DEPARTMENT OF ECONOMICS
AND BUSINESS ECONOMICS
AARHUS UNIVERSITY

CREATES
Center for Research in Econometric Analysis of Time Series

Is the diurnal pattern sufficient to explain the intraday variation in volatility? A nonparametric assessment

Kim Christensen, Ulrich Hounyo and Mark Podolskij

CREATES Research Paper 2017-30

Department of Economics and Business Economics
Aarhus University
Fuglesangs Allé 4
DK-8210 Aarhus V
Denmark

Email: oekonomi@au.dk
Tel: +45 8716 5515

Is the diurnal pattern sufficient to explain the intraday variation in volatility? A nonparametric assessment*

Kim Christensen[†] Ulrich Hounyo[†] Mark Podolskij[‡,†]

July, 2017

Abstract

In this paper, we propose a nonparametric way to test the hypothesis that time-variation in intraday volatility is caused solely by a deterministic and recurrent diurnal pattern. We assume that noisy high-frequency data from a discretely sampled jump-diffusion process are available. The test is then based on asset returns, which are deflated by a model-free jump- and noise-robust estimate of the seasonal component and therefore homoscedastic under the null. The t-statistic (after pre-averaging and jump-truncation) diverges in the presence of stochastic volatility and has a standard normal distribution otherwise. We prove that replacing the true diurnal factor with our estimator does not affect the asymptotic theory. A Monte Carlo simulation also shows this substitution has no discernable impact in finite samples. The test is, however, distorted by small infinite-activity price jumps. To improve inference, we propose a new bootstrap approach, which leads to almost correctly sized tests of the null hypothesis. We apply the developed framework to a large cross-section of equity high-frequency data and find that the diurnal pattern accounts for a rather significant fraction of intraday variation in volatility, but important sources of heteroscedasticity remain present in the data.

JEL Classification: C10; C80.

Keywords: Bipower variation; bootstrapping; diurnal variation; high-frequency data; microstructure noise; pre-averaging; time-varying volatility.

*This paper was previously entitled "Testing for heteroscedasticity in jumpy and noisy high-frequency data: A resampling approach." We appreciate the thoughtful feedback received from two anonymous referees, the associate editor, and Yacine Aït-Sahalia as part of the revision process, which helped to significantly improve the paper. We also thank for comments on an earlier version of this paper made by conference participants at the 10th Computational and Financial Econometrics (CFE) conference in Seville, Spain; the Vienna-Copenhagen (VieCo) 2017 conference in Vienna, Austria; the 10th Annual SoFiE meeting in New York, USA; and at seminars in Aarhus, Aix-Marseille, Albany, California Polytechnic, Connecticut, Cornell, Erasmus (Rotterdam), Glasgow, Manchester and Nottingham. The authors acknowledge research funds from the Danish Council for Independent Research (DFF – 4182-00050). In addition, Podolskij thanks the Villum Foundation for the grant "Ambit fields: Probabilistic properties and statistical inference." This work was also supported by CREATES, which is funded by the Danish National Research Foundation (DNRF78). Please address correspondence to: uhounyo@econ.au.dk.

[†]Aarhus University, Department of Economics and Business Economics, CREATES, Fuglesangs Allé 4, 8210 Aarhus V, Denmark.

[‡]Aarhus University, Department of Mathematics, Ny Munkegade 118, 8000 Aarhus C, Denmark.

1 Introduction

There is a widespread agreement in the literature that any dynamic model of volatility should—at a minimum—account for two distinct features in order to explain the formation of diffusive risk in financial markets. On the one hand, a mean-reverting but highly persistent stochastic component is needed at the *interday* horizon to capture volatility clustering (e.g., Fama, 1965; Mandelbrot, 1963). On the other, a pervasive diurnal effect is required as one of the most critical determinants to describe the recurrent behavior of *intraday* volatility. In stock markets, for example, there is a tendency for absolute (or squared) price changes during the course of a trading day to form a so-called "U"- or reverse "J"-shape with notably larger fluctuations near the opening and closing of the exchange than around lunch time (see, e.g., Harris, 1986; Wood, McInish, and Ord, 1985, for early-stage documentation of this attribute). In addition to these effects, volatility may exhibit large, sudden shifts around the release of important economic news, such as macroeconomic information (e.g., Andersen and Bollerslev, 1998).

A recent strand of work, fueled by access to high-frequency data and complimentary theory for model-free measurement of volatility, has taken a more detailed close-up of these components and largely confirmed their presence.[1] The diurnal U-shape, in particular, has emerged as a potent—if not predominant—source of within-day variation in volatility. It is therefore common to formulate *parametric* models of time-varying volatility targeted for high-frequency analysis (be it in continuous- or discrete-time) as a composition of a stochastic and deterministic process (with suitable restrictions imposed to ensure the parameters are separately identified). A standard approach is to assume that the stochastic process is constant within a day but is evolving randomly between them (thus enabling volatility clustering), while the deterministic part is a smooth periodic function that is allowed to change within the day but is otherwise time-invariant (thus capturing the diurnal effect), see, e.g., Andersen and Bollerslev (1997, 1998); Boudt, Croux, and Laurent (2011); Engle and Sokalska (2012) and references therein.

Indeed, a major motivation behind the preferred use of realized measures of return variation that are temporally aggregated to the daily frequency is to avoid dealing with the diurnal effect, since it is widely believed to make them intrinsically robust against its presence. However, as stressed by Andersen, Dobrev, and Schaumburg (2012); Dette, Golosnoy, and Kellermann (2016) diurnal effects inject a strong Jensen's inequality-type bias in some of these estimators; an effect that is reinforced and magnified with a high "volatility-of-stochastic volatility" (e.g., Christensen, Oomen, and Podolskij, 2014). This can, for instance, alter the finite sample properties of jump tests designed to operate at *either* the intraday *or* interday horizon (e.g., Andersen, Bollerslev, and Dobrev, 2007; Barndorff-Nielsen and Shephard, 2006; Lee and Mykland, 2008) and make them significantly leaned

[1] A comprehensive list of papers in this field, including several reviews of the literature, is available at the webpage of the Oxford-Man Institute of Quantitative Finance's Realized Library: http://realized.oxford-man.ox.ac.uk/research/literature.

toward the alternative and cause spurious jump detection as result.[2] As such, further investigation of diurnal effects appears warranted.

In this paper, we develop a *nonparametric* framework to assess if diurnal effects can, in fact, explain all of the intraday variation in volatility, as stipulated by such a setup. A casual inspection of high-frequency data does not offer conclusive evidence about the validity of this conjecture. In concrete applications, inference is obscured by microstructure noise at the tick-by-tick frequency (e.g., Hansen and Lunde, 2006) and the existence of price jumps that are potentially very small and highly active (e.g., Aït-Sahalia and Jacod, 2012a).[3] Moreover, even if stochastic volatility is truly present, in practice its components may be so persistent that it is acceptable (and convenient) to regard it as absent on small time scales.[4]

Consistent with the above, we model the asset log-price as a general arbitrage-free Itô semimartingale, which is contaminated by microstructure noise. In our framework, the asymptotic theory is (mainly) infill, i.e. the process is assumed to be observed on a fixed time interval with mesh tending to zero.

There are several existing tests of constant volatility available in the high-frequency volatility area (see, e.g., Dette, Podolskij, and Vetter, 2006; Dette and Podolskij, 2008; Vetter and Dette, 2012). To our knowledge, none allow for the joint disturbance of jumps and microstructure noise, nor do they directly study the extension to diurnal variation advocated here. We formulate a test on the back of log-returns that are only homoscedastic under the null, after they are filtered for diurnal effects. We then study a jump-robust version of the pre-averaged bipower variation, where we extend the bivariate central limit theorem of Podolskij and Vetter (2009a) to the jumpy setting (see Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008; Jacod, Li, Mykland, Podolskij, and Vetter, 2009; Zhang, Mykland, and Aït-Sahalia, 2005; Zhang, 2006, for further work on noise-robust volatility estimation). The test is constructed via the asymptotic distribution implied by a transformation of such statistics and an application of Cauchy-Schwarz for a particular—but standard—choice of the parameters. As an aside, we add that a slightly different configuration of our t-statistic (based on the comparison of suitably non-truncated and truncated statistics) can serve as a basis for a jump test, which is robust to diurnal effects, but we do not pursue this idea in the present paper.

As the diurnal pattern is unknown in practice, we follow the "two-stage" approach of Andersen and Bollerslev (1997). In the first stage, the diurnal factors are pre-estimated. We propose a nonparametric estimator, which is both inherently jump- and noise-robust and therefore applicable

[2]The effects are deeply intertwined, however, because jumps can also induce substantial biases in and distort estimates of both integrated variance (e.g., Barndorff-Nielsen and Shephard, 2004; Christensen, Oomen, and Podolskij, 2014) and the diurnal pattern (e.g., Andersen, Bollerslev, and Das, 2001; Boudt, Croux, and Laurent, 2011).

[3]In view of this, a related topic is whether a diffusive component is needed in the first place to represent risk formation in financial asset prices. While the prevailing evidence is slightly mixed, it appears largely affirmative (see, e.g., Aït-Sahalia and Jacod, 2012b; Kolokolov and Renò, 2017; Todorov and Tauchen, 2010).

[4]Of course, the locally constant approximation of stochastic volatility is one of the most heavily exploited in the analysis of high-frequency data, see, e.g., Mykland and Zhang (2009).

in practice. It extends previous work of, e.g., Andersen and Bollerslev (1997); Andersen, Dobrev, and Schaumburg (2012); Boudt, Croux, and Laurent (2011); Taylor and Xu (1997); Todorov and Tauchen (2012) to the noisy setting (see also Hecq, Laurent, and Palm, 2012). In the second stage, the estimator is inserted and therefore replaces the true value in the deflation step. The estimate contains a sampling error, however, which afflicts the calculation and may propagate through the system and invalidate the analysis. This problem has largely gone unnoticed (or at least been ignored) in previous work.[5] Here, we show that our estimator has a sampling error of sufficiently small order to not affect the asymptotic theory.

A simulation study reveals that this substitution also has no discernable impact in finite samples. The test is, however, severely distorted by the presence of small infinite-activity price jumps, which it understandably appears to confuse with stochastic volatility. To improve inference, we suggest a new bootstrap approach. It is of independent interest and can be viewed as an overlapping version of the wild blocks of blocks bootstrap by Hounyo, Gonçalves, and Meddahi (2017). We prove the first-order validity of the bootstrap, while in simulations it helps to restore an almost correctly sized test.

The paper is structured as follows. In Section 2, we introduce our theoretical framework and the main assumptions. We extend the asymptotic theory of the pre-averaged bipower variation and construct a jump- and noise-robust test of the hypothesis that all intraday variation in volatility is captured by the diurnal pattern. In Section 3, we propose an estimator of the latent diurnal pattern and show that the sampling error of the feasible statistic is sufficiently small to not affect the previous results. In Section 4, we introduce the bootstrap and show its consistency for testing the null hypothesis in a noisy jump-diffusion setting. In Section 5, we present the Monte Carlo results, while an empirical illustration is conducted in Section 6. We conclude in Section 7. The mathematical proofs and some auxiliary results are relegated to the Appendix.

2 Theoretical setup

We let X denote a latent efficient log-price defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, P)$ and recorded in the window $[0, T]$, where T is the number of days in the sample and the subinterval $[t-1, t]$ is the tth day, for $t = 1, \dots T$. Throughout, T is mainly fixed and the asymptotic theory is infill, so we often impose $T = 1$ as a normalization, but please note the important digression in Section 3, where a long-span analysis ($T \to \infty$) is required, and there the additional notation and interpretation of $[0, T]$ is helpful.

As consistent with the no-arbitrage restriction (e.g., Delbaen and Schachermayer, 1994), we

[5] A notable exception is Todorov and Tauchen (2012). In the supplemental material to that paper, available at Econometrica's website, the authors treat the impact of diurnal filtering on their realized Laplace transform estimator of volatility, but microstructure noise is supposed to be absent.

model X as an Itô semimartingale:

$$X_t = X_0 + \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s + J_t, \quad t \in [0, T], \tag{1}$$

where $(a_t)_{t\geq 0}$ is a predictable, locally bounded drift process, $(\sigma_t)_{t\geq 0}$ is an adapted, càdlàg volatility process, while $(W_t)_{t\geq 0}$ is a Brownian motion.

J_t is a jump process defined by the equation:

$$J_t = \left(\delta 1_{\{|\delta|\leq 1\}}\right) \star \left(\underline{\mu}_t - \underline{\nu}_t\right) + \left(\delta 1_{\{|\delta|>1\}}\right) \star \underline{\mu}_t, \tag{2}$$

where $\underline{\mu}$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ and $\underline{\nu}$ is a predictable compensator of $\underline{\mu}$, such that $\underline{\nu}(\mathrm{d}s, \mathrm{d}x) = \mathrm{d}s \otimes \lambda(\mathrm{d}x)$ and λ is a σ-finite measure.

As explained above, the main idea of the paper is to construct a test, which tells whether a diurnal component is adequate to describe the evolution of within-day volatility.[6] To this end, we need to put some structure on the problem, starting with:

Assumption (D1): $\sigma_t = \sigma_{sv,t}\sigma_{u,t}$.

$\sigma_{sv,t}$ and $\sigma_{u,t}$ represent two distinct sources of time-varying volatility in many financial return series. The first term, $\sigma_{sv,t}$, denotes a stochastic process, which allows for randomness in the evolution of σ_t over time. The second term, $\sigma_{u,t}$, is a deterministic seasonal component that represents the diurnal pattern.

The multiplicative structure means our test can be formed by deflating the log-return series with $\sigma_{u,t}$ and checking if the outcome is homoscedastic, as we do in Section 2.2.

Assumption (D2): $\left(\sigma^2_{sv,t}\right)_{t\geq 0}$ is stationary with $E(\sigma^2_{sv,t}) = \sigma^2$ and $\sum_{k=0}^{\infty} \operatorname{cov}\left(\sigma^2_{sv,t}, \sigma^2_{sv,t+k}\right) < \infty$.

Assumption (D3): $(\sigma^2_{u,t})_{t\geq 0}$ is a continuously differentiable 1-periodic function with bounded derivative for $t \to 0$ and $t \to 1$ and normalized such that $\int_0^1 \sigma^2_{u,s}\mathrm{d}s = 1$.

Assumption (D2) – (D3) are sufficient to ensure identification of both volatility components from the data.[7] Apart from the stationarity of the stochastic volatility, the former also restricts its memory, which implies the process is ergodic. The latter says the diurnal component has to be recurrent, so that we can gradually infer it from gathering a larger sample.[8] Taken together, the conditions imply that an average of (an estimate of) volatility sampled at a fixed time of the day $s \in (0, 1)$, i.e. $T^{-1}\sum_{t=1}^{T} \sigma^2_{t-1+s} = \sigma^2_{u,s} T^{-1} \sum_{t=1}^{T} \sigma^2_{sv,t-1+s} \xrightarrow{p} \sigma^2_{u,s}\sigma^2$, as $T \to \infty$, thereby delivering $\sigma^2_{u,s}$ after suitable normalization.

[6] We do not speak about the dynamics of volatility during market close. Thus, our framework is consistent with random changes in between-day volatility.

[7] The normalization in Assumption (D3) is known as a "standardization condition," which ensures that the decomposition in Assumption (D1) is unique under (locally) constant volatility, see, e.g., Andersen and Bollerslev (1997); Boudt, Croux, and Laurent (2011); Taylor and Xu (1997).

[8] It can be extended to accommodate a day-of-the-week effect (e.g., Andersen and Bollerslev, 1998).

In Section 3, we use this idea to recover σ_u from noisy high-frequency data, so that we can compute the test in practice, but—for the moment—we treat it as observed.

We further rule out jumps in $\sigma_{sv,t}$:

Assumption (V): $\sigma_{sv,t}$ is of the form:

$$\sigma_{sv,t} = \sigma_0 + \int_0^t \tilde{a}_s \mathrm{d}s + \int_0^t \tilde{\sigma}_s \mathrm{d}W_s + \int_0^t \tilde{v}_s \mathrm{d}B_s, \tag{3}$$

where $(\tilde{a}_t)_{t\geq 0}$, $(\tilde{\sigma}_t)_{t\geq 0}$ and $(\tilde{v}_t)_{t\geq 0}$ are adapted, càdlàg stochastic processes, while $(B_t)_{t\geq 0}$ is a standard Brownian motion that is independent of W.

Assumption (V) is common in the realized volatility literature (see, e.g., Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008; Gonçalves and Meddahi, 2009; Mykland and Zhang, 2009; Christensen, Podolskij, and Vetter, 2013; Hounyo, 2017). It facilitates the control of some approximation errors in the proofs, but it can potentially be relaxed. In recent work, Christensen, Podolskij, Thamrongrat, and Veliyev (2016) operate with a power variation-based statistic and impose a weaker set of assumptions, which allows for rather unrestricted jump dynamics in volatility (see their equation (2.3) and Theorem (3.2), which is based on Assumption (H1) from Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006)). It may be possible to extend our setting in that direction, but we leave a full exploration of it for future research.

