldsymbol{X}^{(q_A)} = \boldsymbol{x}^{(q_A)}\Big]\, \mathrm{d}\boldsymbol{x}^{(q_A)}.$$

Lemma 4: Let Assumptions A2 to A5 hold as well as the following bandwidths conditions: (i) $Th_q^{q+2s} \to 0$, (ii) $Th_q^{3q/2}\, b^{3/2} \to \infty$, (iii) $Th_q^{2s+q/2}\, b^{1/2} \to 0$, (iv) $Th_q^{q/2}\, b^{2s+1/2} \to 0$, (v) $h_q^{2s-q}\, b^{-1} \to 0$, (vi) $h_q^{-q}\, b^{2s-1} \to 0$, (vii) $T\, h_{q_A}^{q_A}\, b\, \big(h_{q_A}^{2s} + b^{2s}\big) \to 0$ and $(viii)$ $T(\ln T)^{-1} h_q^{q}\, h_{q_A}^{q_A}\, b^2 \to \infty$. It then follows that, under the null $\mathbb{H}_0$,

$$\bar{\Omega}^{-1} \left\{ h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \left[\frac{\bar{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - f_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})}{\bar{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} \right]^2 - h_q^{-q/2}\, b^{-1/2}\, \mu_1 \right\} \xrightarrow{d} N(0,1),$$

where $\bar{\Omega}^2 \equiv 2\, C_2(K)\, C_2(\bar{\boldsymbol{W}}) \int \pi^2(y, \boldsymbol{x})\, \mathrm{d}y\, \mathrm{d}\boldsymbol{x}$.

Lemma 5: Let Assumptions A2 to A5 hold as well as the following bandwidths conditions: (i) $Th_d^{d+2s} \to 0$, (ii) $Th_d^{3d/2}\, b^{3/2} \to \infty$, (iii) $Th_d^{2s+d/2}\, b^{1/2} \to 0$, (iv) $Th_d^{d/2}\, b^{2s+1/2} \to 0$, (v) $h_d^{2s-d}\, b^{-1} \to 0$ and (vi) $h_d^{-d}\, b^{2s-1} \to 0$ for $d \in \{q, q_A\}$ as well as (vii) $Th_{q_A}^{q_A}\, b\, \big(h_{q_A}^{2s} + b^{2s}\big) \to 0$, and $(viii)$ $T(\ln T)^{-1} h_q^{-q}\, h_{q_A}^{3q_A}\, b^2 \to \infty$. It then follows that

$$\bar{\Omega}^{-1} \left\{ h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \left[\frac{\bar{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) - f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{\bar{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}) \right]^2 - h_q^{q/2}\, h_{q_A}^{-q_A} b^{-1/2} \mu_2 \right\} = o_p(1).$$

Lemma 6: Let the bandwidth conditions (i) to $(viii)$ in Lemma 5 hold. It then follows from Assumptions A2 to A5 that, under the null $\mathbb{H}_0$,

$$\bar{\Omega}^{-1} \left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)})\, \frac{\bar{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})\, \bar{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{\bar{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} - h_q^{q/2 - q_A}\, b^{-1/2} \mu_3 \right] = o_p(1),$$

where $\bar{\epsilon}_{Y|\boldsymbol{X}^{(\cdot)}}(Y_t|\boldsymbol{X}_t^{(\cdot)}) \equiv \bar{f}_{Y|\boldsymbol{X}^{(\cdot)}}(Y_t|\boldsymbol{X}_t^{(\cdot)}) - f_{Y|\boldsymbol{X}^{(\cdot)}}(Y_t|\boldsymbol{X}_t^{(\cdot)})$.

B.2 Proof of Theorem 1

(i) We first observe that

$$\widehat{\Omega}_T^2 - \Omega^2 = 2\, C_2(K)\, C_2(\boldsymbol{W}) \left[\frac{1}{T} \sum_{t=1}^{T} \frac{\pi^2(Y_t, \boldsymbol{X}_t^{(q)})}{f_{Y,\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)})} - \int \pi^2(y, \boldsymbol{x}^{(q)})\, \mathrm{d}y\, \mathrm{d}\boldsymbol{x}^{(q)} + \frac{1}{T} \sum_{t=1}^{T} \frac{\pi^2(Y_t, \boldsymbol{X}_t^{(q)})}{f_{Y,\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)}) \widehat{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)})} \Big(\widehat{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)}) - f_{Y,\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)}) \Big) \right].$$

is of order $o_p(1)$ and hence we treat them as asymptotically equivalent in what follows. Under the null that $f_{Y|X^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) = f_{Y|X^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})$ almost surely, it follows that, up to a term of order $o_p(1)$,

$$\begin{aligned}
\Lambda_T = {} & \Omega^{-1} \left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t^{(q)})}{f^2_{Y|\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)})}\, \widehat{\epsilon}^{\,2}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - h_q^{-q/2}\, b^{-1/2}\, \mu_1 \right] \\
& + \Omega^{-1} \left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t^{(q)})}{f^2_{Y|\boldsymbol{X}^{(q_A)}}(Y_t, \boldsymbol{X}_t^{(q_A)})}\, \widehat{\epsilon}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) - h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\, \mu_2 \right] \\
& - 2\, \Omega^{-1} \left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t^{(q)})}{f^2_{Y|\boldsymbol{X}^{(q)}}(Y_t, \boldsymbol{X}_t^{(q)})}\, \widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})\, \widehat{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) - h_q^{q/2 - q_A}\, b^{-1/2}\, \mu_3 \right] \\
& + \Omega^{-1}\, h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \Big[\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - \widehat{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) \Big]^2
\end{aligned}$$

$$\times \left[\frac{1}{\widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} - \frac{1}{f^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} \right]$$

$$- \Omega^{-1}\left[h_q^{-q/2}\, b^{-1/2}\,(\widehat{\mu}_{1,T}-\mu_1) - h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\,(\widehat{\mu}_{2,T}-\mu_2) + 2\, h_q^{q/2-q_A}\, b^{-1/2}\,(\widehat{\mu}_{3,T}-\mu_3) \right]$$

$$= \Lambda^{(0)}_{1,T} + \Lambda^{(0)}_{2,T} + \Lambda^{(0)}_{3,T}, \tag{11}$$

where $\Lambda^{(0)}_{1,T}$ is the sum of the first three terms on the right-hand side of (11). Lemmata 1 to 3 yield the asymptotic normality of $\Lambda^{(0)}_{1,T}$ under the null and ensure that $\Lambda^{(0)}_{2,T} = o_p(1)$ due to bandwidth conditions (vii) and $(viii)$.

It thus remains to show that $\Lambda^{(0)}_{3,T}$ is also of order $o_p(1)$. We start with

$$\begin{aligned} h_q^{-q/2}\, b^{-1/2}\,(\widehat{\mu}_{1,T}-\mu_1) &= C_1(K)\, C_1(\boldsymbol{W})\, h_q^{-q/2}\, b^{-1/2}\, \frac{1}{T}\sum_{t=1}^{T} \frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})\left(\widehat{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)}) - f_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})\right)}{\widehat{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)}) f_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})} \\ &\quad + h_q^{-q/2}\, b^{1/2}\, C_1(\boldsymbol{W})\, \frac{1}{T}\sum_{t=1}^{T} \frac{1}{f_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)})} \Bigg\{ \mathbb{E}\Big[\pi(Y_s,\boldsymbol{X}_s^{(q)})|\boldsymbol{X}_s^{(q)} = \boldsymbol{X}_t^{(q)}\Big] \\ &\qquad\qquad - \frac{\frac{1}{T}\sum_{s=1}^{T} \boldsymbol{W}_{h_q}(\boldsymbol{X}_s^{(q)} - \boldsymbol{X}_t^{(q)})\,\pi(Y_s,\boldsymbol{X}_s^{(q)})}{\frac{1}{T}\sum_{s=1}^{T} \boldsymbol{W}_{h_q}(\boldsymbol{X}_s^{(q)} - \boldsymbol{X}_t^{(q)})} \Bigg\} \\ &\quad + h_q^{-q/2}\, b^{1/2}\, C_1(\boldsymbol{W})\, \frac{1}{T}\sum_{t=1}^{T} \frac{\mathbb{E}\Big[\pi(Y_s,\boldsymbol{X}_s^{(q)})|\boldsymbol{X}_s^{(q)} = \boldsymbol{X}_t^{(q)}\Big]\left(\widehat{f}_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)}) - f_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)})\right)}{f_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)})\widehat{f}_{\boldsymbol{X}^{(q)}}(\boldsymbol{X}_t^{(q)})} \\ &= o_p(1). \end{aligned}$$

The last equality follows from the fact that the second and third terms are of smaller order than the first term, whereas $\inf_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})} f_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})$ is bounded away from zero in a compact set $\mathcal{C}(Y,\boldsymbol{X}^{(q)}) \subset \mathbb{R}^{q+1}$ and the degenerate U-statistic

$$\frac{1}{T}\sum_{t=1}^{T} \frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})\left(\widehat{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)}) - f_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})\right)}{\widehat{f}_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)}) f_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})} = o_p(h_q^{q/2}\, b^{1/2}).$$

In addition, it follows along the same lines that $h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\,(\widehat{\mu}_{2,T}-\mu_2)$ and $h_q^{q/2-q_A}\, b^{-1/2}\,(\widehat{\mu}_{3,T}-\mu_3)$ are also of order $o_p(1)$.

(ii) Consider the following expansion under the alternative hypothesis $\mathbb{H}_A$

$$\begin{aligned} \Lambda_T &= \Omega^{-1}\left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{\widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})}\, \widehat{\epsilon}^{\,2}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - h_q^{-q/2}\, b^{-1/2}\, \widehat{\mu}_{1,T} \right] \\ &\quad + \Omega^{-1}\left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{\widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})}\, \widehat{\epsilon}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) - h_q^{q/2}\, h_{q_A}^{-q_A}\, b^{-1/2}\, \widehat{\mu}_{2,T} \right] \\ &\quad - 2\,\Omega^{-1}\left[h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{\widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})}\, \widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})\, \widehat{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) - h_q^{q/2-q_A}\, b^{-1/2}\, \widehat{\mu}_{3,T} \right] \\ &\quad + \Omega^{-1}\, h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{\widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})} \left[f_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) \right]^2 \\ &= \Lambda^{(1)}_{1,T} + \Lambda^{(1)}_{2,T} + \Lambda^{(1)}_{3,T} + \Lambda^{(1)}_{4,T}. \end{aligned}$$

Under the alternative, $f_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})$ differs from $f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})$ in a subset of nonzero Lebesgue measure. This means that the first and third bias terms will no longer converge to μ_1 and μ_3 almost surely, and hence $\Lambda^{(1)}_{1,T} + \Lambda^{(1)}_{2,T} + \Lambda^{(1)}_{3,T}$ becomes of order $O_p\left(h_q^{-q/2}\, b^{-1/2}\right)$. In addition, $\Lambda^{(1)}_{4,T}$ is now of order $O_p\left(T\, h_q^{q/2}\, b^{1/2}\right)$, which ensures unit asymptotic power. ■

B.3 Proof of Theorem 2

The result ensues from Lemmata 4 to 6 along the same lines as in the proof of Theorem 1. ■