In some of our results, we also assume that the volatility is bounded away from zero. In particular, we sometimes adopt the following condition:

Assumption (V'): $\sigma_{sv,t} > 0$ and $\sigma_{u,t} > 0$, for all $t \geq 0$.

At last, we impose that:

Assumption (J): There exists a sequence of stopping times $(\widetilde{\tau}_n)_{n=1}^{\infty}$ increasing to ∞ and a deterministic nonnegative function $\tilde{\gamma}_n$ such that $\int_{\mathbb{R}} \tilde{\gamma}_n(x)^\beta \lambda(\mathrm{d}x) < \infty$ and $||\delta(\omega, t, x)|| \wedge 1 \leq \tilde{\gamma}_n(x)$, for all (ω, t, x) with $t \leq \tilde{\tau}_n(\omega)$, where $\beta \in [0, 2]$.

β captures the activity of the jump process. As β approaches two, the jumps are smaller but more vibrant. As explained by Todorov and Bollerslev (2010), the harder they are to distinguish from the diffusive part of X. Below, we impose Assumption (J) to hold for any $\beta \in [0, 1)$, thus restricting attention to jump processes with sample paths of finite length.

2.1 Microstructure noise

The presence of market frictions (such as price discreteness, rounding errors, bid-ask spreads, gradual response of prices to block trades and so forth) prevent us from observing the true, efficient

log-price process X_t. Instead, we observe a noisy version Y_t, which we assume is given by

$$Y_t = X_t + \epsilon_t, \tag{4}$$

where ϵ_t is a noise term that collects the market microstructure effects. We assume that ϵ_t is independently distributed and independent of X_t, such that

$$E(\epsilon_t) = 0 \quad \text{and} \quad E\big(\epsilon_t^2\big) = \sigma_{u,t}^2 \omega^2, \tag{5}$$

for any t, where Y_t is observed.

As consistent with, e.g, Bandi and Russell (2006); Kalnina and Linton (2008), the second moment of the noise is allowed to be heteroscedastic and exhibit diurnal variation. We assume it is identical to the volatility diurnality, which conveniently makes the detrended noise asymptotically i.i.d (cf. (12)). We return to this later in Remark 3, where we highlight the impact of weakening it to a general form of heteroscedasticity.

About the noise distribution, we follow Podolskij and Vetter (2009a):

Assumption (A): (i) ϵ is distributed symmetrically around zero, and (ii) for any $0 > a > -1$, it holds that $E(|\epsilon_t|^a) < \infty$.

Assumption (A'): Cramer's condition is fulfilled, that is $\limsup_{t\to\infty} |\chi(t)| < 1$, where χ denotes the characteristic function of ϵ.

2.2 Test of heteroscedasticity

To develop a test of the "no heteroscedasticity after diurnal correction" assumption, we partition the sample space Ω into the following two subsets:

$$\Omega_{\mathcal{H}_0} = \{\omega : \sigma_{sv,t} \text{ is constant for } t \geq 0\}, \tag{6}$$

and $\Omega_{\mathcal{H}_a} = \Omega_{\mathcal{H}_0}^{\complement}$. The null hypothesis can then formally be defined as $\mathcal{H}_0 : \omega \in \Omega_{\mathcal{H}_0}$, whereas the alternative is $\mathcal{H}_a : \omega \in \Omega_{\mathcal{H}_a}$.

Our goal is to find a test with a prescribed asymptotic significance level and with power going to one to test the hypothesis that $\omega \in \Omega_{\mathcal{H}_0}$. The key challenge we address is how to construct such a test, when X—apart from being driven by a Brownian component—is subject to diurnal variation, potentially discontinuous and observed with measurement error. The solution is based on computing a set of estimators, which reveal information about the presence of time-variation in the stochastic volatility $\sigma_{sv,t}$ robustly to the above features.

The differential form of (1) scaled by $\sigma_{u,t}$ yields:

$$\frac{\mathrm{d}X_t}{\sigma_{u,t}} = \frac{a_t}{\sigma_{u,t}}\mathrm{d}t + \frac{\sigma_t}{\sigma_{u,t}}\mathrm{d}W_t + \frac{\mathrm{d}J_t}{\sigma_{u,t}}, \tag{7}$$

or

$$\mathrm{d}X_t^d = a_t^d \mathrm{d}t + \sigma_{sv,t}\mathrm{d}W_t + \mathrm{d}J_t^d, \tag{8}$$

where a superscript d is used to represent a process that has been adjusted by the seasonal component of volatility.[9]

Then, we study the quadratic variation of X^d:

$$[X^d]_t = \int_0^t \sigma^2_{sv,s} ds + \sum_{s \leq t} |\Delta X^d_s|^2, \tag{9}$$

where $\int_0^t \sigma^2_{sv,s} \mathrm{d}s$ is the integrated variance of X^d, while $\sum_{s \leq t} |\Delta X^d_s|^2$ is the sum of the squared deflated jumps, where $\Delta X^d_s = X^d_s - X^d_{s-}$.

We note that if the stochastic volatility process is constant, say $\sigma_{sv,t} = \sigma$, (1) reduces to

$$X_t = X_0 + \int_0^t a_s \mathrm{d}s + \sigma \sigma_{u,t} \big(W_t - W_0\big) + \big(\delta 1_{\{|\delta| \leq 1\}}\big) \star \Big(\underline{\mu}_t - \underline{\nu}_t\Big) + \big(\delta 1_{\{|\delta| > 1\}}\big) \star \underline{\mu}_t, \tag{10}$$

while

$$[X^d]_t = \sigma^2 t + \sum_{s \leq t} |\Delta X^d_s|^2. \tag{11}$$

The construction of the t-statistic now progresses in three steps. Firstly, we account for microstructure noise by doing local pre-averaging of Y^d. Secondly, we tease out the continuous part of the quadratic variation by suitably removing the jump component in (11). Thirdly, we develop a fully feasible theory by proposing a statistic that can replace σ_u in the computations.

2.3 The pre-averaging approach

In this section, we confine the clock to $t \in [0, 1]$, i.e. we set $T = 1$. In our simulations and empirical work, we implement the test "day-by-day," so that here the unit interval is naturally interpreted as a trading day's worth of data.

The noisy log-price Y_t is observed at regular time points $t_i = i/n$, for $i = 0, \ldots, n$. Then, the deflated intraday log-returns (at frequency n) can be computed as:

$$\Delta^n_i Y^d \equiv Y^d_{i/n} - Y^d_{(i-1)/n}, \quad i = 1, \ldots, n. \tag{12}$$

As $Y^d_t = X^d_t + \epsilon^d_t$, we can split $\Delta^n_i Y^d$ into

$$\Delta^n_i Y^d = \Delta^n_i X^d + \Delta^n_i \epsilon^d, \tag{13}$$

where $\Delta^n_i X^d = X^d_{i/n} - X^d_{(i-1)/n}$ denotes the n-frequency return of the efficient log-price, while $\Delta^n_i \epsilon^d = \epsilon^d_{i/n} - \epsilon^d_{(i-1)/n}$ is the change in the microstructure component.

To lessen the noise, we adopt the pre-averaging approach of Jacod, Li, Mykland, Podolskij, and Vetter (2009); Podolskij and Vetter (2009a,b). To describe it, we let k_n be a sequence of positive integers and g a real-valued function. k_n represents the length of a pre-averaging window, while g assigns a weight to those noisy log-returns that are inside it. g is defined on $[0, 1]$, such that $g(0) =$

[9]Note that many parts of this paper can be applied to both the raw and deflated log-returns series (e.g., the pre-averaging theory in the next subsection). We base the theoretical exposition on the seasonally adjusted version to minimize the notational load, while we present results for both series in the empirical application.

$g(1) = 0$ and $\int_0^1 g(s)^2 \mathrm{d}s > 0$. We assume g is continuous and piecewise continuously differentiable with a piecewise Lipschitz derivative g'. A canonical function that fulfills these restrictions is $g(x) = \min(x, 1-x)$.

We introduce the notation:

$$\phi_1(s) = \int_s^1 g'(u)g'(u-s)\mathrm{d}u \quad \text{and} \quad \phi_2(s) = \int_s^1 g(u)g(u-s)\mathrm{d}u, \tag{14}$$

and for $i = 1, 2$, we let $\psi_i = \phi_i(0)$. For instance, if $g(x) = \min(x, 1-x)$, it follows that $\psi_1 = 1$ and $\psi_2 = 1/12$.

Also, we write:

$$\psi_1^n = k_n \sum_{j=1}^{k_n} \left(g\left(\frac{j}{k_n}\right) - g\left(\frac{j-1}{k_n}\right)\right)^2 \quad \text{and} \quad \psi_2^n = \frac{1}{k_n} \sum_{j=1}^{k_n-1} g^2\left(\frac{j}{k_n}\right). \tag{15}$$

In the appendix, after freezing the volatility locally, ψ_1^n and ψ_2^n appear in the conditional expectation of the squared pre-averaged return in (17). As $n \to \infty$,

$$\psi_1^n \to \psi_1 \quad \text{and} \quad \psi_2^n \to \psi_2, \tag{16}$$

while $\psi_i^n - \psi_i = O\big(n^{-1/2}\big)$, for $i = 1, 2$, so we can work with ψ_i and not worry about the effect of this substitution in the asymptotic theory. In contrast, ψ_i^n can differ a lot from ψ_i, if k_n is small, so as a practical guide it is better to work with (15).

The pre-averaged return, say $\Delta_i^n \bar{Y}^d$, is then found by computing a weighted sum of consecutive n-frequency deflated log-returns over a block of size k_n:

$$\Delta_i^n \bar{Y}^d = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \Delta_{i+j-1}^n Y^d, \quad i = 1, \ldots, n - k_n + 2. \tag{17}$$

As readily seen, pre-averaging entails a slight "loss" of summands compared to n. Thus, while the original sample size is n, there are only $n - k_n + 2$ elements in $(\Delta_i^n \bar{Y}^d)_{i=1}^{n-k_n+2}$. It follows from the decomposition in (13) that $\Delta_i^n \bar{Y}^d = \Delta_i^n \bar{X}^d + \Delta_i^n \bar{\epsilon}^d$ and, as shown by Vetter (2008),

$$\Delta_i^n \bar{X}^d = O_p\left(\sqrt{\frac{k_n}{n}}\right) \quad \text{and} \quad \Delta_i^n \bar{\epsilon}^d = O_p\left(\frac{1}{\sqrt{k_n}}\right). \tag{18}$$

Thus, the noise is dampened, thereby reducing its influence on $\Delta_i^n \bar{Y}^d$. As an outcome, we retrieve a basically noise-free estimate, which can substitute the efficient log-return $\Delta_i^n X^d$ in subsequent computations, taking proper account of the dependence introduced in $(\Delta_i^n \bar{Y}^d)_{i=1}^{n-k_n+2}$.[10] The reduction increases with larger k_n, but too much pre-averaging also impedes the accuracy of estimators of the quadratic variation, yielding a trade-off in selecting k_n. To strike a balance and get an efficient

[10] If k_n is even, it follows with the above definition of $g(x) = \min(x, 1-x)$ that the pre-averaged returns in (17) can be rewritten as $\Delta_i^n \bar{Y}^d = \frac{1}{k_n} \sum_{j=1}^{k_n/2} Y^d_{\frac{i+k_n/2+j}{n}} - \frac{1}{k_n} \sum_{j=1}^{k_n/2} Y^d_{\frac{i+j}{n}}$. Thus, the sequence $(2\Delta_i^n \bar{Y}^d)_{i=1}^{n-k_n+2}$ can be interpreted as constituting a new set of increments from a price process that is constructed by simple averaging of the rescaled noisy log-price series, $(Y^d_{i/n})_{i=0}^n$, in a neighbourhood of i/n, thus making the use of the term pre-averaging and the associated notation transparent.

$n^{-1/4}$ rate of convergence, Jacod, Li, Mykland, Podolskij, and Vetter (2009) propose to set:

$$k_n = \theta\sqrt{n} + o\big(n^{-1/4}\big), \tag{19}$$

for some $\theta \in (0, \infty)$. With this choice, the orders of $\Delta_i^n \bar{X}^d$ and $\Delta_i^n \bar{\epsilon}^d$ are balanced and equal to $O_p\big(n^{-1/4}\big)$. An example of (19) used throughout this paper is $k_n = \big[\theta\sqrt{n}\big]$.

2.3.1 The pre-averaged bipower variation

With the pre-averaged return series, $(\Delta_i^n \bar{Y}^d)_{i=1}^{n-k_n+2}$, available, Podolskij and Vetter (2009a) propose the bipower variation statistic:

$$BV(Y^d, l, r)^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{N_n} y(Y^d, l, r)_i^n, \tag{20}$$

where $l, r \geq 0$, $y(Y^d, l, r)_i^n = |\Delta_i^n \bar{Y}^d|^l |\Delta_{i+k_n}^n \bar{Y}^d|^r$, $N_n = n - 2k_n + 2$ and $\mu_p = E(|N(0,1)|^p)$.[11] In the following, if we write $BV(l, r)^n$ and $y(l, r)_i^n$, we assume that they are implicitly defined with respect to Y^d. Podolskij and Vetter (2009a) show that under suitable regularity conditions, in particular that X is a continuous Itô semimartingale (i.e., X follows (103)), then as $n \to \infty$

$$BV(l, r)^n \xrightarrow{p} BV(l, r) = \int_0^1 \left(\theta\psi_2 \sigma_{sv,s}^2 + \frac{1}{\theta}\psi_1 \omega^2\right)^{\frac{l+r}{2}} \mathrm{d}s, \tag{21}$$

and

$$n^{1/4} \begin{pmatrix} BV(l_1, r_1)^n - BV(l_1, r_1) \\ BV(l_2, r_2)^n - BV(l_2, r_2) \end{pmatrix} \xrightarrow{d_s} MN\big(0, \Sigma\big), \tag{22}$$

with $l_1, r_1, l_2, r_2 \geq 0$, where "$\xrightarrow{d_s}$" is stable convergence, $\Sigma = \big(\Sigma_{ij}^{l_1,r_1,l_2,r_2}\big)_{1 \leq i,j \leq 2}$ the conditional covariance matrix of the limiting process $n^{1/4}\big(BV(l_1, r_1)^n, BV(l_2, r_2)^n\big)^\intercal$, and $^\intercal$ the transpose.[12]

2.3.2 A truncated pre-averaged bipower variation

The estimator in (20) can also be made jump-robust in both the stochastic limit and its asymptotic distribution, but—as explained by Podolskij and Vetter (2009a)—this puts strong restrictions on l and r. Firstly, the central limit theory in (22) is not valid for the popular choice $l = r = 1$. Indeed, Vetter (2010) shows that this estimator is not even mixed Gaussian, which severely constrains our ability to draw inference. Secondly, the version with $l = r = 2$ as implemented below, does not converge to the limit in (21), if X jumps, and while that is true for the pre-averaged (1,1)-bipower variation, asymptotically, it is well-known that the latter typically has a pronounced upward bias in finite samples (e.g., Christensen, Oomen, and Podolskij, 2014). Thus, to achieve a better jump-robustness and enlarge the feasible set of powers for which we can do hypothesis testing, we follow

[11] In order to avoid a finite sample bias in the construction of $BV(l, r)^n$, we only divide it by N_n (the number of summands in the estimator) in our simulations and empirical work. We stick with n in the theoretical parts of the paper, as it involves less notation.

[12] The formal definition of Σ is given in Appendix A.

Corsi, Pirino, and Renò (2010) in the no-noise and finite-activity jump setting by combining the bipower idea with the truncation approach of Mancini (2009); Jacod and Protter (2012); Jing, Liu, and Kong (2014).

To introduce our t-statistic for the homoscedasticity test, we therefore start by deriving a result as above for a truncated pre-averaged bipower variation, which verifies that the probability limit and asymptotic distribution of this new estimator are identical to those given by (21) and (22) in the general setting, where X follows the Itô semimartingale in (1). Thus, we propose to set:

$$\check{BV}(l,r)^n = n^{\frac{l+r}{4}-1}\frac{1}{\mu_l\mu_r}\sum_{i=1}^{N_n}\check{y}(l,r)_i^n, \tag{23}$$

where $\check{y}(l,r)_i^n = |\Delta_i^n\bar{Y}^d|^l 1_{\left\{|\Delta_i^n\bar{Y}^d|<\upsilon_n\right\}}|\Delta_{i+k_n}^n\bar{Y}^d|^r 1_{\left\{|\Delta_{i+k_n}^n\bar{Y}^d|<\upsilon_n\right\}}$ and $1_{\{\cdot\}}$ is the indicator function, which discards pre-averaged log-returns that exceed a predetermined level

$$\upsilon_n = \alpha u_n^{\varpi}, \text{ for } \alpha > 0 \text{ and } \varpi \in (0,1/2), \tag{24}$$

such that $u_n = k_n/n$.