B.4 Proof of Theorem 3

(i) The local linear estimator based on realized measures rather than integrated variances reads

$$\begin{aligned}\widehat{\boldsymbol{\beta}}_T^{(M)}(y,\boldsymbol{x}^{(q)}) &= \widehat{\boldsymbol{\beta}}_T(y,\boldsymbol{x}^{(q)}) + \left(\mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{H}_{\boldsymbol{x}^{(q)}}\right)^{-1}\left(\mathcal{H}'_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}_M^{(q)}}\mathcal{Y}_{y_M} - \mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{Y}_y\right)\\
&+\left[\left(\frac{1}{T}\mathcal{H}'_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}_M^{(q)}}\mathcal{H}_{\boldsymbol{x}_M^{(q)}}\right)^{-1} - \left(\frac{1}{T}\mathcal{H}'_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{H}_{\boldsymbol{x}^{(q)}}\right)^{-1}\right]\frac{1}{T}\mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{Y}_y\\
&+\left[\left(\frac{1}{T}\mathcal{H}'_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}_M^{(q)}}\mathcal{H}_{\boldsymbol{x}_M^{(q)}}\right)^{-1} - \left(\frac{1}{T}\mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{H}_{\boldsymbol{x}^{(q)}}\right)^{-1}\right]\frac{1}{T}\left(\mathcal{H}_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}_M^{(q)}}\mathcal{Y}_{y_M} - \mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{Y}_y\right),\end{aligned}$$

where the index M denotes reliance on realized measures and $\frac{1}{T}\left(\mathcal{H}'_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}_M^{(q)}}\mathcal{Y}_{y_M} - \mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{Y}_y\right)$ is a column vector given by

$$\begin{pmatrix} \frac{1}{T}\sum_{t=1}^T\left[\prod_{j=1}^q W_{h_q}(X_{jt,M}-x_j)K_b(Y_{t,M}-y) - \prod_{j=1}^q W_{h_q}(X_{jt}-x_j)K_b(Y_t-y)\right]\\ \frac{1}{T}\sum_{t=1}^T\left[\prod_{j=1}^q W_{h_q}(X_{jt,M}-x_j)K_b(Y_{t,M}-y)(X_{1t,M}-x_1) - \prod_{j=1}^q W_{h_q}(X_{jt}-x_j)K_b(Y_t-y)(X_{1t}-x_1)\right]\\ \vdots\\ \frac{1}{T}\sum_{t=1}^T\left[\prod_{j=1}^q W_{h_q}(X_{jt,M}-x_j)K_b(Y_{t,M}-y)(X_{qt,M}-x_q) - \prod_{j=1}^q W_{h_q}(X_{jt}-x_j)K_b(Y_t-y)(X_{qt}-x_q)\right]\end{pmatrix} \tag{12}$$

We start by bounding the first row of (12), namely,

$$\begin{aligned}&\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})}\left|\frac{1}{T}\sum_{t=1}^T\left(\prod_{j=1}^q W_{h_q}(X_{jt,M}-x_j)K_b(Y_{t,M}-y) - \prod_{j=1}^q W_{h_q}(X_{jt}-x_j)K_b(Y_t-y)\right)\right|\\
&\le \sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})}\frac{1}{T}\sum_{t=1}^T\left|\sum_{i=1}^q\prod_{j=1}^q\frac{1}{h_q^{q+1}b}W'\left(\frac{\widetilde{X}_{jt,M}-x_j}{h_q}\right)K\left(\frac{\widetilde{Y}_{t,M}-y}{b}\right)N_{i,t,M}\right|\\
&+\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})}\frac{1}{T}\sum_{t=1}^T\left|\prod_{j=1}^q\frac{1}{h_q^q b^2}W\left(\frac{\widetilde{X}_{jt,M}-x_j}{h_q}\right)K'\left(\frac{\widetilde{Y}_{t,M}-y}{b}\right)N_{0,t,M}\right|\\
&+\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})}\frac{1}{T}\sum_{t=1}^T\left|\sum_{i=1}^q\prod_{j=1}^q\frac{1}{h_q^{q+1}b^2}W'\left(\frac{\widetilde{X}_{jt,M}-x_j}{h_q}\right)K'\left(\frac{\widetilde{Y}_{t,M}-y}{b}\right)N_{0,t,M}\,N_{i,t,M}\right|,\end{aligned} \tag{13}$$

where $\widetilde{X}_{jt,M}\in(X_{jt,M},X_{jt})$. As for the first term on the right-hand side of (13), it turns out that

$$\begin{aligned}&\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})}\frac{1}{T}\sum_{t=1}^T\left|\sum_{i=1}^q\prod_{j=1}^q\frac{1}{h_q^{q+1}b}W'\left(\frac{\widetilde{X}_{jt,M}-x_j}{h_q}\right)K\left(\frac{\widetilde{Y}_{t,M}-y}{b}\right)N_{i,t,M}\right|\\
&\le q\sup_{i,t}N_{i,t,M}\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})}\frac{1}{T}\sum_{t=1}^T\left|\prod_{j=1}^q\frac{1}{h_q^{q+1}b}W'\left(\frac{\widetilde{X}_{jt,M}-x_j}{h_q}\right)K\left(\frac{\widetilde{Y}_{t,M}-y}{b}\right)\right|\\
&= O_p(1)\sup_{i,t}N_{i,t,M}.\end{aligned}$$

In view of Assumption A6 and Lemma 1 in Corradi et al. (2009),

$$\begin{aligned}\Pr\left(\sup_{1\le t\le T}T^{-\frac{1}{k-1}}a_M^{1/2}\,|N_{i,t,M}|>\varepsilon\right) &\le \sum_{t=1}^T\Pr\left(T^{-\frac{1}{k-1}}a_M^{1/2}\,|N_{i,t,M}|>\varepsilon\right) \le \varepsilon^{-k}\,T^{1-\frac{k}{k-1}}\,a_M^{k/2}\,\mathbb{E}\,|N_{i,t,M}|^k\\
&\le \varepsilon^{-k}\,T^{-\frac{1}{k-1}}\,a_M^{k/2}\,O(a_M^{-k/2})\to 0, \qquad \text{as } M,T\to\infty\end{aligned}$$

meaning that $\sup_{1\le t\le T}|N_{i,t,M}| = O_p\left(T^{\frac{1}{k-1}}\,a_M^{-1/2}\right)$. It is straightforward to show using a similar argument that the second and third terms on the right-hand side of (13), as well as $\frac{1}{T}\left(\mathcal{H}'_{\boldsymbol{x}_M^{(q)}}\mathcal{W}_{\boldsymbol{x}_M^{(q)}}\mathcal{Y}_{y_M} - \mathcal{H}'_{\boldsymbol{x}^{(q)}}\mathcal{W}_{\boldsymbol{x}^{(q)}}\mathcal{Y}_y\right)$, are also of order $O_p\left(T^{\frac{1}{k-1}}\,a_M^{-1/2}\right)$, uniformly on the compact set $\mathcal{C}(Y,\boldsymbol{X}^{(q)})$. Also, it follows from the fact that $\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(d)}}(y|\boldsymbol{x}^{(d)}) = \widehat{\beta}^{(M)}_{0T}(y,\boldsymbol{x}^{(d)})$ with $d\in\{q,q_A\}$ that

$$\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(d)})}\left|\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(d)}}(y|\boldsymbol{x}^{(d)}) - \widehat{f}_{Y|\boldsymbol{X}^{(d)}}(y|\boldsymbol{x}^{(d)})\right| = O_p\left(T^{\frac{1}{k-1}}\,a_M^{-1/2}\right). \tag{14}$$

It is now immediate to see that

$$
\begin{aligned}
\Omega\left(\widehat{\Lambda}_T^{(M)} - \widehat{\Lambda}_T\right) &= h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left[\left(\frac{\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(q)}}(Y_{t,M}|\boldsymbol{X}^{(q)}_{t,M}) - \widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(q_A)}}(Y_{t,M}|\boldsymbol{X}^{(q_A)}_{t,M})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right)^2 \pi(Y_{t,M}, \boldsymbol{X}^{(q)}_{t,M}) \right. \\
&\qquad \left. - \left(\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}^{(q)}_t) - \widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right)^2 \pi(Y_t, \boldsymbol{X}^{(q)}_t) \right] \\
&\quad + h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left(\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(q)}}(Y_{t,M}|\boldsymbol{X}^{(q)}_{t,M}) - \widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(q_A)}}(Y_{t,M}|\boldsymbol{X}^{(q_A)}_{t,M}) \right)^2 \pi(Y_{t,M}, \boldsymbol{X}^{(q)}_{t,M}) \\
&\qquad \times \left[\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t) - \widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(q_A)}}(Y_{t,M}|\boldsymbol{X}^{(q_A)}_{t,M})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(q_A)}}(Y_{t,M}|\boldsymbol{X}^{(q_A)}_{t,M})} \right] \\
&\quad + O_p\left(h_q^{-q/2}\, b^{-1/2}\, T^{\frac{1}{k-1}}\, a_M^{-1/2} \right) \\
&= A_{T,M} + B_{T,M} + O_p\left(h_q^{-q/2}\, b^{-1/2}\, T^{\frac{1}{k-1}}\, a_M^{-1/2} \right) \\
&= A_{T,M} + B_{T,M} + o_p\left(T^{\frac{k+1}{2(k-1)}}\, a_M^{-1/2} \right), \qquad (15)
\end{aligned}
$$

where the last term captures the contribution of the bias terms, namely, $(\widehat{\mu}_{i,T,M} - \widehat{\mu}_{i,T}) = O_p\left(T^{\frac{1}{k-1}}\, a_M^{-1/2}\right)$ for $i \in \{1, 2, 3\}$. Now,

$$
\left| \pi(Y_{t,M}, \boldsymbol{X}^{(q)}_{t,M}) - \pi(Y_t, \boldsymbol{X}^{(q)}_t) \right| \leq \left(\sup_{\mathcal{C}(Y, \boldsymbol{X}^{(q)})} \sum_{i=0}^{q} \partial_i \pi_{t,M} \right) \left(\sup_{i,t} |N_{i,t,M}| \right) = O_p\left(T^{\frac{1}{k-1}}\, a_M^{-1/2} \right),
$$

and so letting $\pi_{t,M} \equiv \pi(Y_{t,M}, \boldsymbol{X}^{(q)}_{t,M})$ and $\pi(Y_t, \boldsymbol{X}_t) \equiv \pi(Y_t, \boldsymbol{X}^{(q)}_t)$ yields

$$
h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left(\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}^{(q)}_t) - \widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right)^2 (\pi_{t,M} - \pi(Y_t, \boldsymbol{X}_t)) = O_p\left(h_q^{-q/2}\, b^{-1/2}\, T^{\frac{1}{k-1}}\, a_M^{-1/2} \right). \qquad (16)
$$

It also follows from (14) that, for $d = \{q, q_A\}$,

$$
h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left(\frac{\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(d)}}(Y_{t,M}|\boldsymbol{X}^{(d)}_{t,M}) - \widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_{t,M}|\boldsymbol{X}^{(d)}_{t,M})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right)^2 \pi_{t,M} = O_p\left(h_q^{q/2}\, b^{1/2}\, T^{\frac{k+1}{k-1}}\, a_M^{-1} \right), \qquad (17)
$$

whereas

$$
\begin{aligned}
&h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left(\frac{\widehat{f}_{Y|\boldsymbol{X}^{(d)}}(Y_{t,M}|\boldsymbol{X}^{(d)}_{t,M}) - \widehat{f}_{Y|\boldsymbol{X}^{(d)}}(Y_t|\boldsymbol{X}^{(d)}_t)}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right)^2 \pi_{t,M} \\
&\quad \leq \left(\sup_{i,t} N^2_{i,t,M} \right) h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left(\frac{\sum_{i=0}^{q} \partial_i \widehat{\boldsymbol{\beta}}_{0T}(\widetilde{Y}_{t,M}|\widetilde{\boldsymbol{X}}^{(d)}_{t,M})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right)^2 \pi_{t,M} = O_p\left(T^{\frac{k+1}{k-1}}\, h_q^{q/2}\, b^{1/2}\, a_M^{-1} \right). \qquad (18)
\end{aligned}
$$

and

$$
\begin{aligned}
&h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \left(\frac{\widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}^{(q)}_t) - \widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right) \left(\frac{\widehat{f}^{(M)}_{Y|\boldsymbol{X}^{(d)}}(Y_{t,M}|\boldsymbol{X}^{(d)}_{t,M}) - \widehat{f}_{Y|\boldsymbol{X}^{(q)}}(Y_{t,M}|\boldsymbol{X}^{(d)}_{t,M})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}^{(q_A)}_t)} \right) \\
&\quad = O_p\left(T^{\frac{k+1}{2(k-1)}}\, \sqrt{\ln T}\, a_M^{-1/2} \right). \qquad (19)
\end{aligned}
$$

Altogether, (16) to (19) also imply that the two other cross-terms are of order $O_p\left(T^{\frac{k+1}{2(k-1)}}\, \sqrt{\ln T}\, a_M^{-1/2}\right)$, so that $A_{T,M} = O_p\left(T^{\frac{k+1}{2(k-1)}}\, \sqrt{\ln T}\, a_M^{-1/2}\right)$. Finally, given that $B_{T,M}$ is of smaller probability order than $A_{T,M}$, it suffices to follow the same steps as in the proof of Theorem 1(i) to complete the proof of statement (i).

(*ii*) Under the alternative $\mathbb{H}_A$, $f_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})$ and $f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})$ differ in a subset of nonzero Lebesgue measure. This implies that the terms in (16) and (19) become of order $O_p\left(T^{\frac{k}{k-1}}\, h_q^{q/2}\, b^{1/2}\, a_M^{-1/2}\right)$ under the alternative, though there is no change in the probability orders of (17) and (18). Altogether, this shows that $\widehat{\Lambda}_T^{(M)} - \widehat{\Lambda}_T = O_p\left(T^{\frac{k}{k-1}}\, h_q^{q/2}\, b^{1/2}\, a_M^{-1/2}\right)$ under $\mathbb{H}_A$, completing the proof. ■

B.5 Proof of Theorem 4

The result ensues along the same lines as in the proof of Theorem 3.

B.6 Proof of Lemma 1

Under the null $\mathbb{H}_0$, $f_{Y|\boldsymbol{X}^{(q)}}(y|\boldsymbol{x}^{(q)})$ coincides almost surely with $f_{Y|\boldsymbol{X}^{(q_A)}}(y|\boldsymbol{x}^{(q_A)})$ and hence

$$
\begin{aligned}
\widehat{\Lambda}_{1,T} &\equiv \Omega^{-1}\left\{ h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \left[\frac{\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t \mid \boldsymbol{X}_t^{(q)})}{\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t \mid \boldsymbol{X}_t^{(q_A)})} \right]^2 - h_q^{-q/2}\, b^{-1/2}\, \mu_1 \right\} \\
&= \Omega^{-1}\left\{ h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \left[\frac{\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t \mid \boldsymbol{X}_t^{(q)})}{f_{Y|\boldsymbol{X}^{(q)}}(Y_t \mid \boldsymbol{X}_t^{(q_A)})} \right]^2 - h_q^{-q/2}\, b^{-1/2}\, \mu_1 \right\} \\
&\quad + \frac{h_q^{q/2}\, b^{1/2}}{\Omega} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)})\, \widehat{\epsilon}^{\,2}_{Y|\boldsymbol{X}^{(q)}}(Y_t \mid \boldsymbol{X}_t^{(q)}) \left[\frac{1}{\widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t \mid \boldsymbol{X}_t^{(q_A)})} - \frac{1}{f^{2}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t \mid \boldsymbol{X}_t^{(q_A)})} \right] \\
&= \widehat{\Lambda}_{11,T} + \widehat{\Lambda}_{12,T}.
\end{aligned}
$$

As per $\widehat{\Lambda}_{12,T}$, the bandwidth condition (*vii*) in Lemmata 1 and 2 ensures that

$$
\sup_{\mathcal{C}(Y,\boldsymbol{X}^{(q)})} \left\{ \widehat{f}^{\,2}_{Y|\boldsymbol{X}^{(q_A)}}(y|\boldsymbol{x}^{(q_A)}) - f^{2}_{Y|\boldsymbol{X}^{(q_A)}}(y|\boldsymbol{x}^{(q_A)}) \right\} = O_p\left(T^{-1/2}\sqrt{\ln T}\, h_{q_A}^{-q_A/2}\, b^{-1/2}\right),
$$

and so it follows from the bandwidth condition (*viii*) in Lemma 1 that $\widehat{\Lambda}_{12,T} = o_p(1)$ given that

$$
h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \left[\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)}) - \widehat{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)}) \right]^2 = O_p\left(h_q^{-q/2}\, b^{-1/2}\right).
$$

As $\widehat{\Lambda}_{11,T}$ concerns only $\boldsymbol{X}^{(q)}$, we hereafter suppress the superscript index from the conditioning state vector and let $m(\boldsymbol{x}, y) = \mathbb{E}\big[K_b(Y_t - y) \,\big|\, \boldsymbol{X}_t = \boldsymbol{x}\big]$. By the same reasoning as in the proof of Theorem 1 in Aït-Sahalia et al. (2009b), the bandwidth conditions (*i*) to (*vi*) ensure that

$$
\begin{aligned}
I_T &= h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \pi(Y_t, \boldsymbol{X}_t^{(q)}) \left[\frac{\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t \mid \boldsymbol{X}_t^{(q)})}{f_{Y|\boldsymbol{X}^{(q)}}(Y_t \mid \boldsymbol{X}_t^{(q)})} \right]^2 \\
&= h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t)}{f^2_{Y,\boldsymbol{X}}(Y_t, \boldsymbol{X}_t)} \left(\frac{1}{T} \sum_{\tau=1}^{T} \boldsymbol{W}_{h_q}(\boldsymbol{X}_\tau - \boldsymbol{X}_t)\big[K_b(Y_\tau - Y_t) - m(\boldsymbol{X}_t, Y_t)\big] \right)^2 + O_p(T^{-1/2}\sqrt{\ln T}\, h_q^{-q/2}\, b^{-1/2}) \\
&= \widetilde{I}_T + o_p(1).
\end{aligned}
$$

Letting now

$$
\phi(t, \tau, k) = \frac{1}{T^2} \frac{\pi(Y_t, \boldsymbol{X}_t)}{f^2_{Y,\boldsymbol{X}}(Y_t, \boldsymbol{X}_t)} \boldsymbol{W}_{h_q}(\boldsymbol{X}_\tau - \boldsymbol{X}_t)\Big[K_b(Y_\tau - Y_t) - m(\boldsymbol{X}_\tau, Y_t)\Big]\, \boldsymbol{W}_{h_q}(\boldsymbol{X}_k - \boldsymbol{X}_t)\, \Big[K_b(Y_k - Y_t) - m(\boldsymbol{X}_k, Y_t)\Big]
$$

and $\bar{\phi}(t, \tau, k) = \phi(t, \tau, k) + \phi(t, k, \tau) + \phi(\tau, t, k) + \phi(\tau, k, t) + \phi(k, t, \tau) + \phi(k, \tau, t)$ yields

$$
\begin{aligned}
\widetilde{I}_T &= h_q^{q/2}\, b^{1/2} \sum_{t<\tau<k}^{T} \bar{\phi}(t, \tau, k) + h_q^{q/2}\, b^{1/2} \sum_{t\neq\tau}^{T} \Big[\phi(t, \tau, \tau) + \phi(\tau, t, \tau) + \phi(\tau, \tau, t)\Big] + h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \phi(t, t, t) \\
&= \widetilde{I}_{1,T} + \widetilde{I}_{2,T} + \widetilde{I}_{3,T}.
\end{aligned}
$$

As in Aït-Sahalia et al. (2009b), we must now demonstrate the following statements to conclude the proof; the only difference is that we must also account for the higher dimensionality of the conditioning set ($q > 1$).

(a) $\widetilde{I}_{1,T} = (T-2)\, h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \bar{\phi}(t,\tau) + o_p(1)$, where $\bar{\phi}(t,\tau) = \int \bar{\phi}(t,\tau,k)\,\mathrm{d}F(y_k, \boldsymbol{x}_k)$.

(b) $\widetilde{I}_{2,T} = \frac{1}{2}\, T(T-1)\, h_q^{q/2}\, b^{1/2}\, \widetilde{\phi}(0) + o_p(1)$, where $\widetilde{\phi}(0) = \mathbb{E}[\widetilde{\phi}(t)]$, $\widetilde{\phi}(t) = \int \widetilde{\phi}(t,\tau)\,\mathrm{d}F(y_\tau, \boldsymbol{x}_\tau)$, and $\widetilde{\phi}(t,\tau) = \phi(t,t,\tau) + \phi(t,\tau,t) + \phi(\tau,t,t) + \phi(\tau,t,\tau) + \phi(\tau,\tau,t) + \phi(t,\tau,\tau)$.

(c) $\widetilde{I}_{3,T} = o_p(1)$.

(d) It also holds that

$$\tfrac{1}{2}\, T(T-1)\, h_q^{q/2}\, b^{1/2}\, \widetilde{\phi}(0) = h_q^{-q/2}\, b^{-1/2}\, C_1(K)\, C_1(\boldsymbol{W}) \int \pi(y,\boldsymbol{x})\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x} - h_q^{-q/2}\, b^{1/2}\, C_1(\boldsymbol{W}) \int \mathbb{E}\big[\pi(Y,\boldsymbol{X}) \,\big|\, \boldsymbol{X} = \boldsymbol{x}\big]\,\mathrm{d}\boldsymbol{x} + o(1) \tag{20}$$

and that

$$\Omega^2 = \lim_{T\to\infty} \mathrm{Var}\left[(T-2) h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \bar{\phi}(t,\tau)\right] = 2\, C_2(K)\, C_2(\boldsymbol{W}) \int \pi^2(y,\boldsymbol{x})\,\mathrm{d}y\,\mathrm{d}\boldsymbol{x}. \tag{21}$$

(e) $(T-2)\, h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \bar{\phi}(t,\tau) \xrightarrow{d} N(0,\Omega^2)$.