Theorem 2.1 *Let l_1, r_1, l_2 and r_2 be four positive real numbers and X be given by* (1)*. Suppose that Assumption (J) holds for some $\beta \in [0, \min\{1, l_1, r_1, l_2, r_2\})$ and that $\left(\frac{l_1+r_1-1}{2(l_1+r_1-\beta)} \vee \frac{l_2+r_2-1}{2(l_2+r_2-\beta)}\right) \leq \varpi < 1/2$. Furthermore, we assume (D1), (V), (A), and impose the moment condition $E(|\epsilon_t|^s) < \infty$, for some $s > (3 \vee 2(r_1+l_1) \vee 2(r_2+l_2))$. If any l_i or r_i is in $(0,1]$, we postulate (V′), otherwise either (V′) or (A′). In addition, suppose that $k_n \to \infty$ as $n \to \infty$ such that* (19) *holds. Then, as $n \to \infty$,*

$$n^{1/4}\begin{pmatrix} \check{BV}(l_1,r_1)^n - BV(l_1,r_1) \\ \check{BV}(l_2,r_2)^n - BV(l_2,r_2) \end{pmatrix} \xrightarrow{d_s} MN\big(0,\Sigma\big). \tag{25}$$

Theorem 2.1 shows that (23) is robust to the jump part in its limiting distribution. Note that Σ is identical to the matrix in (22). To our knowledge, the result is new with the main innovations being the statistic is (23) and the underlying process is a general Itô semimartingale given by (1). It extends Theorem 3 of Podolskij and Vetter (2009a) to discontinuous X by establishing a joint asymptotic distribution, as in (22), for the class of truncated pre-averaged bipower variation. In previous work, Jing, Liu, and Kong (2014) prove—under some regularity conditions—the consistency and CLT for the truncated pre-averaged realized variance, i.e. the statistic of the form $\check{BV}(2,0)^n$, when X follows (1). Our paper generalizes the latter article to the bipower setting with—subject to the above constraint—arbitrary powers.

The lower bound on ϖ is determined by an interplay between the bipower parameters and the activity of the jump process. The crude intuition is that small jumps tend to resemble Brownian motion, so if the threshold vanishes too slowly, it can impair the jump-robustness, and this effect is aggravated for larger bipowers. We therefore normally work with ϖ close to a half in practice. In our Monte Carlo, $l_1 = r_2 = 2$, $l_2 = r_2 = 1$ and $\beta = 0.5$, so that any $\varpi \in [1/3, 1/2)$ is valid. Throughout, we always set $\varpi = 0.49$.

The above enables extraction of an essentially noise-free and jump-robust estimate of the continuous piece of the quadratic variation in (9) and thus facilitates the construction of a test for the presence of time-variation in $\sigma_{sv,t}$. An implication of (25) is that for any $l_1, r_1, l_2, r_2 \geq 0$, which adhere to the conditions of Theorem 2.1 and such that $l_1 + r_1 > l_2 + r_2$, as $n \to \infty$,

$$\begin{aligned} &\check{BV}(l_1, r_1)^n - (\check{BV}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}} \xrightarrow{p} BV(l_1, r_1) - (BV(l_2, r_2))^{\frac{l_1+r_1}{l_2+r_2}} \\ &= \int_0^1 \left(\theta\psi_2\sigma_{sv,s}^2 + \frac{1}{\theta}\psi_1\omega^2\right)^{\frac{l_1+r_1}{2}} \mathrm{d}s - \left[\int_0^1 \left(\theta\psi_2\sigma_{sv,s}^2 + \frac{1}{\theta}\psi_1\omega^2\right)^{\frac{l_2+r_2}{2}} \mathrm{d}s\right]^{\frac{l_1+r_1}{l_2+r_2}} \geq 0, \end{aligned} \tag{26}$$

with equality if and only if $\sigma_{sv,t}$ is constant. We thus build a test of $\mathcal{H}_0$ via the infeasible t-statistic:

$$T_{\text{inf.}}^n = \frac{n^{1/4}\Big(\check{BV}(l_1, r_1)^n - (\check{BV}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}}\Big)}{\sqrt{V}} \xrightarrow{d} N(0, 1), \tag{27}$$

where

$$V = \Sigma_{11} - 2\left(\frac{l_1 + r_1}{l_2 + r_2}\right)\big(\check{BV}(l_2, r_2)^n\big)^{\frac{l_1+r_1}{l_2+r_2}-1}\Sigma_{12} + \left(\frac{l_1 + r_1}{l_2 + r_2}\right)^2 (\check{BV}(l_2, r_2)^n)^{2\left(\frac{l_1+r_1}{l_2+r_2}-1\right)}\Sigma_{22}. \tag{28}$$

Note that the convergence in (27) holds only under $\mathcal{H}_0$, while under $\mathcal{H}_a$ it follows from (26) that $n^{1/4}\Big(\check{BV}(l_1, r_1)^n - (\check{BV}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}}\Big) \to \infty$. This way we can determine if X^d has homoscedastic or heteroscedastic volatility with asymptotically correct size and power tending to one, as $n \to \infty$. To render the test feasible, we propose a consistent estimator of Σ in Section 4, which can be plugged into (28). It is both inherently robust to heteroscedasticity and positive semi-definite.

3 A local estimator of diurnal variance

In the previous section, we pretended the diurnal component of volatility was available to deflate the noisy log-return series (i.e., (7)). In practice, σ_u is unobserved. We here propose a nonparametric jump- and noise-robust estimator of it and state appropriate conditions, under which the sampling error—induced by this estimation—is asymptotically negligible, so that it does not thwart the results in Section 2 (and 4).

It turns out to be impossible to recover the latent diurnal variance on a fixed time interval. We thus resort to a long-span asymptotic theory, which extracts information about it by pooling high-frequency data across days.

As above, we suppose that on day t we record Y at equidistant time points $t_i = t - 1 + i/n$, for $i = 0, 1, \ldots, n$ and write the associated n-frequency log-returns as:

$$\Delta_{(t-1)n+i}^n Y \equiv Y_{t-1+i/n} - Y_{t-1+(i-1)/n}, \quad \text{for } t = 1, \ldots, T \text{ and } i = 1, \ldots, n. \tag{29}$$

As in Zhang, Mykland, and Aït-Sahalia (2005), we operate within a two time scale framework, where the "slow" scale uses a coarser set of m-frequency returns, where $m < n$, i.e. $\Delta_{(t-1)m+j}^m Y = Y_{t-1+j/m} - Y_{t-1+(j-1)/m}$, for $t = 1, \ldots, T$ and $j = 1, \ldots, m$, which is reserved for diurnal variance

estimation, while the "fast" scale is based on all observed n-frequency returns and is intended for a bias-correction. Throughout, we assume m is a divisor of n, so that $\{j/m\}_{j=0}^{m} \subseteq \{i/n\}_{i=0}^{n}$.

In the following, we say a process $(b_t)_{t\geq 0}$ is bounded in L^p, if

$$\sup_{t\in\mathbb{R}_+} E\big[|b_t|^p\big] < \infty. \tag{30}$$

Assumption (D4): $(a_t)_{t\geq 0}$, $(\tilde{a}_t)_{t\geq 0}$, $(\tilde{\sigma}_t)_{t\geq 0}$ and $(\tilde{v}_t)_{t\geq 0}$ are bounded in L^4.

Assumption (D4) adds some regularity to the driving processes in X, which is necessary here as $T \to \infty$ in the asymptotic theory, and so we cannot appeal to the standard "localization" procedure (e.g., Jacod, 2008) to bound various terms in the proofs.

Now, we set:

$$\hat{\sigma}^2_{u,s} = \frac{1}{T}\sum_{t=1}^{T}(\sqrt{m}\Delta^m_{(t-1)m+j}Y)^2 - \frac{m}{T}\sum_{t=1}^{T}[\hat{\text{var}}(\epsilon_{t-1+(j-1)/m}) + \hat{\text{var}}(\epsilon_{t-1+j/m})], \quad \text{for } s \in [(j-1)/m, j/m), \tag{31}$$

where $\hat{\text{var}}(\epsilon_{t-1+(j-1)/m})$ is a consistent estimator of $\text{var}(\epsilon_{t-1+(j-1)/m})$, which has to converge at a rate faster than m^{-1}, e.g.

$$\hat{\text{var}}(\epsilon_{t-1+(j-1)/m}) = -\frac{1}{T}\frac{1}{n/m-1}\sum_{t=1}^{T}\sum_{i=1}^{n/m}\left[\left(\Delta^n_{i+\left[t-1+\frac{j-1}{m}\right]_n}Y\right)\left(\Delta^n_{i-1+\left[t-1+\frac{j-1}{m}\right]_n}Y\right)\right]. \tag{32}$$

As readily seen, $\hat{\sigma}^2_{u,s}$ is based directly on the raw noisy high-frequency data. It does not require jump-truncation nor pre-averaging and is therefore trivial to compute.[13] Due to its reliance on the squared normalized noisy high-frequency increment, however, it accumulate a bias from the microstructure noise, which the second term in (31) cancels out by computing a local block-wise estimator of the noise variance.

While it appears counterintuitive, $\hat{\sigma}^2_{u,s}$ is also jump-robust, as we show below. The intuition is that in our model, there are no fixed points of discontinuity in X, so that the influence of any jumps is intrinsically averaged away, as $m \to \infty$, $T \to \infty$ and $n \to \infty$.

Proposition 3.1 *Assume that X is given by* (1)*. Moreover, we suppose Assumption (D1) – (D4), (V), (V'), (A) and (A'). If $m \to \infty$ and $T \to \infty$, as $n \to \infty$, then it holds that*

$$\hat{\sigma}^2_{u,s} = \sigma^2_{u,s} + O_P(mT^{-1/2}) + O_P(m^{3/2}n^{-1/2}T^{-1/2}). \tag{33}$$

Next, we note that:

$$\sqrt{\hat{\sigma}^2_{u,s}} - \sqrt{\sigma^2_{u,s}} = \hat{\sigma}_{u,s} - \sigma_{u,s} \simeq \frac{1}{2\sqrt{\sigma^2_{u,s}}}\big(\hat{\sigma}^2_{u,s} - \sigma^2_{u,s}\big), \tag{34}$$

[13] It is naturally also possible to pre-average and follow up with jump-truncation. This may lead to a better rate of convergence for the diurnal variance estimator. Here, we do not follow this line of thought, as it requires extra tuning parameters, and because the current setup appears to work reasonably well in practice.

with $\sigma^2_{u,s}$ bounded away from 0. Thus, we can write:

$$\hat{\sigma}_{u,s} = \sigma_{u,s} + O_P(mT^{-1/2}) + O_P(m^{3/2}n^{-1/2}T^{-1/2}). \tag{35}$$

It therefore follows that if

$$m \propto n^{\delta_2} \quad \text{and} \quad T \propto n^{\delta_3}, \tag{36}$$

where $\delta_2 \in (0, 1/2]$ and $\delta_3 > 1/2+2\delta_2$ (for a fixed value of δ_2), then the error induced from estimating σ_u does not alter the analysis in Section 2 and 4. That is, neither Theorem 2.1 or (27) are affected, nor is the bootstrap applied to $\check{y}(l,r)^n_i = |\Delta^n_i \bar{Y}^d|^l 1_{\{|\Delta^n_i \bar{Y}^d| < \upsilon_n\}} |\Delta^n_{i+k_n} \bar{Y}^d|^r 1_{\{|\Delta^n_{i+k_n} \bar{Y}^d| < \upsilon_n\}}$, where (with a slight abuse of notation) we redefine

$$\Delta^n_i \bar{Y}^d = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \Delta^n_{i+j-1} Y^d, \quad i = 1, \ldots, n-k_n+2, \tag{37}$$

to be based on:

$$\Delta^n_{i+j-1} Y^d = \frac{\Delta^n_{i+j-1} Y}{\hat{\sigma}_{u,\frac{i+j-1}{n}}}, \tag{38}$$

with $\hat{\sigma}_{u,\frac{i+j-1}{n}}$ from (31).

Remark 1 *If the noise is autocorrelated but not heteroscedastic,* $\hat{\omega}^2 = \hat{\text{var}}(\epsilon_{t-1+(j-1)/m})$ *given by* (32) *is no longer a consistent estimator of* $\omega^2 = \text{var}(\epsilon)$*. Indeed, when the noise is a stationary* q*-dependent sequence (for known* $q > 0$*), the statistic defined in* (32) *estimates the quantity* $2\big(\rho(0) - \rho(1)\big)$*, where* $\rho(k) = cov(\epsilon_1, \epsilon_{1+k})$*. Hautsch and Podolskij (2013, Lemma 2) discuss an estimator of* $\rho(k)$*,* $k = 0, \ldots, q+1$*, which is obtained from a simple recursion formula. Building on their result, we can deduce an estimator of* ω^2 *in this alternative setup:*

$$\hat{\omega}^2 = -\sum_{k=1}^{q+1} k\hat{\gamma}(k), \tag{39}$$

where

$$\hat{\gamma}(k) = \frac{1}{T}\frac{1}{n/m-k}\sum_{t=1}^{T}\sum_{i=1}^{n/m-k}\left(\Delta^n_{i+\left[t-1+\frac{j-1}{m}\right]_n Y}\right)\left(\Delta^n_{i+k+\left[t-1+\frac{j-1}{m}\right]_n Y}\right), \text{ for } k = 0, \ldots, q+1. \tag{40}$$

Then,

$$\hat{\omega}^2 \xrightarrow{p} \omega^2 = \rho(0). \tag{41}$$

The Monte Carlo and empirical analysis is based on

$$\hat{\sigma}^2_{u,s} \equiv \frac{1}{T}\sum_{t=1}^{T}(\sqrt{m}\Delta^m_{j+(t-1)m} Y)^2 - 2m\hat{\omega}^2, \quad \text{for } s \in [(j-1)/m, j/m), \tag{42}$$

where $\hat{\omega}^2$ *is* (39) *with* $q = 3$*.*

4 The bootstrap

In this section, we improve the quality of inference in our test of heteroscedasticity in the noisy jump-diffusion setting by relying on the bootstrap, when computing critical values for the t-statistic. This is warranted by the Monte Carlo in Section 5, which reveals that in small samples, the feasible version of (27) (cf. (66)) is poorly approximated by the standard normal. Next, we propose a bootstrap estimator of the conditional covariance matrix of the limiting process $n^{1/4}\big(\check{BV}(l_1,r_1)^n,\check{BV}(l_2,r_2)^n\big)^\intercal$, i.e. Σ. As the bootstrap estimator is positive semi-definite by construction, it renders our test implementable.

We build on a series of papers in the high-frequency volatility area. In particular, Gonçalves and Meddahi (2009) propose the wild bootstrap for realized variance, in a framework where the asset price is observed without error. Gonçalves, Hounyo, and Meddahi (2014) and Hounyo, Gonçalves, and Meddahi (2017) extend their work to accommodate noise. The latter studies the pre-averaged realized variance estimator—i.e., $BV(2,0)^n$—proposed by Jacod, Li, Mykland, Podolskij, and Vetter (2009), where the pre-averaged returns are both overlapping and heteroscedastic due to stochastic volatility. In this context, a block bootstrap applied to $(\Delta_i^n\bar{Y}^d)_{i=1}^{n-k_n+2}$ appears natural.

Nevertheless, such a scheme is only consistent if $\sigma_{sv,t}$ is constant. As shown by Hounyo, Gonçalves, and Meddahi (2017), the problem is that $|\Delta_i^n\bar{Y}^d|^2$ are heterogeneously distributed under time-varying volatility.[14] In particular, their mean and variance are unequal. This creates a bias term in the blocks of blocks bootstrap variance estimator. To cope with both dependence and heterogeneity of $|\Delta_i^n\bar{Y}^d|^2$, they combine the wild bootstrap with the blocks of blocks bootstrap. The procedure exploits that heteroscedasticity can be handled by the former, while the latter can replicate serial dependence in the data. Hounyo (2017) generalizes Hounyo, Gonçalves, and Meddahi (2017) to a broad class of covariation estimators in a general setting that accommodates jumps, microstructure noise, irregularly spaced high-frequency data and non-synchronous trading. Also, Dovonon, Gonçalves, Hounyo, and Meddahi (2014) develop a new local Gaussian bootstrap for high-frequency jump testing, but market microstructure noise is supposed to be absent. Here, we allow for noise and concentrate on heteroscedasticity.

The bootstrap version of $\check{BV}(l,r)^n$ is

$$\check{BV}(l,r)^{n*}=n^{\frac{l+r}{4}-1}\frac{1}{\mu_l\mu_r}\sum_{i=1}^{N_n}\check{y}(l,r)_i^{n*}, \tag{43}$$

where $(\check{y}(l,r)_i^{n*})_{i=1}^{N_n}$ is a bootstrap sample from $(\check{y}(l,r)_i^n)_{i=1}^{N_n}$.