B.6.1 Proof of statement (a)

It follows from the Hoeffding decomposition that

$$\widetilde{I}_{1,T} = h_q^{q/2}\, b^{1/2} \sum_{t<\tau<k} \Phi(t,\tau,k) + (T-2)\, h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \bar{\phi}(t,\tau), \tag{22}$$

where $\Phi(t,\tau,k) = \bar{\phi}(t,\tau,k) - \bar{\phi}(t,\tau) - \bar{\phi}(t,k) - \bar{\phi}(\tau,k)$. To show that the first term on the right-hand side of (22) is of order $o_p(1)$, it suffices to apply Lemma 5(i) in Aït-Sahalia et al. (2009b) with $\delta = 1/3$. This results in $\mathbb{E}\Big(\widetilde{I}_{1,T}^{\,2}\Big) = O\left(T^{-1}\, h_q^{3q/2}\, b^{-3/2}\right)$, which is of order $o(1)$ by condition (ii). ■

B.6.2 Proof of statement (b)

As before, applying the Hoeffding decomposition yields

$$\begin{aligned} h_q^{q/2}\, b^{1/2}\, \widetilde{I}_{2,T} &= h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \widetilde{\phi}(t,\tau) \\ &= h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \Big[\widetilde{\phi}(t,\tau) - \widetilde{\phi}(t) - \widetilde{\phi}(\tau) - \widetilde{\phi}(0)\Big] + (T-1)\, h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \Big[\widetilde{\phi}(t) - \widetilde{\phi}(0)\Big] \\ &\quad + \tfrac{1}{2}\, T(T-1)\, h_q^{q/2}\, b^{1/2}\, \widetilde{\phi}(0). \end{aligned}$$

Lemma 5(ii) in Aït-Sahalia et al. (2009b) with $\delta = 1$ then dictates that

$$h_q^{q/2}\, b^{1/2} \sum_{t<\tau} \Big[\widetilde{\phi}(t,\tau) - \widetilde{\phi}(t) - \widetilde{\phi}(\tau) - \widetilde{\phi}(0)\Big] = O_p\left(T^{-1}\, h_q^{-5q/4}\, b^{-5/4}\right),$$

which is of order $o_p(1)$ due to the bandwidth condition (ii). Under Assumption A4, the central limit for β-mixing processes ensures that

$$(T-1)\, h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \Big[\widetilde{\phi}(t) - \widetilde{\phi}(0)\Big] = O_p\left(T^{-1}\, h_q^{-q}\, b^{-1}\right) = o_p(1).$$ ■

B.6.3 Proof of statement (c)

It is immediate to see that

$$\widetilde{I}_{3,T} = h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \phi(t,t,t) = O_p\left(T\, h_q^{-3q/2}\, b^{-3/2}\right),$$

which is of order $o_p(1)$ by condition (ii). ■

B.6.4 Proof of statement (d)

As for (20) and (21), the result follows along similar lines of the proof of claim (d) in Aït-Sahalia et al. (2009b). ■

B.6.5 Proof of statement (e)

It suffices to apply Fan and Li's (1999) central limit theorem for degenerate U-statistics of absolutely regular processes for U-statistics to obtain the desired result (see Amaro de Matos and Fernandes, 2007). See also Aït-Sahalia et al. (2009b) and Gao and Hong (2008) for alternative central limit theorems that deal with degenerate U-statistics of α-mixing processes. ■

B.7 Proof of Lemma 2

Let $\bar{\psi}(t,\tau)$ and $\widetilde{\psi}(0)$ respectively denote the counterparts of $\bar{\phi}(t,\tau)$ and $\widetilde{\phi}(0)$ once we substitute

$$\psi(t,\tau,k) = \frac{1}{T^2}\,\frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{f^2_{Y,\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})}\,\Big\{\widetilde{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_\tau^{(q_A)}-\boldsymbol{X}_t^{(q_A)})\Big[K_b(Y_\tau-Y_t)-m(\boldsymbol{X}_\tau^{(q_A)},Y_t)\Big] \times\ \widetilde{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_k^{(q_A)}-\boldsymbol{X}_t^{(q_A)})\Big[K_b(Y_k-Y_t)-m(\boldsymbol{X}_k^{(q_A)},Y_t)\Big]\Big\}$$

for $\phi(t,\tau,k)$. Applying the same argument we put forth in the proof of Lemma 1 then yields

$$\begin{aligned}
J_T &= h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{f^2_{Y,\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})}\Big[\widehat{f}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})-f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})\Big]^2\\
&= (T-2)\,h_q^{q/2}\,b^{1/2}\sum_{t<\tau}\bar{\psi}(t,\tau)+\tfrac{1}{2}\,T(T-1)\,h^{q/2}\,b^{1/2}\,\widetilde{\psi}(0)+o_p(1),
\end{aligned}$$

whose first term on the right-hand side satisfies the central limit theorem for U-statistics. In addition,

$$\begin{aligned}
\tfrac{1}{2}\,\widetilde{\psi}(0) &= \frac{1}{T^2}\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,\widetilde{\boldsymbol{W}}^2_{h_{q_A}}(\boldsymbol{x}_j^{(q_A)}-\boldsymbol{x}_i^{(q_A)})\Big(K_b(y_j-y_i)-m(\boldsymbol{x}_j^{(q_A)},y_i)\Big)^2\,\mathrm{d}F(y_i,\boldsymbol{x}_i^{(q)}\,\mathrm{d}F(y_j,\boldsymbol{x}_j^{(q_A)})\\
&= \frac{1}{T^2}\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,\widetilde{\boldsymbol{W}}^2_{h_{q_A}}(\boldsymbol{x}_j^{(q_A)}-\boldsymbol{x}_i^{(q_A)})K_b^2(y_j-y_i)\,\mathrm{d}F(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}F(y_j,\boldsymbol{x}_j^{(q_A)})\\
&\quad -2\frac{1}{T^2}\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,\widetilde{\boldsymbol{W}}^2_{h_{q_A}}(\boldsymbol{x}_j^{(q_A)}-\boldsymbol{x}_i^{(q_A)})K_b(y_j-y_i)m(\boldsymbol{x}_j^{(q_A)},y_i)\,\mathrm{d}F(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}F(y_j,\boldsymbol{x}_j^{(q_A)})\\
&\quad -\frac{1}{T^2}\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,\widetilde{\boldsymbol{W}}^2_{h_{q_A}}(\boldsymbol{x}_j^{(q_A)}-\boldsymbol{x}_i^{(q_A)})m^2(\boldsymbol{x}_j^{(q_A)},y_i)\,\mathrm{d}F(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}F(\boldsymbol{x}_j^{(q_A)})\\
&= \frac{1}{T^2}\,\frac{1}{h_{q_A}^{q_A}\,b}\int\widetilde{\boldsymbol{W}}^2(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\int\boldsymbol{K}^2(v)\,\mathrm{d}v\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f(y_i,\boldsymbol{x}_i^{(q_A)})}\,f(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)}\,\big\{1+O\big(h^2_{q_A}+b^s\big)\big\}\\
&\quad -\frac{1}{T^2}\,\frac{1}{h_{q_A}^{q_A}}\int\widetilde{\boldsymbol{W}}^2(\boldsymbol{u})\,\mathrm{d}\boldsymbol{u}\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,f(y_i,\boldsymbol{x}_i^{(q)})\,m^2(\boldsymbol{x}_i^{(q_A)},y_i)\,f(\boldsymbol{x}_i^{(q_A)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)}\,\big\{1+O(h^2_{q_A}+b^s)\big\},
\end{aligned}$$

where the last equality follows from a Taylor expansion with $\boldsymbol{u} = (\boldsymbol{x}_j^{(q_A)}-\boldsymbol{x}_i^{(q_A)})/h_{q_A}$ and $v = (\boldsymbol{y}_j-\boldsymbol{y}_i)/b$ given that $\mathbb{E}\Big[K_b(y_j-y_i)|Y=y_i,\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big] = m(\boldsymbol{x}_i^{(q_A)},y_i) = f(y_i|\boldsymbol{x}_i^{(q_A)})\,\big\{1+O(h^2_{q_A}+b^s)\big\}$. In addition, it ensues from

$$\begin{aligned}
\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f(y_i,\boldsymbol{x}_i^{(q_A)})}\,f(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)} &= \int\pi(y_i,\boldsymbol{x}_i^{(q)})\,f(\boldsymbol{x}_i^{(q_B)}|y_i,\boldsymbol{x}_i^{(q_A)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)}\\
&= \int\mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|Y=y_i,\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q_A)}
\end{aligned}$$

and

$$\begin{aligned}
\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,f^2(y_i|\boldsymbol{x}_i^{(q_A)})\,f(y_i,\boldsymbol{x}_i^{(q)})\,f(\boldsymbol{x}_i^{(q_A)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)} &= \int\pi(y_i,\boldsymbol{x}_i^{(q)})\,f(y_i,\boldsymbol{x}_i^{(q_B)}|\boldsymbol{x}_i^{(q_A)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)}\\
&= \int\mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}\boldsymbol{x}_i^{(q_A)},
\end{aligned}$$

that

$$\tfrac{1}{2}\,T(T-1)\,h_q^{q/2}\,b^{1/2}\,\widetilde{\psi}(0) = h_q^{q/2}\,h_{q_A}^{-q_A}\,b^{-1/2}\,C_1(K)\,C_1(\widetilde{\boldsymbol{W}}) \int \mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)}|Y=y_i,\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}\boldsymbol{x}_i^{(q_A)}\,\mathrm{d}y_i$$
$$+\,h_q^{q/2}\,h_{q_A}^{-q_A}\,b^{1/2}\,C_1(\widetilde{\boldsymbol{W}}) \int \mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}\boldsymbol{x}_i^{(q_A)} + o(1),$$

completing the proof. ■

B.8 Proof of Lemma 3

Let

$$\varphi(t,\tau,k) = \frac{1}{T^2}\,\frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})}{f_{Y,\boldsymbol{X}^{(q)}}(Y_t,\boldsymbol{X}_t^{(q)})\,f_{Y,\boldsymbol{X}^{(q_A)}}(Y_t,\boldsymbol{X}_t^{(q_A)})}\,\boldsymbol{W}_{h_q}(\boldsymbol{X}_\tau^{(q)}-\boldsymbol{X}_t^{(q)})\Big[K_b(Y_\tau-Y_t)-m(\boldsymbol{X}_\tau^{(q)},Y_t)\Big]$$
$$\times\,\widetilde{\boldsymbol{W}}_{h_{q_A}}(\boldsymbol{X}_k^{(q_A)}-\boldsymbol{X}_t^{(q_A)})\Big[K_b(Y_k-Y_t)-m(\boldsymbol{X}_k^{(q_A)},Y_t)\Big].$$

Proceeding along the same line as in the proof of Lemma 1 then yields

$$h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\frac{\pi(Y_t,\boldsymbol{X}_t^{(q)})\,\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})\,\widehat{\epsilon}_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})}{f_{Y|\boldsymbol{X}^{(q)}}(Y_t|\boldsymbol{X}_t^{(q)})\,f_{Y|\boldsymbol{X}^{(q_A)}}(Y_t|\boldsymbol{X}_t^{(q_A)})} = (T-2)\,h_q^{q/2}\,b^{1/2}\sum_{t<\tau}\bar{\varphi}(t,\tau)$$
$$+\,\tfrac{1}{2}\,T(T-1)\,h_q^{q/2}\,b^{1/2}\,\widetilde{\varphi}(0)+o_p(1).$$

Let now $\boldsymbol{u}=(\boldsymbol{x}_j^{(q_A)}-\boldsymbol{x}_i^{(q_A)})/h_{q_A}$, $v=(y_j-y_i)/b$, and $\boldsymbol{z}=(\boldsymbol{x}_j^{(q)}-\boldsymbol{x}_i^{(q)})/h_q$. Given that under the null $\mathbb{H}_0$, $\mathbb{E}\Big[K_b(y_j-y_i)|Y=y_i,\boldsymbol{X}^{(q)}=\boldsymbol{x}_i^{(q)}\Big]=m(\boldsymbol{x}_i^{(q)},y_i)=m(\boldsymbol{x}_i^{(q_A)},y_i)$, it follows that

$$\tfrac{1}{2}\,\widetilde{\varphi}(0) = T^{-2}h_q^{-q_A}\,b^{-1}\,C_1(K)\,\widetilde{\boldsymbol{W}}(0)\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f(y_i,\boldsymbol{x}_i^{(q_A)})}\,f(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)}\,\big\{1+O\big(h_{q_A}^2+b^s\big)\big\}$$
$$-\,T^{-2}h_q^{-q_A}\,\widetilde{\boldsymbol{W}}(0)\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2(y_i,\boldsymbol{x}_i^{(q_A)})}\,f(y_i,\boldsymbol{x}_i^{(q)})\,m^2(\boldsymbol{x}_i^{(q_A)},y_i)\,f(\boldsymbol{x}_i^{(q_A)})\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q)}\,\big\{1+O\big(h_{q_A}^2+b^s\big)\big\}$$
$$= T^{-2}h_q^{-q_A}\,b^{-1}\,C_1(K)\,\widetilde{\boldsymbol{W}}(0)\int\mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|Y=y_i,\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q_A)}\,\big\{1+O\big(h_{q_A}^2+b^s\big)\big\}$$
$$-\,T^{-2}h_q^{-q_A}\,\widetilde{\boldsymbol{W}}(0)\int\mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}\boldsymbol{x}_i^{(q_A)}\,\big\{1+O\big(h_{q_A}^2+b^s\big)\big\}.$$

This means that

$$\tfrac{1}{2}\,T(T-1)\,h_q^{q/2}\,b^{1/2}\,\widetilde{\varphi}(0) = h_q^{q/2-q_A}\,b^{-1/2}\,C_1(K)\,\widetilde{\boldsymbol{W}}(0)\int\mathbb{E}\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|Y=y_i,\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}y_i\,\mathrm{d}\boldsymbol{x}_i^{(q_A)}$$
$$-\,h_q^{q/2-q_A}\,\widetilde{\boldsymbol{W}}(0)\int\Big[\pi(y_i,\boldsymbol{x}_i^{(q)})|\boldsymbol{X}^{(q_A)}=\boldsymbol{x}_i^{(q_A)}\Big]\,\mathrm{d}\boldsymbol{x}_i^{(q_A)},$$

which completes the proof. ■

B.9 Proof of Lemma 4

As in the proof of Lemma 1, we suppress the superscript index from the conditioning state vector for notational simplicity. Letting $\bar{m}(\boldsymbol{x},y)=\mathbb{E}[K_b(Y_t-y)|\boldsymbol{X}_t=\boldsymbol{x}]$ then yields

$$\bar{I}_T = h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\pi(Y_t,\boldsymbol{X}_t)\left(\frac{\bar{f}_{Y|\boldsymbol{X}}(Y_t|\boldsymbol{X}_t)-f_{Y|\boldsymbol{X}}(Y_t|\boldsymbol{X}_t)}{f_{Y|\boldsymbol{X}^{(q)}}(Y_t\,\big|\,\boldsymbol{X}_t^{(q)})}\right)^2$$
$$= h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\frac{\pi(Y_t,\boldsymbol{X}_t)}{f_{Y,\boldsymbol{X}}^2(Y_t,\boldsymbol{X}_t)}\left[\frac{1}{T}\sum_{\tau=1}^{T}\bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_\tau-\boldsymbol{X}_t)\Big(K_b(Y_\tau-Y_t)-\bar{m}(\boldsymbol{X}_t,Y_t)\Big)\right]^2$$
$$+\,h_q^{q/2}\,b^{1/2}\sum_{t=1}^{T}\pi(Y_t,\boldsymbol{X}_t)\Big(\bar{m}(\boldsymbol{X}_t,Y_t)-f_{Y|\boldsymbol{X}}(Y_t|\boldsymbol{X}_t)\Big)^2$$

$$- 2h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t)}{f_{Y,\boldsymbol{X}}(Y_t, \boldsymbol{X}_t)} \Big(\bar{m}(\boldsymbol{X}_t, Y_t) - f_{Y|\boldsymbol{X}}(Y_t|\boldsymbol{X}_t)\Big) \frac{1}{T} \sum_{\tau=1}^{T} \bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_\tau - \boldsymbol{X}_t)\Big(K_b(Y_\tau - Y_t) - \bar{m}(\boldsymbol{X}_t, Y_t)\Big)$$
$$= \bar{I}_{1,T} + \bar{I}_{2,T} + \bar{I}_{3,T}.$$

It then follows from condition (iii) that $\bar{I}_{2,T} = O\left(T\, h_q^{2s+q/2}\, b^{1/2}\right) = o(1)$. If one accounts for the fact that $\bar{\boldsymbol{W}}$ and K are both of order $s > 2$, a similar argument as in the proof of Lemma 1 then gives way to

$$\begin{aligned}
\bar{I}_{1,T} &= h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t)}{f^2_{Y,\boldsymbol{X}}(Y_t, \boldsymbol{X}_t)} \left[\frac{1}{T} \sum_{\tau=1}^{T} \bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_\tau - \boldsymbol{X}_t)\Big(K_b(Y_\tau - Y_t) - \bar{m}(\boldsymbol{X}_t, Y_t)\Big)\right]^2 \\
&\quad + O_p\left(T^{-1}\, h_q^{-2q}\, b^{-1}\, \ln T\right) + O_p\left(h_q^{s-q/2}\, b^{-1/2}\right) \\
&= \widetilde{\bar{I}}_{1,T} + o_q(1)
\end{aligned}$$

given that conditions (ii) and (v) hold. In addition, it also turns out that

$$\begin{aligned}
\bar{I}_{3,T} &= -2\, h_q^{q/2}\, b^{1/2} \sum_{t=1}^{T} \frac{\pi(Y_t, \boldsymbol{X}_t)}{f_{Y,\boldsymbol{X}}(Y_t, \boldsymbol{X}_t)} \Big(\bar{m}(\boldsymbol{X}_t, Y_t) - f_{Y|\boldsymbol{X}}(Y_t|\boldsymbol{X}_t)\Big) \frac{1}{T} \sum_{\tau=1}^{T} \bar{\boldsymbol{W}}_{h_q}(\boldsymbol{X}_\tau - \boldsymbol{X}_t)\Big(K_b(Y_\tau - Y_t) - \bar{m}(\boldsymbol{X}_t, Y_t)\Big) \\
&\quad + O_p\left(h_q^{-q/2}\, b^s\, \ln T\right) \\
&= \widetilde{\bar{I}}_{3,T} + o_p(1)
\end{aligned}$$

in view of condition (vi). To complete the proof, it now suffices to develop a similar argument as in the proof of Lemma 1 accounting for the fact that both kernels are of order s. ■

B.10 Proofs of Lemmata 5 and 6

We omit the proofs because they are almost exactly the same as the proofs of Lemmata 2 and 3. It indeed suffices to apply the same line of reasoning to derive the results in a straightforward manner. ■

B.11 Proof of Theorem 5

We denote by $\Pr_*$ the probability distribution induced by the bootstrap sampling, with expectation and variance operators given by $\mathbb{E}_*$ and Var_*, respectively. In addition, we also let $O_p^*(1)$ and $o_p^*(1)$ denote the orders of magnitude according to the bootstrap-induced probability law.

Both local-linear and kernel smoothing results follow straightforwardly once we prove the bootstrap versions of Lemmata 1 to 3 and of Lemmata 4 to 6, respectively. As the proofs are very similar, in what follows, we restrict attention to the bootstrap test based on local linear smoothing. We start with the bootstrap counterpart of Lemma 1. As in the latter's proof, it turns out that

$$\begin{aligned}
\widehat{\Lambda}^*_{1,\mathcal{T}} &= h_{*q}^{q/2}\, b_*^{1/2} \sum_{t=1}^{\mathcal{T}} \pi(Y_t^*, \boldsymbol{X}_t^{*(q)}) \left[\frac{\widehat{f}^*_{Y|\boldsymbol{X}^{(q)}}(Y_t^* \mid \boldsymbol{X}_t^{*(q)}) - f_{Y|\boldsymbol{X}^{(q)}}(Y_t^* \mid \boldsymbol{X}_t^{*(q)})}{\widehat{f}^*_{Y|\boldsymbol{X}^{(q_A)}}(Y_t^* \mid \boldsymbol{X}_t^{*(q_A)})}\right]^2 - h_{*q}^{-q/2}\, b_*^{-1/2}\, \mu_1 \\
&= h_{*q}^{q/2}\, b_*^{1/2} \sum_{t=1}^{\mathcal{T}} \frac{\pi(Y_t^*, \boldsymbol{X}_t^{*(q)})}{f^2_{Y,\boldsymbol{X}^{(q)}}(Y_t^*, \boldsymbol{X}_t^{*(q)})} \left(\frac{1}{\mathcal{T}} \sum_{\tau=1}^{\mathcal{T}} \boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_\tau^{*(q)} - \boldsymbol{X}_t^{*(q)})\Big[K_{b_*}(Y_\tau^* - Y_t^*) - m(\boldsymbol{X}_\tau^{*(q)}, Y_t^*)\Big]\right)^2 \\
&\quad -h_{*q}^{-q/2}\, b_*^{-1/2}\, \mu_1 + o_p^*(1).
\end{aligned}$$

Let now

$$\begin{aligned}
\phi_*(k, j, i) = \mathcal{T}^{-2}\, \frac{\pi(Y_k^*, \boldsymbol{X}_k^{*(q)})}{f^2_{Y,\boldsymbol{X}^{(q)}}(Y_k^*, \boldsymbol{X}_k^{*(q)})}\, &\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_j^{*(q)} - \boldsymbol{X}_k^{*(q)})\Big[K_{b_*}(Y_j^* - Y_k^*) - m(\boldsymbol{X}_j^{*(q)}, Y_k^*)\Big] \\
&\times \boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_i^{*(q)} - \boldsymbol{X}_k^{*(q)})\Big[K_{b_*}(Y_i^* - Y_k^*) - m(\boldsymbol{X}_i^{*(q)}, Y_k^*)\Big].
\end{aligned}$$