We apply a bootstrap to $\check{y}(l,r)_i^n$, which replicates their dependence and heterogeneity. As suggested by Hounyo, Gonçalves, and Meddahi (2017), we merge the wild bootstrap with block-based resampling. However, our bootstrap is new, and it can be viewed as an overlapping version of their algorithm. We name it "the overlapping wild blocks of blocks bootstrap." We note that

[14] This feature is highlighted by the asymptotic distribution of $\Delta_i^n\bar{Y}^d$ in (90) below.

the degree of overlap among the blocks to be bootstrapped plays a major role in efficiency: the nonoverlapping block-based approach is less efficient than a partial or full-overlap block (e.g., Dudek, Leśkow, Paparoditis, and Politis, 2014).

To describe this approach, let b_n be a sequence of integers, which will denote the bootstrap block size, such that for some $\delta_1 \in (0,1)$:

$$b_n = O\big(n^{\delta_1}\big). \tag{44}$$

We divide $(\check{y}(l,r)_i^n)_{i=1}^{N_n}$ into overlapping blocks of size b_n. The total number of such blocks is $N_n - b_n + 1$. The bootstrap is based on $N_n - 2b_n + 2$ of them. In particular, we look at overlapping blocks within the set $(\check{y}(l,r)_i^n)_{i=1}^{N_n-b_n}$ (there is $J_n = N_n - 2b_n + 1$ many such blocks) and the last block containing the elements $\check{y}(l,r)_{N_n-b_n+1}^n, \ldots, \check{y}(l,r)_{N_n}^n$. The bootstrap sample is constructed by properly combining the first J_n blocks.

To explain this setup and avoid confusion, note that the main ingredient behind the theoretical validity of the suggested resampling scheme is that we center all bootstrap draws from a block of b_n consecutive observations, say the jth that holds $\check{y}(l,r)_j^n, \ldots, \check{y}(l,r)_{j+b_n-1}^n$, around a local average of data in the $(j+b_n)$th block (which is thus shifted to the right and consists of $\check{y}(l,r)_{j+b_n}^n, \ldots, \check{y}(l,r)_{j+2b_n-1}^n$), as given by $\bar{B}_{j+b_n}$ in (45) below. This principle is no longer applicable starting with the block that covers the elements $\check{y}(l,r)_{N_n-2b_n+2}^n, \ldots, \check{y}(l,r)_{N_n-b_n+1}^n$, because the centering here demands a local average to be computed from $\check{y}(l,r)_{N_n-b_n+2}^n, \ldots, \check{y}(l,r)_{N_n+1}^n$, and the last observation is not available.

Let $u_1, \ldots, u_{J_n+1}$ be i.i.d. random variables, whose distribution is independent of the original sample. We denote by $\mu_q^* = E^*\big(u_j^q\big)$ its qth order moments.[15] Then,

$$\bar{B}_j = \frac{1}{b_n}\sum_{i=1}^{b_n} \check{y}(l,r)_{i-1+j}^n, \quad j = 1, \ldots, N_n - b_n + 1, \tag{45}$$

is the average of the data in the jth block consisting of $\check{y}(l,r)_j^n, \ldots, \check{y}(l,r)_{j+b_n-1}^n$. Next, we generate the overlapping wild blocks of blocks bootstrap observations by:

$$\check{y}(l,r)_m^{n*} - \bar{\bar{B}}^{N_n} = \begin{cases} \frac{1}{\sqrt{b_n}}\sum_{j=1}^{m}\big(\check{y}(l,r)_m^n - \bar{B}_{b_n+j}\big)u_j, & \text{if} \quad m \in I_1^n, \\ \frac{1}{\sqrt{b_n}}\sum_{j=1}^{b_n}\big(\check{y}(l,r)_m^n - \bar{B}_{m+j}\big)u_{m+j-b_n}, & \text{if} \quad m \in I_2^n, \\ \frac{1}{\sqrt{b_n}}\sum_{j=1}^{N_n-b_n+1-m}\big(\check{y}(l,r)_m^n - \bar{B}_{J_n+1-j+b_n}\big)u_{J_n+1-j}, & \text{if} \quad m \in I_3^n, \\ \frac{1}{\sqrt{b_n}}\big(\check{y}(l,r)_m^n - \bar{B}_{N_n-b_n+1}\big)u_{J_n+1}, & \text{if} \quad m \in I_4^n, \end{cases} \tag{46}$$

[15] As usual in the bootstrap literature, P^* (E^* and var^*) denotes the probability measure (expected value and variance) induced by the resampling, conditional on a realization of the original time series. In addition, for a sequence of bootstrap statistics Z_n^*, we write (i) $Z_n^* = o_{p^*}(1)$ or $Z_n^* \xrightarrow{p^*} 0$, as $n \to \infty$, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{n\to\infty} P[P^*(|Z_n^*| > \delta) > \varepsilon] = 0$, (ii) $Z_n^* = O_{p^*}(1)$ as $n \to \infty$, if for all $\varepsilon > 0$ there exists an $M_\varepsilon < \infty$ such that $\lim_{n\to\infty} P[P^*(|Z_n^*| > M_\varepsilon) > \varepsilon] = 0$, and (iii) $Z_n^* \xrightarrow{d^*} Z$ as $n \to \infty$, if conditional on the sample Z_n^* converges weakly to Z under P^*, for all samples contained in a set with probability P converging to one.

where

$$\bar{\bar{B}}^{N_n} = \frac{1}{N_n}\sum_{i=1}^{N_n} \check{y}(l,r)_i^n, \tag{47}$$

and

$$\begin{aligned} I_1^n &= \{1,\ldots,b_n-1\}, & I_2^n &= \{b_n,\ldots,J_n\}, \\ I_3^n &= \{J_n+1,\ldots,N_n-b_n\}, & I_4^n &= \{N_n-b_n+1,\ldots,N_n\}. \end{aligned} \tag{48}$$

It is interesting to note that if we were to center $\check{y}(l,r)_m^n$ around the grand mean $\bar{\bar{B}}^{N_n}$, instead of the localized block average $\bar{B}_{j+m}$, it would yield a bootstrap observation

$$\check{y}(l,r)_m^{n*} - \bar{\bar{B}}^{N_n} = \Big(\check{y}(l,r)_m^n - \bar{\bar{B}}^{N_n}\Big)\eta_m, \tag{49}$$

for $m \in I_2^n$ (the main set), where $\eta_m = \frac{1}{\sqrt{b_n}}\sum_{j=1}^{b_n} u_{m+j-b_n}$. Therefore, under the assumption that $E(u_j) = 0$ and $\mathrm{var}(u_j) = 1$, we find that $E(\eta_m) = 0$, $\mathrm{var}(\eta_m) = 1$, and $\mathrm{cov}(\eta_m, \eta_{m-k}) = \big(1 - \frac{k}{b_n}\big)1_{\{k\le b_n\}}$. Thus, our approach is related to the dependent wild bootstrap of Shao (2010) (see also, e.g., Hounyo (2014)), who extends the traditional wild bootstrap of Wu (1986); Liu (1988) to the time series setting, and it is the special case, where the kernel function is assumed to be Bartlett (see Assumption 2.1 in Shao, 2010).

The idea of the new centering $\bar{B}_{j+m}$ is to deal with the mean heterogeneity of $\check{y}(l,r)_m^n$. As shown by Hounyo, Gonçalves, and Meddahi (2017), for the case of squared pre-averaged returns $y(2,0)_m^n$, centering the non-overlapping wild blocks of blocks bootstrap around the corresponding grand mean $N_n^{-1}\sum_{i=1}^{N_n} y(2,0)_i^n$ does not work, when $\sigma_{sv,t}$ is time-varying. In this paper, we show that generating the bootstrap observations as in (46) does yield an asymptotically valid bootstrap for $(\check{BV}(l_1,r_1)^n, \check{BV}(l_2,r_2)^n)^\intercal$, even if $\sigma_{sv,t}$ is not constant.

As in Shao (2010) and Hounyo (2014), the dependence between neighboring observations $\check{y}(l,r)_m^n$ and $\check{y}(l,r)_{m'}^n$ is not only preserved, if m and m' belong to a particular block, as typical in block-based resampling. Indeed, if $|m-m'| < b_n$, $\check{y}(l,r)_m^{n*}$ and $\check{y}(l,r)_{m'}^{n*}$ are conditionally dependent (except for the last b_n data).

A common feature of the block-based bootstrap, in particular the non-overlapping wild blocks of blocks approach of Hounyo, Gonçalves, and Meddahi (2017), is that if the sample size N_n is not a multiple of b_n, then one has to either take a shorter bootstrap sample or use a fraction of the last resampled block. This leads to some inaccuracy, when b_n is large relative to N_n. In contrast, for the overlapping version proposed in this paper, the size of the bootstrap sample is always equal to the original sample size.

Write

$$\bar{\bar{B}}^{N_n*} = \frac{1}{N_n}\sum_{i=1}^{N_n} \check{y}(l,r)_i^{n*}, \tag{50}$$

as the average value of the bootstrap observations. A closer inspection of $\bar{\bar{B}}^{N_n*}$ suggests that we

can rewrite the centered bootstrap sample mean $\bar{\bar{B}}^{N_n*} - \bar{\bar{B}}^{N_n}$ as follows

$$N_n\Big(\bar{\bar{B}}^{N_n*} - \bar{\bar{B}}^{N_n}\Big) = \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} b_n\big(\bar{B}_j - \bar{B}_{j+b_n}\big)u_j. \tag{51}$$

Thus,

$$\begin{aligned} \check{BV}(l,r)^{n*} &= \check{BV}(l,r)^{n} + n^{\frac{l+r}{4}-1}\frac{1}{\mu_l\mu_r}\frac{1}{\sqrt{b_n}}\sum_{j=1}^{J_n} b_n\big(\bar{B}_j - \bar{B}_{j+b_n}\big)u_j \\ &= \check{BV}(l,r)^{n} - \frac{1}{\sqrt{b_n}}\sum_{j=1}^{J_n} \check{\Delta B}(l,r)^n_j u_j, \end{aligned} \tag{52}$$

where

$$\check{\Delta B}(l,r)^n_j = \check{B}(l,r)^n_{j+b_n} - \check{B}(l,r)^n_j, \tag{53}$$

such that

$$\check{B}(l,r)^n_j = n^{\frac{l+r}{4}-1}\frac{1}{\mu_l\mu_r}\sum_{i=1}^{b_n} \check{y}(l,r)^n_{i-1+j}. \tag{54}$$

We can now derive the first and second bootstrap moment of $n^{1/4}\begin{pmatrix} \check{BV}(l_1,r_1)^{n*} \\ \check{BV}(l_2,r_2)^{n*} \end{pmatrix}$. The following Lemma states the formulas.

Lemma 4.1 *Assume that $\check{y}(l,r)^{n*}_m$ are generated as in* (46)*. Then, it follows that*

$$E^*\big(\check{BV}(l,r)^{n*}\big) = \check{BV}(l,r)^{n} - \frac{1}{\sqrt{b_n}}\sum_{j=1}^{J_n} \check{\Delta B}(l,r)^n_j E^*(u_j), \tag{55}$$

Also, for $1 \le i,j \le 2$,

$$\mathrm{cov}^*\big(n^{1/4}\check{BV}(l_i,r_i)^{n*}, n^{1/4}\check{BV}(l_j,r_j)^{n*}\big) = \frac{\sqrt{n}}{b_n}\sum_{k=1}^{J_n} \check{\Delta B}(l_i,r_i)^n_k \check{\Delta B}(l_j,r_j)^n_k \mathrm{var}^*(u_k). \tag{56}$$

Equation (55) of Lemma 4.1 implies that with $E^*(u_j) = 0$, $\check{BV}(l,r)^{n*}$ is an unbiased estimator of $\check{BV}(l,r)^n$, i.e. $E^*\big(\check{BV}(l,r)^{n*}\big) = \check{BV}(l,r)^n$. The second part shows that the bootstrap covariance of $n^{1/4}\check{BV}(l_i,r_i)^{n*}$ and $n^{1/4}\check{BV}(l_j,r_j)^{n*}$ depends on the variance of u. In particular, if we select $\mathrm{var}^*(u) = 1/2$ as in Hounyo, Gonçalves, and Meddahi (2017):

$$\mathrm{var}^*\left(n^{1/4}\begin{pmatrix} \check{BV}(l_1,r_1)^{n*} \\ \check{BV}(l_2,r_2)^{n*} \end{pmatrix}\right) = \check{\Sigma}^n, \tag{57}$$

where $\check{\Sigma}^n = \big(\check{\Sigma}^{l_1,r_1,l_2,r_2,n}_{ij}\big)_{1\le i,j\le 2}$ and

$$\check{\Sigma}^{l_i,r_i,l_j,r_j,n}_{ij} = \frac{\sqrt{n}}{2b_n}\sum_{k=1}^{J_n} \check{\Delta B}(l_i,r_i)^n_k \check{\Delta B}(l_j,r_j)^n_k. \tag{58}$$

Note that based on (58), we can rewrite $\check{\Sigma}^n$ as

$$\check{\Sigma}^n = \frac{\sqrt{n}}{2b_n} \sum_{j=1}^{J_n} \check{\xi}_j \check{\xi}_j^{\intercal}, \tag{59}$$

where $\check{\xi}_j \equiv \left(\Delta\check{B}(l_1, r_1)_j^n, \Delta\check{B}(l_2, r_2)_j^n\right)^{\intercal}$. It follows that if the external random variable u is selected as above, the overlapping wild blocks of blocks bootstrap variance estimator is consistent for the asymptotic variance of $n^{1/4}\big(B\check{V}(l_1, r_1)^n, B\check{V}(l_2, r_2)^n\big)^{\intercal}$ provided $\check{\Sigma}^n$ is a consistent estimator of Σ, as proved in Theorem 4.1 below. Note that $\check{\Sigma}^n$ is related to recent work on asymptotic variance estimation by Mykland and Zhang (2017); see also, e.g., Christensen, Podolskij, Thamrongrat, and Veliyev (2016); Jacod and Todorov (2009); Mancini and Gobbi (2012).

Remark 2 *Note that from* (59)*, we can also rewrite* $\check{\Sigma}^n$ *as follows:*

$$\check{\Sigma}^n = \frac{1}{b_n} \sum_{m=1}^{b_n} \check{\Sigma}_m^n, \tag{60}$$

where

$$\check{\Sigma}_m^n = \frac{\sqrt{n}}{2} \sum_{j=0}^{\lfloor N_n/b_n \rfloor - 2} \check{\xi}_{jb_n+m} \check{\xi}_{jb_n+m}^{\intercal} = \left(\check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}\right)_{1\leq i,j\leq 2}. \tag{61}$$

We deduce that the diagonal elements of $\check{\Sigma}_m^n$*, i.e.* $\check{\Sigma}_{11,m}^{l_1,r_1,l_2,r_2,n}$ *and* $\check{\Sigma}_{22,m}^{l_1,r_1,l_2,r_2,n}$ *are nothing else than the consistent bootstrap variance estimators of the asymptotic variance of* $n^{1/4}B\check{V}(l_1, r_1)^n$ *and* $n^{1/4}B\check{V}(l_2, r_2)^n$*, as proposed by Hounyo (2017).*

The next result shows that under some regularity conditions, the estimator $\check{\Sigma}^n$ converges in probability to Σ in a general Itô semimartingale context.

Theorem 4.1 *Assume that* X *fulfills Assumption (J) for some* $\beta \in [0, 2]$*. Furthermore, suppose that the conditions of Theorem B.1 in Appendix B hold true, when* X *is continuous (i.e.,* X *follows* (103)*), and also if* X *has jumps (i.e.,* X *follows* (1)*) with either*

$$l_1 + r_1 + l_2 + r_2 \leq 4(1 - \delta_1), \quad 0 \leq \beta < 4(1 - \delta_1), \tag{62}$$

or

$$l_1 + r_1 + l_2 + r_2 > 4(1 - \delta_1), \quad 0 \leq \beta < 4(1 - \delta_1), \quad \frac{l_1 + r_1 + l_2 + r_2 - 4(1 - \delta_1)}{2(l_1 + r_1 + l_2 + r_2 - \beta)} \leq \varpi < \frac{1}{2}. \tag{63}$$

Then, as $n \to \infty$*, it holds that*

$$\check{\Sigma}^n \xrightarrow{p} \Sigma, \tag{64}$$

where Σ *is defined in Appendix A.*

In our Monte Carlo studies and empirical application, we take $l_1 = r_1 = 2$ and $l_2 = r_2 = 1$. Here, (63) holds provided $\beta < 4(1 - \delta_1)$. As $1/2 < \delta_1 < 2/3$ by assumption (i.e, $4/3 < 4(1 - \delta_1) < 2$), it therefore suffices that $\beta \in [0, 4/3)$.

Theorem 4.1 implies that in finite samples, we get a consistent and nonnegative estimator of V:

$$\check{V}^n = \check{\Sigma}^n_{11} - 2\left(\frac{l_1+r_1}{l_2+r_2}\right)\left(B\check{V}(l_2,r_2)^n\right)^{\frac{l_1+r_1}{l_2+r_2}-1}\check{\Sigma}^n_{12} + \left(\frac{l_1+r_1}{l_2+r_2}\right)^2\left(B\check{V}(l_2,r_2)^n\right)^{2\left(\frac{l_1+r_1}{l_2+r_2}-1\right)}\check{\Sigma}^n_{22}. \quad (65)$$

Corollary 4.1 *Assume that the conditions from Theorem 4.1 hold true. If X is given by (10), such that Assumption (J) holds for some $\beta \in [0,1)$ and $\left(\frac{1}{2(2-\beta)} \vee \frac{3}{2(4-\beta)}\right) \leq \varpi < 1/2$. Then, if $l_1 + r_1 > l_2 + r_2$ and as $n \to \infty$,*

$$T^n \equiv \frac{n^{1/4}\left(B\check{V}(l_1,r_1)^n - \left(B\check{V}(l_2,r_2)^n\right)^{\frac{l_1+r_1}{l_2+r_2}}\right)}{\sqrt{\check{V}^n}} \xrightarrow{d} N(0,1). \quad (66)$$

Corollary 4.1 delivers the asymptotic normality of the studentized statistic T^n; the feasible version of (27). Note that under the alternative presence of heteroscedasticity, $B\check{V}(l_1,r_1)^n - \left(B\check{V}(l_2,r_2)^n\right)^{\frac{l_1+r_1}{l_2+r_2}}$ converges to a strictly positive random variable. Moreover, as $\check{V}^n$ was shown to be a robust estimator of V even in the presence of stochastic volatility, jumps and noise, we can conclude that the statistic $T^n \to \infty$ if the realization of X^d has a heteroscedastic volatility path. Therefore, appealing to the properties of stable convergence, we deduce that

$$\lim_{n\to\infty} P\left(T^n > z_{1-\alpha} \mid \Omega_{\mathcal{H}_0}\right) = \alpha, \quad (67)$$

$$\lim_{n\to\infty} P\left(T^n > z_{1-\alpha} \mid \Omega_{\mathcal{H}_a}\right) = 1 \quad (68)$$

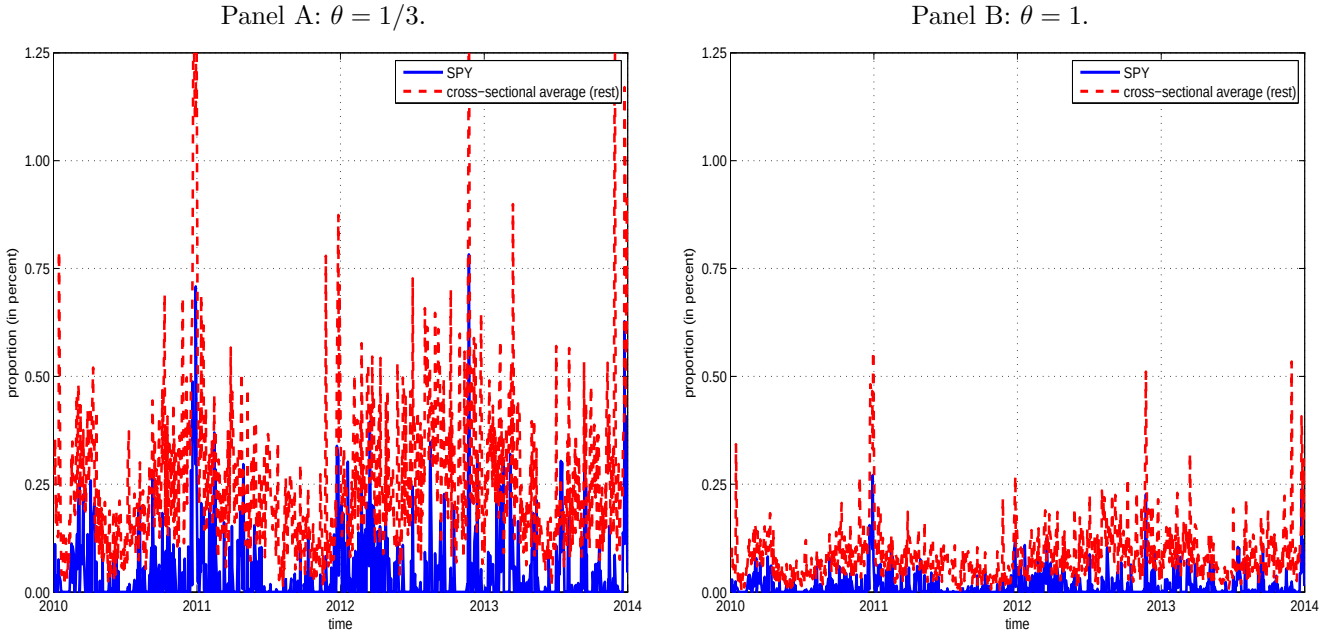
where z_α is the α-quantile of a standard normal distribution. The implication is that we reject $\mathcal{H}_0$, if T^n is significantly positive. While the alternative inference procedure based on (66) does not require any resampling, it possesses inferior finite sample properties, as shown in Section 5.

Remark 3 *The results from Jacod, Podolskij, and Vetter (2010) and Podolskij and Vetter (2009a) indicate that some assumptions can be relaxed. In particular, in Corollary 4.1, if all the powers are even numbers (e.g., $l_1 = 4, r_1 = 0, l_2 = 2$ and $r_2 = 0$), we can prove the results in the general setting of Jacod, Podolskij, and Vetter (2010) with heteroscedastic noise, which is of a general form $E\left(\epsilon^2_t \mid X\right) = \omega^2_t$ and not necessarily restricted to $E\left(\epsilon^2_t \mid X\right) = \omega^2\sigma^2_{u,t}$. Here, the null is modified as*

$$\mathcal{H}_0 : \omega \in \Omega_{\mathcal{H}_0} \cap \left\{\omega : t \longmapsto \text{var}\left(\epsilon^d_t \mid X\right) \text{ is constant on } [0,1]\right\}, \quad (69)$$

where $\epsilon^d_t \equiv \epsilon_t/\sigma_{u,t}$. This suggests that in the presence of heteroscedastic noise of a general form, $\mathcal{H}_0$ is a joint statement about the constancy of both diffusive and the rescaled noise variance. Such information should be useful in practice, because it delivers knowledge about the presence of heteroscedasticity irrespective of its origin. Meanwhile, Figure 1 shows that for our empirical high-frequency data—and the two choices of θ adopted in the paper—$B\check{V}(1,1)^n$ is almost exclusively composed by diffusive volatility. This implies that very little residual noise is left in the data after pre-averaging, which indicates that any rejection of the null is more likely due to genuine time-variation in $\sigma_{sv,t}$. We note that the dampening of the noise is naturally much weaker for $\theta = 1/3$, which is therefore more susceptible to reject $\mathcal{H}_0$ on this ground.

Figure 1: Proportion of microstructure noise in $\breve{BV}(1,1)^n$.



Note. We plot the proportion of $\breve{BV}(1,1)^n$ that is due to residual variation (after pre-averaging) in the microstructure noise. $\breve{BV}(1,1)^n$ is rescaled by $\theta\psi_2^{k_n}$ to provide an estimate of the integrated variance up to a bias term $\psi_1^{k_n}\omega^2/(\theta^2\psi_2^{k_n})$, see (21) and Theorem 2.1. The figure shows the ratio of the bias to the total $\breve{BV}(1,1)^n$ estimate over time for the ticker symbols included in our empirical analysis. ω^2 is replaced by the robust estimator $\hat{\omega}^2$ in (39) computed daily with $q = 3$.

Corollary 4.2 *Assume that the conditions of Theorem 4.1 hold true and the external random variable is chosen as* $u_j \overset{\text{i.i.d.}}{\sim} \big(E^*(u_j), \text{var}^*(u_j)\big)$, *such that* $\text{var}^*(u_j) = 1/2$. *Then, as* $n \to \infty$,

$$\text{var}^*\left(n^{1/4}\begin{pmatrix} \breve{BV}(l_1,r_1)^{n*} \\ \breve{BV}(l_2,r_2)^{n*} \end{pmatrix}\right) = \breve{\Sigma}^n \xrightarrow{p} \Sigma, \tag{70}$$

both in model (103) *and* (1), *where* Σ *is defined in Appendix A.*

Given the consistency of the bootstrap variance estimator, we now prove the associated convergence of the bootstrap distribution of $n^{1/4}\big(\breve{BV}(l_1,r_1)^{n*}, \breve{BV}(l_2,r_2)^{n*}\big)^{\intercal}$.

Theorem 4.2 *Assume that all conditions from Corollary 4.2 hold true and that* $E^*\big(|u_j|\big)^{2+\delta} < \infty$ *for some* $\delta > 0$. *Then, as* $n \to \infty$,

$$\big(\breve{\Sigma}^n\big)^{-1/2} n^{1/4}\begin{pmatrix} \breve{BV}(l_1,r_1)^{n*} - E^*\big(\breve{BV}(l_1,r_1)^{n*}\big) \\ \breve{BV}(l_2,r_2)^{n*} - E^*\big(\breve{BV}(l_2,r_2)^{n*}\big) \end{pmatrix} \xrightarrow{d^*} N(0, I_2), \tag{71}$$

in probability-P, both in model (103) *and* (1). *Moreover, let*

$$S^{n*} = \frac{n^{1/4}\left[\breve{BV}(l_1,r_1)^{n*} - \big(\breve{BV}(l_2,r_2)^{n*}\big)^{\frac{l_1+r_1}{l_2+r_2}} - \left(E^*\big(\breve{BV}(l_1,r_1)^{n*}\big) - \left(E^*\big(\breve{BV}(l_2,r_2)^{n*}\big)\right)^{\frac{l_1+r_1}{l_2+r_2}}\right)\right]}{\sqrt{V}}, \tag{72}$$

where $l_1 + r_1 > l_2 + r_2$. It holds that

$$V^{n*} \equiv \text{var}^* \left[n^{1/4} \Big(\check{BV}(l_1, r_1)^{n*} - \big(\check{BV}(l_2, r_2)^{n*} \big)^{\frac{l_1+r_1}{l_2+r_2}} \Big) \right] \xrightarrow{p} V, \tag{73}$$

and

$$S^{n*} \xrightarrow{d^*} N(0, 1), \tag{74}$$

in probability-P, both in model (10) *and* (1)*.*

Theorem 4.2 shows that the normalized statistic S^{n*} is asymptotically normal both in model (10) and (1). This implies, independently of whether $\mathcal{H}_0$ or $\mathcal{H}_a$ is true, $S^{n*} \xrightarrow{d^*} N(0, 1)$, in probability-$P$. This ensures that the following bootstrap test both controls the size and is consistent under the alternative. Let

$$\mathcal{Z}^{n*} \equiv n^{1/4} \Big[\check{BV}(l_1, r_1)^{n*} - \big(BV(l_2, r_2)^{n*} \big)^{\frac{l_1+r_1}{l_2+r_2}} - \Big(E^* \big(BV(l_1, r_1)^{n*} \big) - \big(E^* (BV(l_2, r_2)^{n*}) \big)^{\frac{l_1+r_1}{l_2+r_2}} \Big) \Big] \tag{75}$$

and

$$\mathcal{Z}^{n} \equiv n^{1/4} \Big(\check{BV}(l_1, r_1)^{n} - \big(\check{BV}(l_2, r_2)^{n} \big)^{\frac{l_1+r_1}{l_2+r_2}} \Big). \tag{76}$$

Remark 4 *We reject $\mathcal{H}_0$ at level α, if $\mathcal{Z}^n > p^*_{1-\alpha}$, where $p^*_{1-\alpha}$ is the $(1-\alpha)$-percentile of the bootstrap distribution of $\mathcal{Z}^{n*}$. Under the conditions of Theorem 4.2, the statistic $\mathcal{Z}^{n*} \xrightarrow{d^*} N(0, V)$, in probability-$P$. Note that as $\mathcal{Z}^n \xrightarrow{d_{st}} N(0, V)$ on $\Omega_{\mathcal{H}_0}$, the fact that $\mathcal{Z}^{n*} \xrightarrow{d^*} N(0, V)$, in probability-$P$, ensures that the test has correct size, as $n \to \infty$. On the other hand, under the alternative (i.e. on $\Omega_{\mathcal{H}_a}$), as $\mathcal{Z}^n$ diverges at rate $n^{1/4}$, but we still have that $\mathcal{Z}^{n*} \xrightarrow{d^*} N(0, V) = O_{p^*}(1)$, the test has unit power asymptotically.*

The above bootstrap test is convenient, as it does not require estimation of the asymptotic variance-covariance matrix Σ, but it may not lead to asymptotic refinements. In order to attain such improvement (or at least be able to prove it), we should base the bootstrap on an asymptotically pivotal t-statistic. To this end, we propose a consistent bootstrap estimator of $\check{\Sigma}^n = \text{var}^* \Big(n^{1/4} \big(\check{BV}(l_1, r_1)^{n*}, \check{BV}(l_2, r_2)^{n*} \big)^\intercal \Big)$. We look at the following adjusted bootstrap version of $\check{\Sigma}^n$ given by $\check{\Sigma}^{n*} = \Big(\check{\Sigma}^{l_1, r_1, l_2, r_2, n*}_{ij} \Big)_{1 \le i, j \le 2}$, where the individual entries of $\check{\Sigma}^{n*}$ are

$$\check{\Sigma}^{l_i, r_i, l_j, r_j, n*}_{ij} = \frac{\sqrt{n}}{b_n} \frac{\text{var}^*(u)}{E^*(u^2)} \sum_{k=1}^{J_n} \check{\Delta B}(l_i, r_i)^{n*}_k \check{\Delta B}^*(l_j, r_j)^n_k, \tag{77}$$

with

$$\check{\Delta B}(l, r)^{n*}_j = \check{\Delta B}(l, r)^n_j u_j, \tag{78}$$

where $\check{\Delta B}(l, r)^n_j$ is from (53) and $(u_j)_{j=1}^{J_n}$ are the external random variables used to generate the bootstrap observations in (46). We can also write

$$\check{\Sigma}^{n*} = \frac{\sqrt{n}}{b_n} \frac{\text{var}^*(u)}{E^*(u^2)} \sum_{j=1}^{J_n} \check{\xi}^*_j \check{\xi}^{*\intercal}_j, \tag{79}$$

where $\check{\xi}_j^* \equiv u_j\Big(\Delta\check{B}(l_1,r_1)_j^n, \Delta\check{B}(l_2,r_2)_j^n\Big)^\intercal$. We can show that $\check{\Sigma}^{n*}$ consistently estimates $\check{\Sigma}^n$ for any choice of external random variable u with $E^*\big(|u_j|^4\big) < \infty$. Next, based on $\check{\Sigma}^{n*}$ we construct a bootstrap studentized variant of (66):

$$T^{n*} \equiv \frac{\mathcal{Z}^{n*}}{\sqrt{\check{V}^{n*}}}, \tag{80}$$

where

$$\check{V}^{n*} = \check{\Sigma}_{11}^{n*} - 2\left(\frac{l_1+r_1}{l_2+r_2}\right)\big(B\check{V}(l_2,r_2)^n\big)^{\frac{l_1+r_1}{l_2+r_2}-1}\check{\Sigma}_{12}^{n*} + \left(\frac{l_1+r_1}{l_2+r_2}\right)^2\big(B\check{V}(l_2,r_2)^n\big)^{2\left(\frac{l_1+r_1}{l_2+r_2}-1\right)}\check{\Sigma}_{22}^{n*}. \tag{81}$$

Theorem 4.3 *Assume that the conditions of Corollary 4.2 are true and the external random variable is chosen as* $u_j \overset{\text{i.i.d.}}{\sim} \big(E^*(u_j), \text{var}^*(u_j)\big)$*, such that* $E^*\big(|u_j|^{4+\delta}\big) < \infty$ *for some* $\delta > 0$*. Then, as* $n \to \infty$*,*

$$\big(\check{\Sigma}^{n*}\big)^{-1/2} n^{1/4} \begin{pmatrix} B\check{V}(l_1,r_1)^{n*} - E^*\big(B\check{V}(l_1,r_1)^{n*}\big) \\ B\check{V}(l_2,r_2)^{n*} - E^*\big(B\check{V}(l_2,r_2)^{n*}\big) \end{pmatrix} \xrightarrow{d^*} N(0, I_2), \tag{82}$$

in probability-P, both in model (103) *and* (1)*. Also,*

$$T^{n*} \xrightarrow{d^*} N(0,1), \tag{83}$$

in probability-P, both in model (10) *and* (1)*.*

Theorem 4.3 shows the asymptotic normality of the studentized statistic T^{n*}. An implication of results in Theorem 4.3 is that we reject $\mathcal{H}_0$ at significance level α, if $T^n > q_{1-\alpha}^*$, where $q_{1-\alpha}^*$ is the $(1-\alpha)$-percentile of the bootstrap distribution of T^{n*}.