Taking conditional expectation over bootstrap samples given $(Y_k, \boldsymbol{X}_k^{(q)})$ then yields

$$\phi_*(j, i) = \mathbb{E}_*\left[\phi_*(k, j, i)\,\Big|\, Y_k, \boldsymbol{X}_k^{(q)}\right]$$

$$= \mathcal{T}^{-2}\,\frac{1}{T}\sum_{k=1}^{T}\frac{\pi(Y_k,\boldsymbol{X}_k^{(q)})}{f^2_{Y,\boldsymbol{X}^{(q)}}(Y_k,\boldsymbol{X}_k^{(q)})}\,\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_j^{*(q)}-\boldsymbol{X}_k^{(q)})\Big[K_{b_*}(Y_j^*-Y_k)-m(\boldsymbol{X}_j^{*(q)},Y_k)\Big]$$

$$\times\,\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_i^{*(q)}-\boldsymbol{X}_k^{(q)})\Big[K_{b_*}(Y_i^*-Y_k)-m(\boldsymbol{X}_i^{*(q)},Y_k)\Big]$$

$$= \mathcal{T}^{-2}\int\frac{\pi(y_k,\boldsymbol{x}_k^{(q)})}{f^2_{Y,\boldsymbol{X}^{(q)}}(y_k,\boldsymbol{x}_k^{(q)})}\,\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_j^{*(q)}-\boldsymbol{x}_k^{(q)})\Big[K_{b_*}(Y_j^*-y_k)-m(\boldsymbol{X}_j^{*(q)},y_k)\Big]$$

$$\times\,\boldsymbol{W}_{h_{*q}}(\boldsymbol{X}_i^{*(q)}-\boldsymbol{x}_k^{(q)})\Big[K_{b_*}(Y_i^*-y_k)-m(\boldsymbol{X}_i^{*(q)},y_k)\Big]\,\mathrm{d}\boldsymbol{x}_k^{(q)}\,\mathrm{d}y_k+o_p^*(\mathcal{T}^{-2}h_{*q}^{-q/2}b_*^{-1/2}),$$

and so

$$\phi_*(0)=\mathbb{E}_*[\phi_*(i,j,j)+\phi_*(j,i,i)]$$

$$=2\,\mathcal{T}^{-2}\,\frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T}\frac{\pi(Y_i,\boldsymbol{X}_i^{(q)})}{f^2_{Y,\boldsymbol{X}^{(q)}}(Y_i,\boldsymbol{X}_i^{q})}\,\boldsymbol{W}^2_{h_{*q}}(\boldsymbol{X}_j^{(q)}-\boldsymbol{X}_i^{(q)})\,\Big(K_{b_*}(Y_j-Y_i)-m(\boldsymbol{X}_j^{(q)},Y_i)\Big)^2$$

$$=2\,\mathcal{T}^{-2}\int\frac{\pi(y_i,\boldsymbol{x}_i^{(q)})}{f^2_{Y,\boldsymbol{X}^{(q)}}(y_i,\boldsymbol{x}_i^{(q)})}\,\boldsymbol{W}^2_{h_{*q}}(\boldsymbol{x}_j^{(q)}-\boldsymbol{x}_i^{(q)})\,\Big(K_{b_*}(y_j-y_i)-m(\boldsymbol{x}_j^{(q)},y_i)\Big)^2\,\mathrm{d}F_{Y,\boldsymbol{X}^{(q)}}(y_i,\boldsymbol{x}_i^{(q)})\,\mathrm{d}F_{Y,\boldsymbol{X}^{(q)}}(y_j,\boldsymbol{x}_j^{(q)})$$

$$+o_p\left(\mathcal{T}^{-2}h_{*q}^{-q/2}b_*^{-1/2}\right).$$

As in statement **(d)** in the proof of Lemma 1, it then follows that

$$\tfrac{1}{2}\,\mathcal{T}(\mathcal{T}-1)\,h_{*q}^{q/2}\,b_*^{1/2}\,\phi_*(0)=h_{*q}^{-q/2}b_*^{-1/2}\mu_1+o_p^*(1)$$

and, as $\mathcal{T}/T\to 0$,

$$(\mathcal{T}-2)\,h_{*q}^{q/2}\,b_*^{1/2}\sum_{j<i}^{\mathcal{T}}\phi_*(j,i)=(\mathcal{T}-2)\,h_{*q}^{q/2}\,b_*^{1/2}\sum_{j<i}^{\mathcal{T}}\Big(\phi_*(j,i)-\mathbb{E}_*\left[\phi_*(j,i)\right]\Big)+o_p^*(1). \tag{23}$$

In view that $\mathrm{Var}_*\left\{(\mathcal{T}-2)\,h_{*q}^{q/2}\,b_*^{1/2}\sum_{t<\tau}\Big(\phi_*(j,i)-\mathbb{E}_*[\phi_*(j,i)]\Big)\right\}=\Omega+o_p(1)$, the first term on the right-hand side of (23) weakly converges to $N(0,\Omega)$ as both T and $\mathcal{T}$ go to infinity, thus mimicking the limiting distribution of $(T-2)\,h_{*q}^{q/2}\,b_*^{1/2}\sum_{j<i}\phi(j,i)$.

Define next $\psi_*(k,j,i)$, $\psi_*(j,i)$ and $\psi_*(0)$ analogously to $\phi_*(k,j,i)$, $\phi_*(j,i)$ and $\phi_*(0)$ for $\boldsymbol{X}^{(q_A)}$, with $\widetilde{\boldsymbol{W}}_{h_{*q_A}}$ replacing $\boldsymbol{W}_{h_{*q}}$. As $\mathcal{T}/T\to 0$, it is possible to show that $(\mathcal{T}-2)h_{*q}^{q/2}b_*^{1/2}\sum_{t<\tau}^{\mathcal{T}}\psi_*(j,i)=o_p^*(1)$ and $\frac{1}{2}\,\mathcal{T}(\mathcal{T}-1)\,h_{*q}^{q/2}\,b_*^{1/2}\,\psi_*(0)=h_{*q}^{q/2}\,h_{*q_A}^{-q_A}\,b_*^{-1/2}\mu_2+o_p^*(1)$. It is also straightforward to derive the bootstrap counterpart of Lemma 3 as well. The statement under the null thereby follows by noting that $\widehat{\mu}^*_{1,\mathcal{T}}=\widehat{\mu}_{1,\mathcal{T}}+o_p(h_{*q}^{-p/2}\,b_*^{-1/2})$, whereas it is immediate to see that $\widehat{\Lambda}^*_{\mathcal{T}}$ diverges at most at rate $O_p\left(\mathcal{T}\,h_{*q}^{q/2}\,b_*^{1/2}\right)$ under the alternative. ■

References

Aït-Sahalia, Y., 1996, Testing continuous-time models of the spot interest rate, Review of Financial Studies 9, 385–426.

Aït-Sahalia, Y., Bickel, P. J., Stoker, T. M., 2001, Goodness-of-fit tests for kernel regression with an application to option implied volatilities, Journal of Econometrics 105, 363412.

Aït-Sahalia, Y., Fan, J., Jiang, J., 2009a, Nonparametric tests of the markov hypothesis in continuous-time models, forthcoming in the Annals of Statistics.

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., Mykland, P., Zhang, L., 2009c, Ultra high frequency volatility estimation with dependent microstructure noise, forthcoming in the Journal of Econometrics.

Allen, F., Qian, J., Qian, M., 2007, China's financial system: Past, present, and future, in: L. Brandt T. Rawski (eds), China's Great Economic Transformation, Cambridge University Press, Cambridge.

Amaro de Matos, J., Fernandes, M., 2007, Testing the Markov property with high frequency data, Journal of Econometrics 141, 44–64.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2001, The distribution of realized exchange rate volatility, Journal of the American Statistical Association 96, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003, Modeling and forecasting realized volatility, Econometrica 71, 529–626.

Andersen, T. G., Dobrev, D., Schaumburg, E., 2009, Jump robust volatility estimation using nearest neighbor truncation, NBER Working Paper No. 15533.

Andrews, D. W., 1997, A conditional Kolmogorov test, Econometrica 65, 1097–1128.

Arcones, M. A., Giné, E., 1992, On the bootstrap of U and V statistics, Annals of Statistics 20, 655–674.

Bandi, F. M., Perron, B., 2006, Long memory and the relation between implied and realized volatility, Journal of Financial Econometrics 4, 636–670.

Bandi, F., Renò, R., 2008a, Nonparametric stochastic volatility, working paper, Johns Hopkins Carey Business School and Università di Siena.

Bandi, F., Renò, R., 2008b, Nonparametric leverage effects, working paper, Johns Hopkins Carey Business School and Università di Siena.

Barndorff-Nielsen, O. E., Hansen, P. H., Lunde, A., Shephard, N., 2008a, 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., Kinnebrock, S., Shephard, N., 2008b, Measuring downside risk: Realised semivariance, in: M. Watson T. Bollerslev J. Russell (eds), forthcoming in Volatility and Time Series Econometrics: Essays in Honor of Robert F. Engle, Oxford University Press, Oxford.

Barndorff-Nielsen, O. E., Shephard, N., 2002, Econometric analysis of realised volatility and its use in estimating stochastic volatility models, Journal of the Royal Statistical Society, series B 64, 243–280.

Barndorff-Nielsen, O. E., Shephard, N., 2004, Power and bipower variation with stochastic volatility and jumps (with discussion), Journal of Financial Econometrics 2, 1–48.

Bickel, P. J., Götze, F., van Zwet, W. R., 1997, Resampling fewer than n observations: Gains, losses, and remedies for losses, Statistica Sinica 7, 1–31.

Bickel, P. J., Rosenblatt, M., 1973, On some global measures of the deviations of density function estimates, Annals of Statistics 1, 1071–1095.

Bretagnolle, J., 1983, Lois limites du bootstrap de certaines fonctionnelles, Annales de l'Institut Henri Poincaré, Section B, Probabilités et Statistiques 19, 281–296.

Carr, P., Geman, H., Madan, D. B., Yor, M., 2005, A causality in variance test and its applications to financial market prices, Finance and Stochastics 9, 453–475.

Chen, X., Linton, O., Robinson, P. M., 2001, The estimation of conditional densities, in: M. L. Puri (ed.), Asymptotics in Statistics and Probability, VSP International Science Publishers, Leiden, pp. 71–84.

Cheung, Y. W., Ng, L. K., 1996, A causality in variance test and its applications to financial market prices, Journal of Econometrics 72, 33–48.

Christensen, B. J., Prabhala, N. R., 1998, The relation between implied and realized volatility, Journal of Financial Economics 50, 125–150.

Corradi, V., Distaso, W., Swanson, N. R., 2009, Predictive inference for integrated volatility, working paper, Warwick University, Imperial College, and Rutgers University.

Cox, J. C., Ingersoll, J. E., Ross, S. A., 1985, A theory of the term structure of interest rates, Econometrica 53, 385–407.

Diebold, F. X., Yilmaz, K., 2009, Measuring financial asset return and volatility spillovers, with application to global equity markets, Economic Journal 119, 158–171.

Diebold, F. X., Yilmaz, K., 2010, Equity market spillovers in the Americas, forthcoming, in: R. Alfaro D. Gray (eds), Financial Stability, Monetary Policy, and Central Banking, Bank of Chile, Santiago.

Engle, R. F., Ito, T., Lin, W.-L., 1990, Meteor showers or heat waves: Heteroskedastic intradaily volatility in the foreign exchange market, Econometrica 58, 525–542.

Engle, R. F., Ng, V., 1988, Measuring and testing the impact of news on volatility, Journal of Finance 48, 1749–1778.

Fan, J.-i., Masry, E., 1997, Local polynomial estimation of regression functions for mixing processes, Scandinavian Journal of Statistics 24, 165–179.

Fan, J.-q., Peng, L., Yao, Q.-W., Zhang, W.-Y., 2009, Approximating conditional density functions using dimension reduction, Acta Mathematicae Applicatae Sinica 25, 445–456.

Fan, J.-q., Yao, Q., Tong, H., 1996, Estimation of conditional densities and sensitivity measures in nonlinear dynamical systems, Biometrika 83, 189–206.

Fan, Y., 1995, Bootstrapping a consistent nonparametric goodness-of-fit test, Econometric Reviews 14, 367–382.