5 Monte Carlo analysis

We here assess the properties of the nonparametric noise- and jump-robust test of deterministic versus stochastic variation in the intraday volatility coefficient that was proposed in Section 2. We also highlight the refinements that can potentially be offered by the bootstrap, as outlined in Section 4, in sample sizes that resemble those, we tend to encounter in practice. We do so via detailed and realistic Monte Carlo simulations, and we start by describing the design of the study.

To simulate the efficient log-price X_t, we adopt the model:

$$\mathrm{d}X_t = a_t\mathrm{d}t + \sigma_t\mathrm{d}W_t + \mathrm{d}J_t, \tag{84}$$

where $X_0 = 0$, $a_t = 0.03$ (per annum) and the other components are defined below.

As above, $\sigma_t = \sigma_{sv,t}\sigma_{u,t}$. To describe $\sigma_{u,t}$, we follow earlier work of Hasbrouck (1999) and Andersen, Dobrev, and Schaumburg (2012) by using the specification:

$$\sigma_{u,t} = C + Ae^{-a_1 t} + Be^{-a_2(1-t)}. \tag{85}$$

We set $A = 0.75$, $B = 0.25$, $C = 0.88929198$ and $a_1 = a_2 = 10$.[16] This produces a pronounced, asymmetric reverse J-shape in $\sigma_{u,t}$ with a value of about 1.8 (1.1) times higher at the start (end) of each simulation compared to the observations in the middle. This is also a good description of the actual intraday volatility pattern observed in our empirical high-frequency data (cf. Panel B in Figure 5).

We assume that $\sigma_{sv,t}$ follows a stochastic volatility two-factor structure (SV2F):[17]

$$\sigma_{sv,t} = s\text{-}exp\left(\beta_0 + \beta_1\tau_{1,t} + \beta_2\tau_{2,t}\right), \tag{86}$$

where

$$\mathrm{d}\tau_{1,t} = \alpha_1\tau_{1,t}\mathrm{d}t + \mathrm{d}B_{1,t}, \qquad \mathrm{d}\tau_{2,t} = \alpha_2\tau_{2,t}\mathrm{d}t + (1+\phi\tau_{2,t})\mathrm{d}B_{2,t}. \tag{87}$$

Here, $B_{1,t}$ and $B_{1,t}$ are two independent standard Brownian motions with $E\left(\mathrm{d}W_t\mathrm{d}B_{1,t}\right) = \rho_1\mathrm{d}t$ and $E\left(\mathrm{d}W_t\mathrm{d}B_{2,t}\right) = \rho_2\mathrm{d}t$.

We follow Huang and Tauchen (2005) and use the parameters $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$ and $\rho_1 = \rho_2 = -0.3$.[18] This means that the first factor becomes a slowly-moving component, which generates persistence in volatility, while the second is a fast mean-reverting process that allows for a sufficient amount of volatility-of-volatility. At the start of each simulation, we initialize τ_1 at random from its stationary distribution, i.e. $\tau_{1,0} \sim N\big(0, -[2\alpha_1]^{-1}\big)$. Meanwhile, τ_2 is started at $\tau_{2,0} = 0$ (e.g., Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008).

In absence of stochastic volatility, i.e. under the null hypothesis of deterministic diurnal variation, we freeze $\sigma_{sv,t}$ at a value equal to the unconditional expectation implied by the above SV2F model, i.e. $\sigma^2 = E(\sigma^2_{sv,t})$.

J_t is a symmetric tempered stable process with Lévy measure:

$$\nu(\mathrm{d}x) = c\frac{e^{-\lambda x}}{x^{1+\beta}}\mathrm{d}x, \tag{88}$$

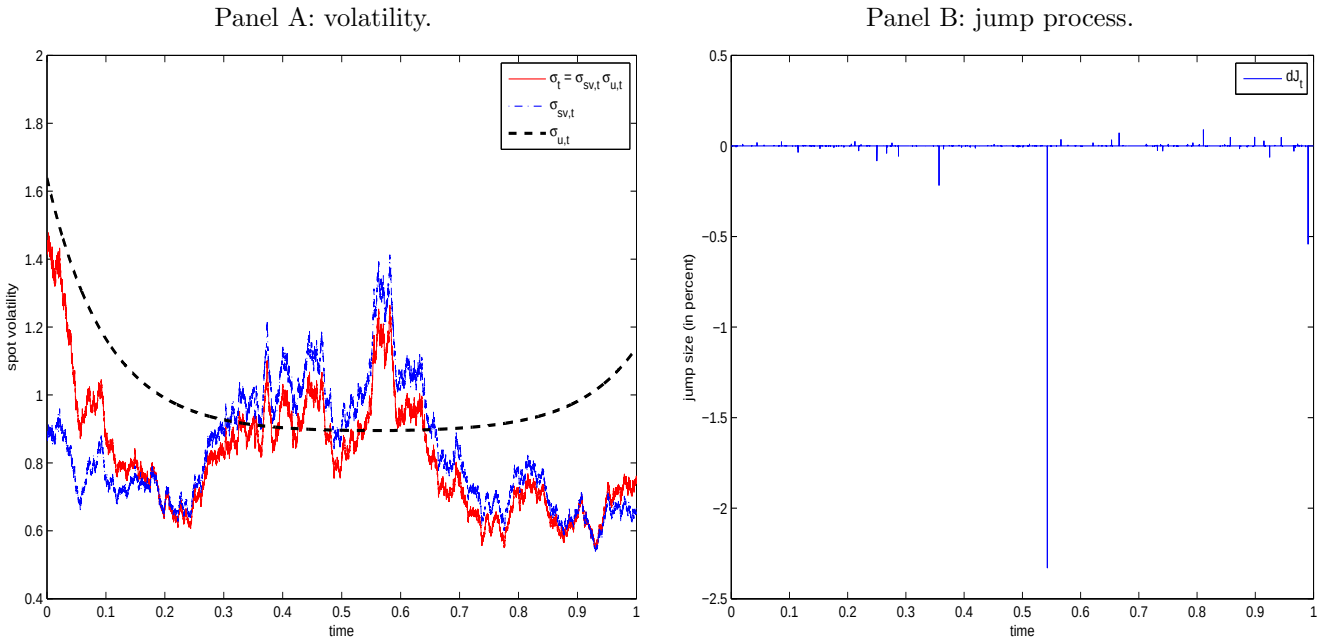
where $c > 0$, $\lambda > 0$, and $\beta \in [0, 2)$ measures the degree of jump activity. We assume $\lambda = 3$ and $\beta = 0.5$. The choice of β produces an infinite-activity, finite-variation process and is consistent with Theorem 4.1. The idea is to subdue X_t to a stream of small jumps that, in contrast to the large ones, are typically difficult to filter via truncation, and which can be confused by the t-statistic with stochastic volatility. We therefore anticipate that this setup induces some size distortions in the test. c is calibrated so that J_t accounts for 20% of the quadratic variation. This parameterization aligns with other papers (e.g., Aït-Sahalia, Jacod, and Li, 2012; Aït-Sahalia and Xiu, 2015; Huang and Tauchen, 2005).

[16] The calibration of C ensures that $\int_0^1 \sigma^2_{u,t}\mathrm{d}t = 1$.

[17] The s-exp function is used to denote the exponential function that has been spliced with a polynomial of linear growth at high values of its argument, i.e. $s\text{-}\exp(x) = e^x$ if $x \leq x_0$ and $s\text{-}\exp(x) = \frac{e^{x_0}}{\sqrt{x_0 - x_0^2 + x^2}}$, if $x > x_0$. As advocated by Chernov, Gallant, Ghysels, and Tauchen (2003), we set $x_0 = \ln(1.5)$.

[18] Note that these parameters are annualized. We assume there are 250 trading days in a year.

Figure 2: Illustration of a simulation.



Note. The figure shows a sample path of the two main ingredients in the jump-diffusion model (from the first simulation of 1,000 replica in total). In Panel A, the volatility is measured relative to its unconditional average. In Panel B, as barely noticeable, the tempered stable jump process has many small increments that are close, but not equal, to zero.

We approximate the continuous time representation of $\sigma_{sv,t}$ using an Euler scheme, while J_t is generated as the difference between two spectrally positive tempered stable processes, which are simulated using the acceptance-rejection algorithm of Baeumer and Meerschaert (2010), as described in Todorov, Tauchen, and Grynkiv (2014).[19] Note that the latter is exact, if $\beta < 1$, as is the case here.

We simulate data for $t \in [0, 1]$ (this is thought of as corresponding to a trading session on a US stock exchange, which spans 6.5 hours), where the discretization step is $\Delta t = 1/23,400$ (i.e., time runs on a one second grid).

In Figure 2, we provide an illustration of a realization of the volatility and jump process from the full model.

A total of $T = 1,000$ Monte Carlo replica is generated. In each simulation, we pollute the efficient price with an additive noise term by setting $Y_{i/n} = X_{i/n} + \epsilon_{i/n}$. To capture the well-known negative serial correlation in log-returns induced by bid-ask bounce in transaction prices and apparent second-order effects present in our real data (cf. Panel A in Figure 5), we follow Kalnina (2011) and model $\epsilon_{i/n}$ (for a given observation frequency n) as an MA(1):

$$\epsilon_{i/n} = \epsilon'_{i/n} + \varphi \epsilon'_{(i-1)/n}, \quad \text{where} \quad \epsilon'_{i/n} \mid (\sigma_t)_{t\in[0,1]} \overset{\text{i.i.d.}}{\sim} N\left(0, \frac{\omega^2}{1+\varphi^2}\right), \tag{89}$$

[19] We thank Viktor Todorov for sharing Matlab code to simulate a tempered stable process.

so that $\text{var}(\epsilon) = \omega^2$.

To gauge how the strength of autocorrelation in ϵ affects our results, we consider $\varphi = 0$, -0.3, -0.5, and -0.9. Of course, the first value corresponds to the i.i.d. noise case. To model the magnitude of ϵ, we set $\omega^2 = \xi^2 \sqrt{\int_0^1 \sigma_t^4 \mathrm{d}t}$, such that the variance of the market microstructure component scales with volatility (e.g., Bandi and Russell, 2006; Kalnina and Linton, 2008). As in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008), we fix $\xi^2 = 0.0001$, 0.001 and 0.01, as motivated by the empirical work of Hansen and Lunde (2006), who find these to be typical sizes of noise contamination for the 30 stocks in the Dow Jones Industrial Average index (see also, e.g., Aït-Sahalia and Yu, 2009).

We inspect both the infeasible setting, where the true—but unknown—diurnal factor $\sigma_{u,t}$ is applied to standardize the noisy high-frequency return in (12) and the feasible version from (38), where the estimator $\hat{\sigma}_{u,t}$ in (31) is used.

We compute $\hat{\sigma}_{u,t}$ based on $m = 78$ (i.e., 5-minute data) and bias-correct with the estimator of ω^2 from (39) with $q = 3$.[20] Although the noise is at most 1-dependent here, our implementation matches the empirical application, where this value of q suffices to capture the autocorrelation found in our real high-frequency data, as evident from Panel A in Figure 5. In Figure 3, we plot $\hat{\sigma}_{u,t}$ alongside $\sigma_{u,t}$ as estimated both under $\mathcal{H}_0$ and $\mathcal{H}_a$. As seen, $\hat{\sigma}_{u,t}$ is roughly unbiased, but it exhibits higher variation around $\sigma_{u,t}$ in the presence of stochastic volatility, which adds measurement error to the calculations.[21]

We then construct noisy deflated log-returns $\Delta_i^n Y^d \equiv Y^d_{i/n} - Y^d_{(i-1)/n}$ at sampling frequency $n = 390$, 780, 1560, 4680, 7800, 11700 and 23400, thereby varying the sample size across a broad range of selections. With the above interpretation of time, the smallest (largest) value of n amounts to observing a new price every minute (second). Such a number of trade arrivals is not unrealistic compared to real high-frequency data, as reported in Section 6.

We pre-average using (17) and (37), which we do locally on a window of size $k_n = [\theta\sqrt{n}]$. We work with $\theta = 1/3$ and $\theta = 1$ (as also done in, e.g., Christensen, Kinnebrock, and Podolskij, 2010).[22] As standard, the weight function is $g(x) = \min(x, 1-x)$.

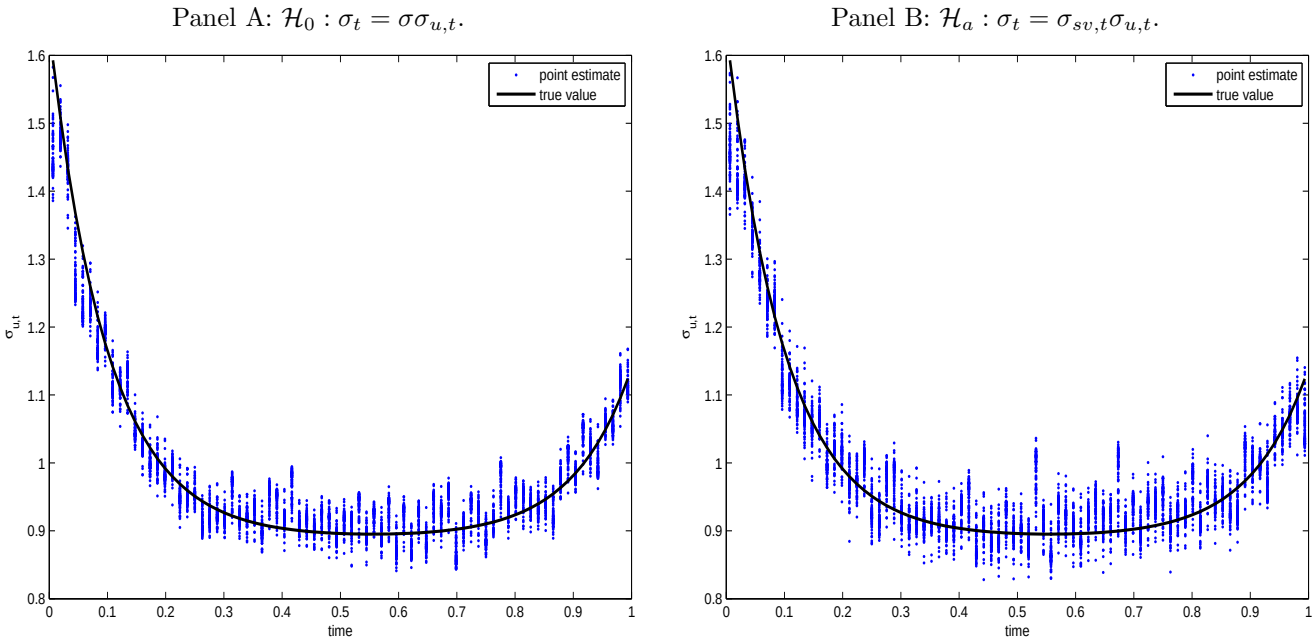
The t-statistic is based on comparing $\check{BV}(l,r)^n$ with $l_1 = r_1 = 2$ and $l_2 = r_2 = 1$. To truncate, we set $v_n = cu_n^{\varpi}$ with $u_n = k_n/n$ in (24), which is adapted to an estimate of the local spot volatility. As in, e.g., Li, Todorov, and Tauchen (2013, 2016), we fix the "rate" parameter $\varpi = 0.49$, while we determine the "scale" dynamically as $c = \Phi(0.999)\sqrt{BV(1,1)^n}$, where $\Phi(0.999)$ is the 99.9%-quantile from the standard normal distribution and $BV(1,1)^n$ is the non-truncated estimator in

[20] The noise plays a limited role at this sampling frequency, but we prefer a low value of m to accommodate all the choices of n in the simulation design. In our empirical work, we exploit 30-second data (i.e., $m = 780$), as the instruments analyzed there are very liquid and, as a result, sample sizes are generally large.

[21] To illustrate the robustness of our approach, we do not filter $\hat{\sigma}_{u,t}$, although it appears natural to exploit the smoothness condition in Assumption (D3) to reduce sampling errors further.

[22] As this introduces a small rounding effect in the relation between θ and k_n, we therefore reset $\theta = k_n/\sqrt{n}$ following the determination of k_n. We apply this "effective" θ in all the subsequent computations, as also advocated in Jacod, Li, Mykland, Podolskij, and Vetter (2009).

Figure 3: Estimation of the diurnal component.



Note. We plot $\sigma_{u,t}$ from (85). An estimate is recovered from $\hat{\sigma}_{u,t}$ proposed in (31) based on $m = 78$. The bias-correction is done via $\hat{\omega}^2$ in (39) with $q = 3$. Panel A depicts estimates under $\mathcal{H}_0 : \sigma_t = \sigma\sigma_{u,t}$, while Panel B is for $\mathcal{H}_a : \sigma_t = \sigma_{sv,t}\sigma_{u,t}$. Each point on a vertical line corresponds to an estimate in that time interval for some combination of n, ξ^2 and φ.

(20). The intuition behind this construction is as follows. Assume there are no jumps in the interval $[i/n, (i + k_n)/n]$. Then, under mild regularity conditions:

$$n^{1/4}\Delta_i^n \bar{Y}^d \overset{a}{\sim} N\left(0, \theta\psi_2\sigma^2_{sv,i/n} + \frac{1}{\theta}\psi_1\omega^2\right). \tag{90}$$

It follows from (21) that $BV(1,1)^n \xrightarrow{p} \int_0^1 \left(\theta\psi_2\sigma^2_{sv,s} + \frac{1}{\theta}\psi_1\omega^2\right)\mathrm{d}s$, so that $\sqrt{BV(1,1)^n}$ is a (jump-robust) measure of the average dispersion (i.e., standard deviation) of the sequence $\Delta_i^n\bar{Y}^d$, while $\Phi(\cdot)$ controls how far out in the tails of the distribution truncation is enforced.[23] On the other hand, while $\Delta_i^n\bar{Y}^d$ is of order $O_p\left(n^{-1/4}\right)$, $\varpi \in (0, 1/2)$ implies that u_n^{ϖ} shrinks at a slower pace than $\Delta_i^n\bar{Y}^d$. Therefore, purely "continuous" returns fall within the boundary of the threshold asymptotically. In contrast, if there are jumps in $[i/n, (i + k_n)/n]$, $\Delta_i^n\bar{Y}^d$ usually has order $O_p(1)$, and such "discontinuous" returns are, eventually, discarded.

The bootstrap inference is done as follows. We resample the pre-averaged high-frequency data $B = 999$ times for each Monte Carlo replication. Application of our bootstrap also requires the selection of the external random variable u. This is an important choice in practice, and consistent with previous work (e.g., Hounyo, 2017; Hounyo, Gonçalves, and Meddahi, 2017) we examine the

[23] While the pre-averaged (1,1)-bipower variation is robust to the presence of jumps in the p-lim, as $n \to \infty$, in practice it tends to be slightly upward biased for a finite value of n, because the jumps are not completely eliminated, see, for example, Christensen, Oomen, and Podolskij (2014).

robustness of our approach by adopting two candidate distributions:[24]

(1.) $u_j \sim N(0, 1/2)$.

(2.)

$$u_j = \begin{cases} \frac{1}{\sqrt{2}}\left(\frac{1-\sqrt{5}}{2}\right), & \text{with probability} \quad p = \frac{\sqrt{5}+1}{2\sqrt{5}} \\ \frac{1}{\sqrt{2}}\left(\frac{1+\sqrt{5}}{2}\right), & \text{with probability} \quad 1-p = \frac{\sqrt{5}-1}{2\sqrt{5}}. \end{cases} \tag{91}$$

In both cases, $E^*(u_j) = 0$ and $\text{var}^*(u_j) = 1/2$, so these are asymptotically valid choices of u_j for the purpose of constructing a bootstrap test based on studentized and unstudentized statistics. The two-point distribution in (2.) was originally proposed by Mammen (1993), and here we just scale it such that its variance is a half.

Estimation of the asymptotic variance-covariance matrix Σ depends on the block size $b_n = O(n^\delta)$ with $1/2 < \delta < 2/3$. Of course, this means nothing other than eventually $b_n = cn^\delta$, for some constant c. There is no available theory, which can help us find optimal choices of c and δ (e.g., via a MSE criterion). Moreover, in finite samples any fixed block size b_n can be achieved from many combinations of c and δ. Set against this upshot, we propose the following. We fix $\delta = 2/3$ at the upper bound (again, the inequality constraint is only binding in the limit). We set $b_n^{\min} = [2n^\delta]$ and $b_n^{\max} = [\min(3n^\delta, N_n/2)]$. The first choice is motivated, since we need at least $b_n \geq 2k_n$ for the estimator to capture the dependence in $(\check{y}(l,r)_i^n)_{i=1}^{N_n}$, while the latter amounts to saying b_n should also not be too large compared to N_n. We then partition $[b_n^{\min}, b_n^{\max}]$ into 30 equidistant subintervals and loop b_n over the integers that are closest to the endpoints. We select an "optimal" value of b_n by using the minimum variance criterion of Politis, Romano, and Wolf (1999) with a two-sided averaging window of length $d = 2$.

In Table 1 – 4, we report the rejection rates—averaged across simulations—of the above jump- and noise-robust test of $\mathcal{H}_0$ at the 5% level of significance. The critical value in each test is found either via the 95% quantile of the standard normal distribution function (labeled CLT), as motivated by the asymptotic theory in Corollary 4.1, or with the help of the bootstrap-based percentile and percentile-t approach—with the headings $z_{\text{wb}\cdot}$ and $t_{\text{wb}\cdot}$—for the two external random variates u introduced above.

Throughout, we highlight the setting with $\varphi = -0.5$, while noting the simulated size and power for other values of φ are generally within ± 1%-point of the numbers reported here (the latter are omitted, but available at request). This is also true for the noise variance parameter, ξ, which changes the results in a limited way, if at all, as gauged by inspection of Panel A – C in each table. As such, neither of the parameters associated to noise has a material effect on the outcome of

[24] We also experimented with a third external random variable using an alternative formulation of the two-point distribution, where $u_j = \pm 1$ with probability $p = 1 - p = 1/2$. The outcome was more or less identical to the results we report based on (2.), so we decided to exclude these results to save space.

Table 1: Infeasible rejection rate at 5% level of significance with $\theta = 0.333$ and $\varphi = -0.5$.

	$\mathcal{H}_0$: deterministic volatility										$\mathcal{H}_a$: stochastic volatility									
	size					avg. block length					power					avg. block length				
	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}
Panel A: $\xi^2 = 0.0001$																				
$n =$ 390	23.0	6.1	6.3	9.3	9.9	136	130	131	132	131	63.1	33.5	33.9	40.0	40.3	136	131	130	132	132
780	17.1	5.1	4.4	7.6	7.9	216	210	211	210	211	63.3	42.2	41.3	45.5	44.8	216	211	212	211	210
1560	12.1	4.4	4.5	5.5	5.7	343	335	336	336	333	65.6	47.4	48.3	50.5	49.7	341	335	334	335	336
4680	10.7	5.0	4.9	5.7	6.3	708	701	700	697	695	70.3	59.4	59.0	60.2	60.1	706	698	700	697	698
7800	10.2	6.0	5.4	7.2	7.5	998	981	986	982	985	74.3	66.2	65.9	66.6	65.7	995	977	980	983	978
11700	10.2	5.6	5.7	6.8	7.2	1,302	1,292	1,291	1,297	1,286	78.5	70.5	71.4	70.2	70.5	1,305	1,289	1,286	1,288	1,278
23400	8.1	5.1	5.5	5.9	5.9	2,068	2,042	2,046	2,040	2,046	83.2	78.7	79.3	79.7	79.7	2,067	2,045	2,045	2,045	2,038
Panel B: $\xi^2 = 0.0010$																				
$n =$ 390	23.7	5.7	6.9	9.4	9.7	136	131	131	132	131	60.6	32.0	31.6	38.4	39.1	136	131	131	131	131
780	16.9	5.0	4.9	6.9	7.2	214	211	210	211	210	62.7	40.5	39.8	44.3	44.2	216	211	211	210	211
1560	13.4	4.3	4.0	5.8	5.7	342	335	336	338	337	63.2	47.2	46.0	49.9	49.7	340	335	336	334	336
4680	10.9	4.8	5.2	6.6	6.6	706	698	701	697	699	69.5	60.3	60.1	60.2	59.6	706	696	700	694	700
7800	9.7	5.1	5.5	6.4	6.2	998	980	988	977	982	72.3	64.6	64.5	64.5	64.4	998	982	986	982	986
11700	10.1	5.7	5.5	6.7	7.3	1,308	1,294	1,287	1,287	1,286	76.9	69.8	70.7	69.6	69.8	1,306	1,284	1,294	1,285	1,294
23400	8.3	5.7	5.5	6.0	5.8	2,062	2,043	2,051	2,053	2,050	83.0	78.5	78.5	78.6	79.0	2,054	2,053	2,051	2,052	2,048
Panel C: $\xi^2 = 0.0100$																				
$n =$ 390	24.7	7.3	7.8	11.3	10.9	136	131	131	131	131	48.2	21.9	21.1	25.5	26.0	135	131	131	130	131
780	16.2	3.7	4.0	6.5	7.0	215	211	211	212	211	52.1	24.6	24.2	29.8	30.2	217	210	211	210	210
1560	14.7	5.2	5.2	8.0	7.9	341	336	334	336	335	49.4	30.8	31.3	35.2	35.8	340	336	335	335	337
4680	12.0	6.2	6.1	7.5	7.7	706	705	698	698	701	60.9	47.5	48.0	49.0	49.0	707	703	703	700	700
7800	10.1	4.0	4.2	5.9	6.2	998	980	991	982	987	65.1	54.9	55.0	55.1	55.0	997	978	979	980	973
11700	9.0	5.7	5.4	5.9	6.1	1,298	1,287	1,295	1,291	1,296	68.0	60.2	59.5	60.5	60.0	1,294	1,290	1,291	1,282	1,290
23400	8.9	5.7	6.1	7.3	6.8	2,055	2,050	2,036	2,045	2,039	73.7	68.3	68.1	68.7	68.3	2,060	2,045	2,041	2,052	2,037

Note. We simulate from a model with drift, volatility, infinite-activity jumps and microstructure noise. We test the hypothesis that $\sigma_t = \sigma\sigma_{u,t}$ is a deterministic function of time (induced by diurnal variation) and report rejection rates both under $\mathcal{H}_0$ (size) and $\mathcal{H}_a$ (power). In the latter, $\sigma_t = \sigma_{sv,t}\sigma_{u,t}$ is also time-varying due to a two-factor SV structure. θ is a tuning parameter that is used to compute the pre-averaging window $k_n = [\theta\sqrt{n}]$, φ is the MA(1) coefficient in the noise process, n is the sample size, and ξ^2 controls the magnitude of noise relative to volatility. CLT is for the asymptotic theory from (66), while $z_{\text{wb}\cdot}$ and $t_{\text{wb}\cdot}$ are rejection rates based on the percentile and percentile-t bootstrap test for two choices of the external random variable u. We made 1,000 Monte Carlo trials with 999 bootstrap replica in each simulation. Further details can be found in Section 5.

Table 2: Infeasible rejection rate at 5% level of significance with $\theta = 1.000$ and $\varphi = -0.5$.

		$\mathcal{H}_0$: deterministic volatility										$\mathcal{H}_a$: stochastic volatility									
		size					avg. block length					power					avg. block length				
		CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}
Panel A: $\xi^2 = 0.0001$																					
$n =$	390	21.1	9.0	9.1	11.4	10.9	135	130	130	130	130	48.3	24.7	23.6	28.8	28.9	135	131	131	131	129
	780	13.7	6.0	5.3	7.1	7.2	215	211	210	210	210	43.4	24.2	23.5	27.7	27.4	217	210	211	210	211
	1560	13.3	6.6	5.7	7.7	8.1	339	336	336	336	334	47.8	34.3	33.7	36.7	36.1	342	335	335	336	334
	4680	10.7	6.8	5.5	7.7	7.2	708	702	698	698	697	51.8	43.2	43.5	43.2	43.4	708	695	697	697	698
	7800	9.1	5.2	5.1	6.1	6.2	996	982	982	974	985	58.4	48.7	48.1	49.0	49.0	994	985	979	982	977
	11700	9.4	5.3	5.4	7.0	6.4	1,296	1,294	1,286	1,285	1,289	62.6	53.3	54.2	53.0	53.2	1,304	1,293	1,286	1,289	1,284
	23400	9.1	6.8	6.7	7.1	7.2	2,060	2,052	2,036	2,059	2,040	67.5	62.3	61.0	60.8	61.5	2,069	2,033	2,036	2,043	2,048
Panel B: $\xi^2 = 0.0010$																					
$n =$	390	21.2	9.6	8.6	10.8	11.3	135	130	131	131	130	49.3	24.3	23.5	28.6	28.1	135	130	131	131	130
	780	13.9	5.9	5.6	7.4	7.6	216	209	211	210	211	43.8	24.2	23.1	27.3	28.3	216	211	211	211	210
	1560	13.3	6.4	6.1	7.8	7.9	339	335	336	334	334	48.0	34.1	33.3	36.3	35.9	340	336	335	336	334
	4680	10.8	6.5	5.8	7.4	7.3	710	699	705	696	702	52.4	43.2	43.8	43.5	43.2	705	700	702	700	699
	7800	9.2	5.2	5.1	5.8	5.9	996	981	988	982	981	58.3	48.2	48.1	47.8	48.1	1,000	981	982	979	981
	11700	9.4	5.3	5.3	7.0	6.4	1,295	1,293	1,285	1,292	1,283	62.4	53.4	53.3	53.7	52.9	1,297	1,289	1,285	1,278	1,287
	23400	9.1	6.5	6.4	7.4	7.3	2,047	2,046	2,026	2,049	2,043	67.3	61.7	62.1	60.6	61.1	2,074	2,045	2,043	2,041	2,052
Panel C: $\xi^2 = 0.0100$																					
$n =$	390	21.6	9.1	9.0	12.9	12.2	135	131	131	131	131	48.9	24.5	24.2	28.6	28.0	136	131	130	131	130
	780	14.1	5.6	5.9	7.5	8.1	215	210	210	210	210	42.2	24.0	23.9	28.0	27.4	215	210	210	210	210
	1560	14.3	7.3	6.6	8.5	8.4	342	335	336	336	336	47.2	33.1	33.0	34.9	35.4	341	335	334	337	334
	4680	10.2	5.5	5.6	7.2	7.3	707	701	703	700	701	52.3	41.2	41.0	42.3	41.8	708	700	701	695	698
	7800	9.4	4.9	5.1	7.1	6.7	998	981	986	979	986	56.9	45.3	45.2	46.1	46.0	994	976	984	975	980
	11700	9.1	6.0	5.9	6.8	6.7	1,298	1,289	1,285	1,287	1,288	61.1	51.5	52.0	51.4	51.1	1,299	1,284	1,283	1,285	1,290
	23400	9.3	7.1	6.9	7.7	7.7	2,056	2,047	2,044	2,044	2,050	65.9	60.1	60.1	59.9	59.9	2,065	2,037	2,044	2,039	2,045

Note. We simulate from a model with drift, volatility, infinite-activity jumps and microstructure noise. We test the hypothesis that $\sigma_t = \sigma\sigma_{u,t}$ is a deterministic function of time (induced by diurnal variation) and report rejection rates both under $\mathcal{H}_0$ (size) and $\mathcal{H}_a$ (power). In the latter, $\sigma_t = \sigma_{sv,t}\sigma_{u,t}$ is also time-varying due to a two-factor SV structure. θ is a tuning parameter that is used to compute the pre-averaging window $k_n = [\theta\sqrt{n}]$, φ is the MA(1) coefficient in the noise process, n is the sample size, and ξ^2 controls the magnitude of noise relative to volatility. CLT is for the asymptotic theory from (66), while $z_{\text{wb}\cdot}$ and $t_{\text{wb}\cdot}$ are rejection rates based on the percentile and percentile-t bootstrap test for two choices of the external random variable u. We made 1,000 Monte Carlo trials with 999 bootstrap replica in each simulation. Further details can be found in Section 5.

Table 3: Feasible rejection rate at 5% level of significance with $\theta = 0.333$ and $\varphi = -0.5$.