Fan, Y., Li, Q., 1999, Central limit theorem for degenerate U-statistics of absolutely regular processes with applications to model specification tests, Journal of Nonparametric Statistics 10, 245–271.

Fan, Y., Linton, O., 2003, Some higher-order theory for a consistent non-parametric model specification test, Journal of Statistical Planning and Inference 109, 125–154.

Feng, L., Seasholes, M. S., 2008, Individual investors and gender similarities in an emerging stock market, Pacific Basin Finance Journal 16, 44–60.

Gao, J., Hong, Y., 2008, Central limit theorems for generalized U-statistics with applications in nonparametric specification, Journal of Nonparametric Statistics 20, 61–73.

Gourieroux, C., Jasiak, J., 2007, Nonlinear causality, with applications to liquidity and stochastic volatility, working paper, CREST, University of Toronto, and York University.

Hall, P., Yao, Q.-W., 2005, Conditional distribution function approximation, and prediction, using dimension reduction, Annals of Statistics 33, 1404–1421.

Hamao, Y., Masulis, R. W., Ng, V., 1990, Correlations in price changes and volatility across international stock markets, Review of Financial Studies 3, 281–308.

Härdle, W., Mammen, E., 1993, Comparing nonparametric vs. parametric regression fits, Annals of Statistics 21, 1926–1947.

Hertz, E., 1998, The Trading Crowd: An Ethnography of the Shanghai Stock Market, Cambridge University Press, Cambridge.

Hong, Y., 2001, A test for volatility spillover with application to exchange rates, Journal of Econometrics 103, 183–224.

Hong, Y., White, H., 2004, Asymptotic distribution theory for nonparametric entropy measures of serial dependence, Econometrica 73, 837–901.

Horowitz, J. L., 2003, Bootstrap methods for Markov processes, Econometrica 71, 1049–1082.

Karolyi, G. A., 1995, A multivariate GARCH model of international transmissions of stock returns and volatility: The case of the United States and Canada, Journal of Business and Economic Statistics 13, 11–25.

King, M., Sentana, E., Wadhwani, S., 1994, Volatility and links between national stock markets, Econometrica 62, 901–933.

King, M., Wadhwani, S., 1990, Transmission of volatility between stock markets, Review of Financial Studies 3, 5–33.

Lin, W.-L., Engle, R. F., Ito, T., 1994, Do bulls and bears move across borders? International transmission of stock returns and volatility, Review of Financial Studies 7, 507–538.

Meddahi, N., 2003, ARMA representation of integrated and realized variances, Econometrics Journal 6, 334–355.

Mei, J., Scheinkman, J. A., Xiong, W., 2009, Speculative trading and stock prices: Evidence from Chinese A-B share premia, Annals of Economics and Finance.

Pantelidis, T., Pittis, N., 2004, Testing Granger causality in variance in the presence of causality in mean, Economic Letters 85, 201–207.

Paparoditis, E., Politis, D. N., 2002, The local bootstrap for Markov processes, Journal of Statistical Planning and Inference 108, 301–328.

Rajarshi, M. B., 1990, Bootstrap in Markov-sequences based on estimates of transition density, Annals of the Institute of Statistical Mathematics 42, 253–268.

Robinson, P. M., 1991, Consistent nonparametric entropy-based testing, Review of Economic Studies 58, 437–453.

Roll, R., 1989, Price volatility, international markets link and implications for regulatory policies, Journal of Financial Services Research 3, 211–246.

Schucany, W. R., Sommers, J. P., 1977, Improvement of kernel-type density estimators, Journal of the American Statistical Association 72, 420–423.

Sensier, M., van Dijk, D., 2004, Testing for volatility changes in U.S. macroeconomic time series, Review of Economics and Statistics 86, 833–839.

van Dijk, D., Osborne, D. R., Sensier, M., 2005, Testing for causality in variance in the presence of breaks, Economic Letters 89, 193–199.

Wongswan, J., 2006, Transmission of information across international equity markets, Review of Financial Studies 19, 1157–1189.

Zhang, L., 2006, Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach, Bernoulli 12, 1019–1043.

Zhang, L., Mikland, P., Aït-Sahalia, Y., 2005, A tale of two time scales: Determining integrated volatility with noisy high frequency data, Journal of the American Statistical Association 100, 1394–1411.

Figure 1
Realized measures of the daily variance of the index returns

The plots display the time series of the realized variance and tripower variation based on the 1-minute and 30-minute index returns as well as the realized kernel estimate of the daily variance based on 1-minute index returns. Intraday index returns refer to continuously compounded returns on the SSE B share index, Topix 100 index, S&P 500 index from January 3, 2000 to December 30, 2005.

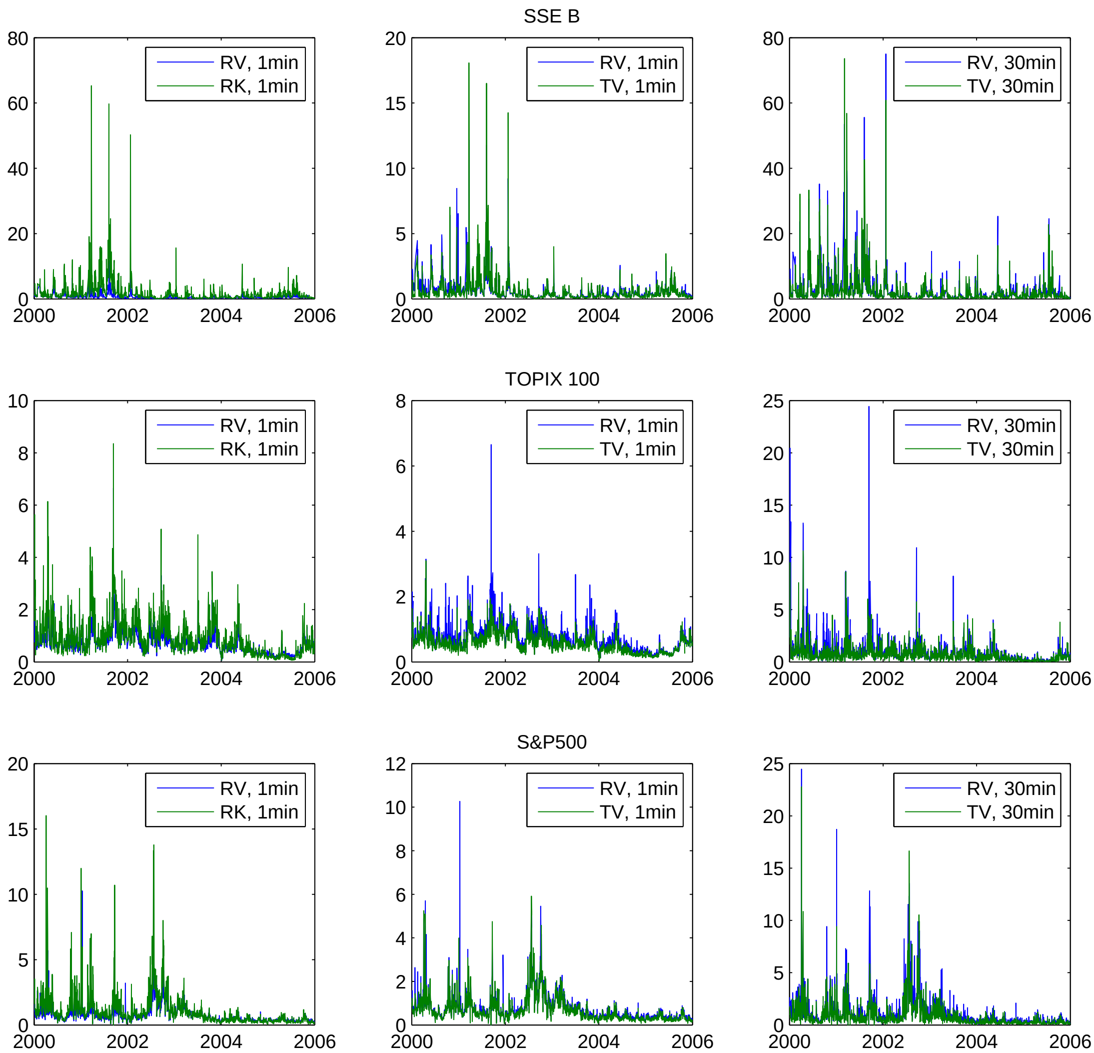

Figure 2
China spillovers to Japan

The plots compare the conditional density of the daily quadratic variation of the Topix 100 index given its past realization at the median value (dashed red line) with the conditional density not only given its past realization at the median value, but also given the past quadratic variation of the SSE B share index at (a) the first quartile, (b) the median, and (c) the third quartile values (solid black line). We measure quadratic variation using a realized kernel estimator based on 1-minute returns and then estimate the conditional densities by means of local-linear smoothing.

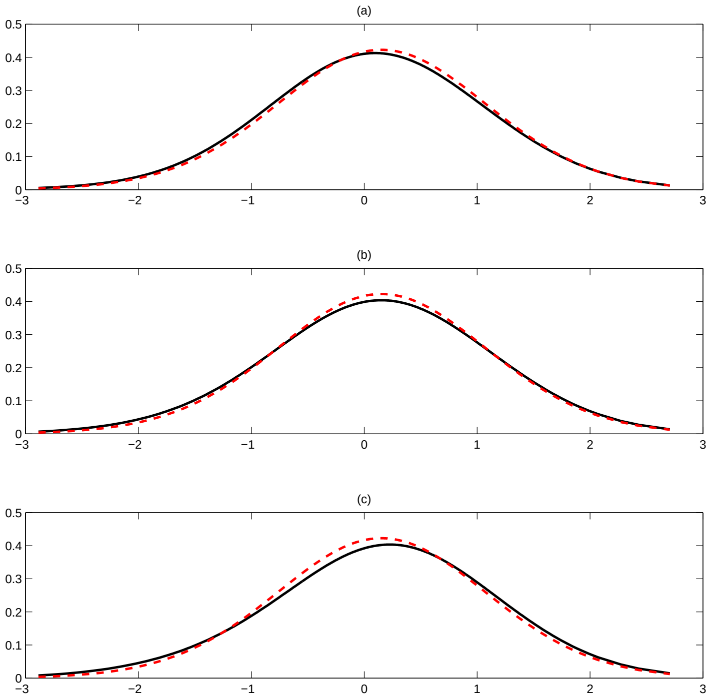

Figure 3
China spillovers to US

The plots compare the conditional density of the daily quadratic variation of the S&P 500 index given its past realization at the median value (dashed red line) with the conditional density not only given its past realization at the median value, but also given the past quadratic variation of the SSE B share index at (a) the first quartile, (b) the median, and (c) the third quartile values (solid black line). We measure quadratic variation using a realized variance estimator based on 15-minute returns and then estimate the conditional densities by means of local-linear smoothing.

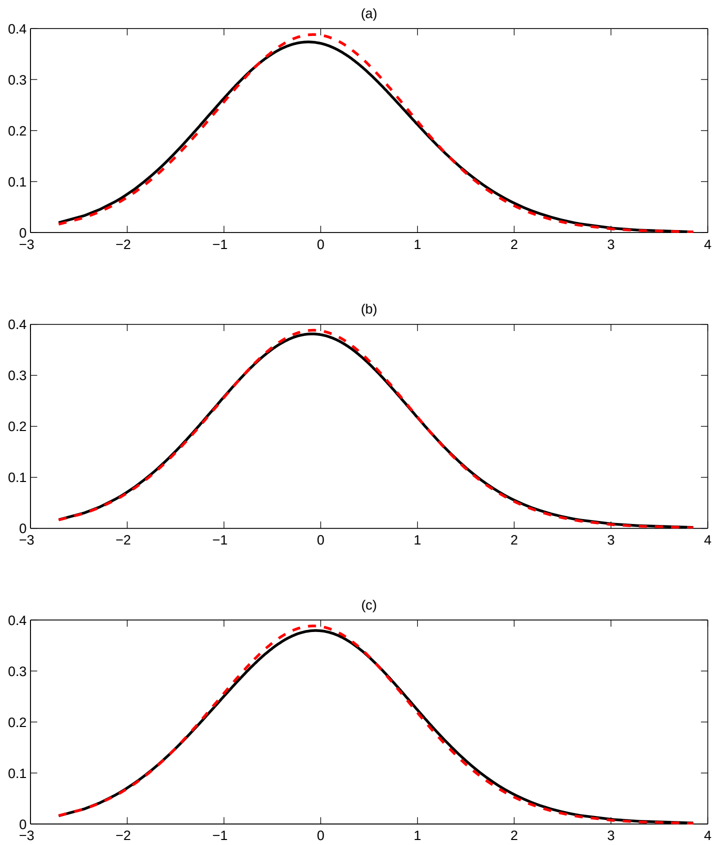

Table 1

Empirical size using bootstrap critical values

To examine empirical size, we simulate intraday returns from two independent CIR processes and then test for conditional independence in variance using bootstrap critical values. We consider tests at the 5% and 10% levels of significance for a sample size of $T = 200$ observations. We vary the number of intraday returns we employ to compute the realized measures in the first step as well as the scaling factors we use to adjust the rule-of-thumb bandwidths in the local linear estimation of the conditional density. In particular, we consider three sample sizes for the estimation of the integrated variance over a given day, $M \in \{144, 288, 576\}$ as well as all possible combinations between the scaling factors $\kappa_b \in \{1, 2.5, 5\}$ and $\kappa_h \in \{1, 2.5, 5\}$. All results rest on 500 Monte Carlo replications and 199 bootstrap artificial samples of 161 daily observations.

(κ_b, κ_h)	$M = 144$		$M = 288$		$M = 576$	
	5%	10%	5%	10%	5%	10%
$(1, 1)$	0.04	0.06	0.03	0.05	0.03	0.06
$(2.5, 1)$	0.01	0.03	0.02	0.04	0.02	0.03
$(5, 1)$	0.01	0.02	0.01	0.03	0.01	0.03
$(1, 2.5)$	0.01	0.02	0.02	0.03	0.02	0.03
$(2.5, 2.5)$	0.01	0.03	0.02	0.04	0.02	0.03
$(5, 2.5)$	0.02	0.04	0.02	0.04	0.02	0.05
$(1, 5)$	0.06	0.10	0.07	0.10	0.07	0.09
$(2.5, 5)$	0.12	0.17	0.13	0.19	0.13	0.18
$(5, 5)$	0.15	0.20	0.17	0.24	0.17	0.22

Table 2
Descriptive statistics for index returns

We collect transactions data for the SSE B share index, the Topix 100 index, and the S&P 500 index from Reuters, available at the Securities Industry Research Centre of Asia-Pacific. The sample spans the period ranging from January 3, 2000 to December 30, 2005. We document the main descriptive statistics for the index percentage returns with continuously compounding at regular sampling intervals of 1 and 30 minutes. The sample does not include overnight returns, so that the first intraday return refers to the opening price that ensues from the pre-sessional auction.

	S&P 500	Topix 100	SSE B share
sampling frequency: 1 minute			
mean	−0.0001	−0.0003	−0.0004
standard deviation	0.0448	0.0531	0.0525
minimum	−1.9020	−1.4095	−1.5255
maximum	1.5562	1.0710	2.3179
skewness	−0.1149	0.0429	0.9108
kurtosis	34.4874	21.4390	58.9127
zero returns	3.58%	5.57%	19.39%
sampling frequency: 30 minutes			
mean	−0.0014	−0.0084	−0.0098
standard deviation	0.2910	0.3170	0.5516
minimum	−3.3035	−4.8155	−6.4782
maximum	3.9838	3.5572	5.0368
skewness	0.0459	−0.2314	0.0450
kurtosis	13.1125	14.6988	15.3793
zero returns	1.72%	1.04%	2.51%

Table 3
Daily volatility transmission to the S&P 500 index

We report the outcome of the bootstrap test for conditional independence of the S&P 500 index daily variance with respect to the daily variances of the Topix 100 index and of the SSE B share index. We estimate daily variances using the following realized measures based on 1-minute and 5-minute returns: realized variance (RV), tripower variation (TV), two-scale realized variance (TS), and realized kernel (RK). In addition, we also compute the realized variance and tripower variation using 15-minute and 30-minute returns. We first standardize the logarithm of the data by their mean and standard deviation and then estimate the conditional densities by means of local linear smoothing with Gaussian-type kernels and scaling factors set to $\kappa_b = \kappa_h = 5/2$. As per the weighting function, we employ a standard multivariate normal density. To obtain critical values, we construct $B = 499$ bootstrap artificial samples of size $\mathcal{T} = \lfloor T^{.65} \rfloor$ by resampling blocks of $\lfloor \mathcal{T}^{1/4} \rfloor$ daily observations.

	RV	TV	TS	RK
Topix 100 index				
1 minute	0.038	0.020	0.050	0.008
5 minutes	0.160	0.224	0.252	0.124
15 minutes	0.010	0.116		
30 minutes	0.014	0.084		
Topix 100 index + VIX index				
1 minute	0.032	0.052	0.234	0.196
5 minutes	0.182	0.208	0.252	0.184
15 minutes	0.036	0.136		
30 minutes	0.046	0.098		
SSE B share index				
1 minute	0.480	0.116	0.028	0.026
5 minutes	0.842	0.250	0.034	0.190
15 minutes	0.010	0.408		
30 minutes	0.008	0.274		
SSE B share index + VIX index				
1 minute	0.546	0.578	0.250	0.216
5 minutes	0.258	0.236	0.094	0.210
15 minutes	0.030	0.104		
30 minutes	0.208	0.106		

Table 4

Daily volatility transmission to the VIX index

We report the outcome of the bootstrap test for conditional independence of the S&P 500 index daily variance with respect to the daily variances of the Topix 100 index and of the SSE B share index. We estimate daily variances using the following realized measures based on 1-minute and 5-minute returns: realized variance (RV), tripower variation (TV), two-scale realized variance (TS), and realized kernel (RK). In addition, we also compute the realized variance and tripower variation using 15-minute and 30-minute returns. We first standardize the logarithm of the data by their mean and standard deviation and then estimate the conditional densities by means of local linear smoothing with Gaussian-type kernels and scaling factors set to $\kappa_b = \kappa_h = 5/2$. As per the weighting function, we employ a standard multivariate normal density. To obtain critical values, we construct $B = 499$ bootstrap artificial samples of size $\mathcal{T} = \lfloor T^{.65} \rfloor$ by resampling blocks of $\lfloor \mathcal{T}^{1/4} \rfloor$ daily observations.

	RV	TV	TS	RK
Topix 100 index				
1 minute	0.002	0.004	0.042	0.052
5 minutes	0.068	0.056	0.368	0.156
15 minutes	0.100	0.182		
30 minutes	0.158	0.240		
Topix 100 index + S&P 500 index				
1 minute	0.020	0.016	0.948	0.380
5 minutes	0.950	0.268	0.202	0.156
15 minutes	0.070	0.428		
30 minutes	0.114	0.724		
SSE B share index				
1 minute	0.074	0.128	0.082	0.064
5 minutes	0.066	0.044	0.254	0.110
15 minutes	0.088	0.170		
30 minutes	0.146	0.098		
SSE B share index + S&P 500 index				
1 minute	0.864	0.792	0.780	0.832
5 minutes	0.788	0.610	0.072	0.728
15 minutes	0.050	0.268		
30 minutes	0.082	0.042		

Table 5
Daily volatility transmission to Japan and China

We report the outcome of the bootstrap test for conditional independence of the S&P 500 index daily variance with respect to the daily variances of the Topix 100 index (Panel A) and of the SSE B share index (Panel B). We estimate daily variances using the following realized measures based on 1-minute and 5-minute returns: realized variance (RV), tripower variation (TV), two-scale realized variance (TS), and realized kernel (RK). In addition, we also compute the realized variance and tripower variation using 15-minute and 30-minute returns. We first standardize the logarithm of the data by their mean and standard deviation and then estimate the conditional densities by means of local linear smoothing with Gaussian-type kernels and scaling factors set to $\kappa_b = \kappa_h = 5/2$. As per the weighting function, we employ a standard multivariate normal density. To obtain critical values, we construct $B = 499$ bootstrap artificial samples of size $\mathcal{T} = \lfloor T^{.65} \rfloor$ by resampling blocks of $\lfloor \mathcal{T}^{1/4} \rfloor$ daily observations.

	RV	TV	TS	RK
Panel A: Transmission to Topix 100 index				
S&P 500 index				
1 minute	0.004	0.016	0.108	0.032
5 minutes	0.046	0.124	0.074	0.052
15 minutes	0.016	0.144		
30 minutes	0.014	0.180		
SSE B share index				
1 minute	0.010	0.022	0.152	0.028
5 minutes	0.160	0.126	0.426	0.026
15 minutes	0.204	0.996		
30 minutes	0.282	0.076		
Panel B: Transmission to SSE B share index				
S&P 500 index				
1 minute	0.238	0.122	0.480	0.144
5 minutes	0.132	0.358	0.122	0.284
15 minutes	0.948	0.274		
30 minutes	0.908	0.774		
Topix 100 index				
1 minute	0.010	0.028	0.370	0.248
5 minutes	0.330	0.138	0.410	0.374
15 minutes	0.288	0.054		
30 minutes	0.756	0.828		

Table 6

Hourly volatility transmission

We report the outcome of the bootstrap test for conditional independence using the following hourly realized measures based on 1-minute returns: realized variance (RV), tripower variation (TV), two-scale realized variance (TS), and realized kernel (RK). We first standardize the logarithm of the data by their mean and standard deviation and then estimate the conditional densities by means of local linear smoothing with Gaussian-type kernels and scaling factors set to $\kappa_b = \kappa_h = 5/2$. As per the weighting function, we employ a standard multivariate normal density. To obtain critical values, we construct $B = 499$ bootstrap artificial samples of size $\mathcal{T} = \lfloor T^{.65} \rfloor$ by resampling blocks of $\lfloor \mathcal{T}^{1/4} \rfloor$ daily observations.

	RV	TV	TS	RK
Panel A: Transmission to the S&P 500 index				
Topix 100 index	0.070	0.092	0.098	0.078
SSE B share index	0.290	0.528	0.686	0.784
Panel B: Transmission to the Topix 100 index				
S&P 500 index	0.178	0.082	0.116	0.130
SSE B share index	0.114	0.082	0.046	0.132
Panel C: Transmission to the SSE B share index				
S&P 500 index	0.592	0.996	0.510	0.906
Topix 100 index	0.154	0.130	0.070	0.358