		$\mathcal{H}_0$: deterministic volatility										$\mathcal{H}_a$: stochastic volatility									
		size					avg. block length					power					avg. block length				
		CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}
Panel A: $\xi^2 = 0.0001$																					
$n =$	390	25.3	6.9	7.3	10.1	10.1	136	130	131	130	131	62.0	34.9	35.7	42.1	41.1	136	131	131	132	131
	780	18.5	5.6	5.5	8.0	8.3	215	210	211	212	211	64.2	43.2	41.9	46.2	47.0	215	210	211	210	208
	1560	13.5	5.1	4.9	6.4	6.3	342	335	335	334	334	66.7	50.5	50.1	52.5	52.4	341	335	334	335	335
	4680	12.0	5.7	5.9	6.5	7.2	712	698	701	700	700	71.5	61.6	61.9	62.3	62.1	707	700	698	699	700
	7800	11.7	7.5	7.4	9.1	9.2	999	978	984	984	981	76.4	68.3	68.2	68.4	68.6	995	979	978	978	982
	11700	11.2	6.4	6.7	8.4	8.2	1,297	1,284	1,283	1,283	1,295	79.5	73.3	72.8	73.5	73.2	1,308	1,291	1,293	1,288	1,284
	23400	10.7	7.1	6.9	8.0	7.8	2,061	2,052	2,049	2,047	2,051	85.7	81.2	82.1	82.1	82.3	2,062	2,043	2,042	2,042	2,049
Panel B: $\xi^2 = 0.0010$																					
$n =$	390	24.6	7.0	6.6	10.7	11.1	136	131	131	132	131	61.6	33.3	33.1	39.2	40.2	135	131	131	131	131
	780	19.0	5.4	5.7	8.0	8.9	215	211	211	211	210	63.7	41.3	39.9	45.6	44.9	216	211	212	211	210
	1560	14.6	4.8	4.5	6.3	6.2	343	336	336	335	335	64.6	48.5	48.1	51.7	51.6	341	336	337	334	336
	4680	12.3	5.8	5.9	7.1	7.4	705	699	700	702	696	70.6	60.9	60.2	60.5	60.3	708	702	702	699	702
	7800	11.8	6.6	6.8	8.3	8.1	998	982	985	983	981	75.0	66.6	66.6	66.9	66.2	998	984	983	978	973
	11700	11.1	6.4	6.8	8.3	7.8	1,311	1,285	1,292	1,287	1,289	79.5	72.8	72.8	71.8	72.5	1,306	1,283	1,293	1,285	1,295
	23400	10.1	6.3	6.5	7.8	8.4	2,057	2,039	2,041	2,049	2,038	84.5	81.0	80.6	80.5	81.2	2,056	2,040	2,054	2,037	2,044
Panel C: $\xi^2 = 0.0100$																					
$n =$	390	24.7	6.6	6.7	10.1	10.6	136	131	130	131	131	49.0	20.6	21.1	26.3	26.7	135	131	131	131	131
	780	18.0	3.6	4.0	5.8	6.0	216	211	210	211	211	52.3	26.3	25.1	31.3	31.6	216	210	211	209	212
	1560	15.0	5.1	5.1	7.7	7.8	343	335	334	334	335	50.2	31.3	30.7	35.2	35.0	341	339	337	337	336
	4680	12.3	6.2	6.4	7.8	8.0	706	702	699	699	697	61.9	48.4	48.1	49.7	49.6	707	702	699	703	700
	7800	10.6	5.4	4.8	6.5	6.4	999	978	989	979	985	65.9	57.7	57.0	58.0	58.1	997	982	981	984	980
	11700	9.4	6.1	6.3	7.0	6.9	1,305	1,287	1,291	1,285	1,293	69.8	62.4	61.4	60.7	61.0	1,301	1,293	1,291	1,279	1,286
	23400	9.9	6.4	6.3	8.4	8.1	2,050	2,033	2,051	2,026	2,041	75.5	69.4	69.1	69.9	70.3	2,058	2,032	2,036	2,038	2,041

Note. We simulate from a model with drift, volatility, infinite-activity jumps and microstructure noise. We test the hypothesis that $\sigma_t = \sigma\sigma_{u,t}$ is a deterministic function of time (induced by diurnal variation) and report rejection rates both under $\mathcal{H}_0$ (size) and $\mathcal{H}_a$ (power). In the latter, $\sigma_t = \sigma_{sv,t}\sigma_{u,t}$ is also time-varying due to a two-factor SV structure. θ is a tuning parameter that is used to compute the pre-averaging window $k_n = [\theta\sqrt{n}]$, φ is the MA(1) coefficient in the noise process, n is the sample size, and ξ^2 controls the magnitude of noise relative to volatility. CLT is for the asymptotic theory from (66), while $z_{\text{wb}\cdot}$ and $t_{\text{wb}\cdot}$ are rejection rates based on the percentile and percentile-t bootstrap test for two choices of the external random variable u. We made 1,000 Monte Carlo trials with 999 bootstrap replica in each simulation. Further details can be found in Section 5.

Table 4: Feasible rejection rate at 5% level of significance with $\theta = 1.000$ and $\varphi = -0.5$.

		$\mathcal{H}_0$: deterministic volatility										$\mathcal{H}_a$: stochastic volatility									
		size					avg. block length					power					avg. block length				
		CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}	CLT	z_{wb1}	z_{wb2}	t_{wb1}	t_{wb2}
Panel A: $\xi^2 = 0.0001$																					
$n =$	390	20.5	8.9	8.6	11.6	11.9	135	131	130	131	130	49.1	24.0	23.7	28.7	28.4	136	131	131	132	130
	780	14.4	5.2	5.0	7.4	7.5	215	211	210	210	209	42.8	24.9	25.1	27.6	27.3	216	210	209	209	211
	1560	14.2	6.4	6.5	7.9	8.1	341	335	337	335	334	48.4	35.3	34.8	37.0	36.9	342	334	335	336	334
	4680	11.7	6.9	6.1	8.3	8.1	710	703	698	700	699	54.1	44.3	44.2	44.5	44.4	710	698	700	694	695
	7800	10.9	6.5	5.8	7.0	7.4	996	983	985	982	975	60.3	51.4	50.6	51.1	51.8	990	981	979	981	976
	11700	10.4	6.2	7.0	7.7	7.9	1,292	1,295	1,296	1,294	1,291	64.2	55.6	56.1	56.2	55.7	1,302	1,284	1,281	1,285	1,292
	23400	10.5	7.8	8.0	8.0	7.8	2,066	2,053	2,045	2,064	2,038	69.0	63.6	63.9	63.0	63.0	2,066	2,046	2,047	2,049	2,036
Panel B: $\xi^2 = 0.0010$																					
$n =$	390	21.0	9.6	8.5	11.3	11.3	135	130	130	131	130	49.9	24.5	24.3	28.3	29.1	135	131	130	131	130
	780	14.4	5.1	5.3	7.4	8.1	215	209	210	209	211	44.9	25.6	25.3	28.8	29.2	215	211	211	211	210
	1560	14.1	6.4	6.3	8.3	8.7	339	335	337	336	334	48.1	34.6	34.6	37.3	36.9	341	335	333	335	334
	4680	11.5	7.2	6.1	8.1	8.0	708	700	700	696	699	53.2	44.2	44.2	44.4	44.7	706	700	699	700	702
	7800	10.2	5.6	5.7	6.9	7.3	999	984	987	983	985	60.2	50.5	49.9	49.6	50.0	996	979	991	975	983
	11700	10.7	5.8	6.3	7.2	7.6	1,295	1,288	1,291	1,287	1,285	63.2	55.8	54.9	55.2	54.6	1,302	1,293	1,291	1,285	1,291
	23400	9.4	7.1	7.6	7.8	7.8	2,056	2,044	2,057	2,048	2,047	68.6	63.6	63.4	62.6	62.6	2,074	2,056	2,052	2,056	2,047
Panel C: $\xi^2 = 0.0100$																					
$n =$	390	20.6	8.5	8.2	12.2	12.2	135	130	130	131	130	49.2	25.4	25.3	29.3	29.3	136	131	130	131	130
	780	14.4	6.0	5.8	7.9	8.4	215	210	211	210	211	41.7	24.2	24.6	27.8	27.3	216	210	210	210	211
	1560	14.3	6.8	7.0	8.8	8.5	342	335	337	334	334	47.4	33.2	33.2	34.4	34.5	340	336	334	335	333
	4680	10.0	6.3	6.0	7.4	7.7	708	698	700	699	695	53.1	42.6	41.7	43.8	43.8	708	698	702	693	701
	7800	9.9	6.2	6.0	7.2	7.2	996	983	984	986	983	57.1	47.3	45.9	47.0	46.9	994	976	980	977	982
	11700	10.1	5.8	6.3	7.2	7.2	1,296	1,287	1,286	1,284	1,288	61.1	52.7	52.5	52.8	52.8	1,297	1,284	1,287	1,287	1,279
	23400	9.9	7.0	7.3	8.3	8.3	2,061	2,040	2,045	2,047	2,039	67.1	60.2	61.5	60.5	60.6	2,069	2,039	2,042	2,050	2,058

Note. We simulate from a model with drift, volatility, infinite-activity jumps and microstructure noise. We test the hypothesis that $\sigma_t = \sigma\sigma_{u,t}$ is a deterministic function of time (induced by diurnal variation) and report rejection rates both under $\mathcal{H}_0$ (size) and $\mathcal{H}_a$ (power). In the latter, $\sigma_t = \sigma_{sv,t}\sigma_{u,t}$ is also time-varying due to a two-factor SV structure. θ is a tuning parameter that is used to compute the pre-averaging window $k_n = [\theta\sqrt{n}]$, φ is the MA(1) coefficient in the noise process, n is the sample size, and ξ^2 controls the magnitude of noise relative to volatility. CLT is for the asymptotic theory from (66), while $z_{\text{wb}\cdot}$ and $t_{\text{wb}\cdot}$ are rejection rates based on the percentile and percentile-t bootstrap test for two choices of the external random variable u. We made 1,000 Monte Carlo trials with 999 bootstrap replica in each simulation. Further details can be found in Section 5.

the t-statistic, illustrating its robustness to market frictions. The only exception to this rule is for $\theta = 1/3$ and $\xi^2 = 0.01$, where a visible drop in the rejection rate under $\mathcal{H}_a$ is noticed, suggesting that narrow pre-averaging is inadequate to counter the microstructure contamination in the presence of an (abnormally) large noise. As expected, the block size b_n increases monotonically with n.

Now, about replacing the latent diurnal volatility $\sigma_{u,t}$ with our statistic $\hat{\sigma}_{u,t}$ from (31), we observe from, say, Table 1 and 3 that it does not modify the properties of the t-statistic much. In particular, the rejection rates are only marginally higher, as was to be expected, because $\hat{\sigma}_{u,t}$ is an estimator, which has an inherent sampling error attached to it. This means it does not completely control for the true seasonal pattern—to the extend $\hat{\sigma}_{u,t}$ deviates from $\sigma_{u,t}$—leaving the rescaled high-frequency log-returns $\Delta_i^n Y^d$ modestly heteroscedastic even under the null. Still, the effect is rather benign, which is remarkable given our naive configuration of $\hat{\sigma}_{u,t}$ with a relatively small choice of m even for larger sample sizes n. We therefore concentrate on the feasible results presented in Table 3 – 4 going forward.

On the other hand, by comparing these two tables we note the pre-averaging window itself, via θ, has a more significant impact on the test, though mostly in small samples. We comment further on that below.

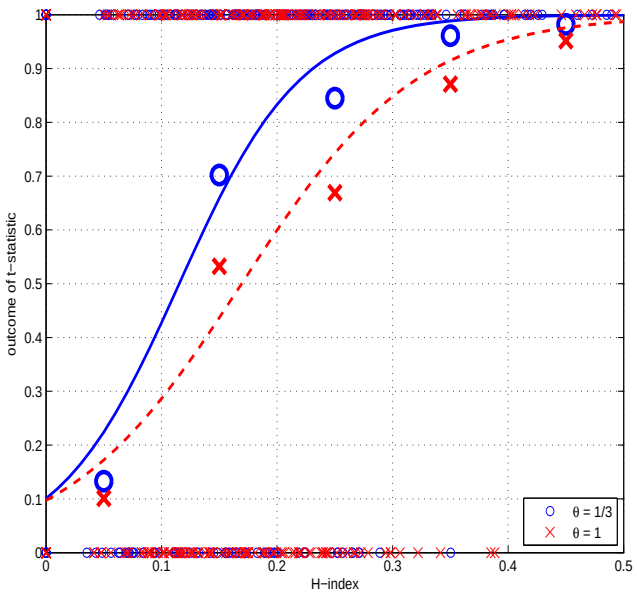
Turning to the analysis of the rejection rates under $\mathcal{H}_0$ of deterministic volatility (size), the tables show the test is oversized. In particular, the CLT-based approach has a pronounced distortion in finite samples, starting at about 25.3% (20.5%) for $\theta = 1/3$ ($\theta = 1$). This is more than five (four) times larger than the nominal level. These rates improve and decline towards 5% as n increases, but remain notably elevated even in fairly large samples.

In contrast, the bootstrap-based approaches are much less biased relative to inference with the asymptotic critical value. The refinement brought about by bootstrapping is often substantial, when the sample size is limited, and the rejection rates are closer to the significance level across the board, albeit they are also mildly inflated initially. The percentile approach appears to possess better size properties compared to the percentile-t, and it settles around 5 – 6% fast. As noted above, the former procedure has the added advantage that it does not require the user to input a—potentially imprecise—estimate of Σ. This also helps to make it slightly less computationally intensive, so as a practical choice we advocate the percentile approach. It is interesting to see that the difference between the two external random variables, in terms of controlling size, is negligible, perhaps with a very weak preference for the one based on the discrete two-point distribution. In the empirical application below, we base our investigation on z_{wb2}.

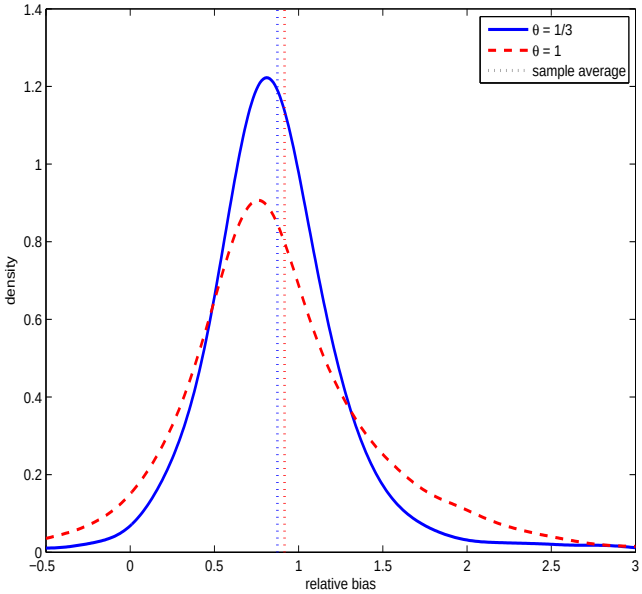
Next, we look at the simulation results under $\mathcal{H}_a$ with stochastic volatility (power). The power exhibited by the various tests is not overwhelming for small n, but it improves steadily towards 100% as n grows large. Still, it stays somewhat less than unity even for $n = 23,400$. It appears the CLT-based test has good power, but this is largely due to the sheer amount of Type I errors committed with this statistic. We observe a notable drop in the rejection rates going from Table 3 (with $\theta = 1/3$) vis-à-vis to Table 4 (with $\theta = 1$), caused by the heavy increase in pre-averaging.

Figure 4: Simulation properties of t-statistic and H-index.

Panel A: assessment of power.



Panel B: relative bias of $\widehat{\text{H}}$-index.



Note. In Panel A, we create an indicator variable I, which equals to one if the t-statistic (based on z_{wb2}) is significant at the 5% nominal level, zero otherwise. We plot it against the true H-index from (92) (small symbol). The curve is from a logistic regression between the two. Also shown are local averages of I (large symbol). In Panel B, we plot the distribution of $\widehat{\text{H}}$-index—defined in (93)—scaled by the H-index. Throughout, the setting is $n = 23,400$, $\xi^2 = 0.001$ and $\varphi = -0.5$.

While this generally renders our testing procedures more resilient to the detrimental effects of microstructure noise, it also smooths out the underlying volatility path and thereby diminishes our ability to uncover genuine heteroscedasticity in the data. It thus highlights a crucial trade-off in practice in terms of selecting θ.

The above suggests our test is not always powerful enough to pick up variation in $\sigma_{sv,t}$. There are several possible explanations of this finding. Firstly, the choice of the tuning parameter θ has a significant impact, as we highlighted above and inspect further below. Secondly, the problem is not trivial. It may just be hard to detect fluctuations in $\sigma_{sv,t}$ from noisy high-frequency data, leaving the jump distortion aside. Thirdly, and although $\sigma_{sv,t}$ is time-varying under the alternative, it may be so persistent that its sample path—which of course differs between simulations—moves about so little (in relative terms) that it appears essentially homoscedastic on an intraday time frame. While this feature partially justifies regarding $\sigma_{sv,t}$ as "almost constant," it also makes it hard for the test to discriminate $\mathcal{H}_a$ from natural sampling variation under $\mathcal{H}_0$ (which it rightfully should do here), at least for the simulated sample sizes.

To shed light on these aspects, we compute:

$$\text{H-index} = 1 - \frac{\left(\int_0^1 \sigma_{sv,s}^2 \mathrm{d}s\right)^2}{\int_0^1 \sigma_{sv,s}^4 \mathrm{d}s}. \tag{92}$$

The H-index compares the square of integrated variance to the integrated quarticity of X^d. It

has the intuitive interpretation that it describes how much $\sigma_{sv,t}$ deviates from $\mathcal{H}_0$ in percent, see, e.g., Podolskij and Wasmuth (2013).[25] We note that H-index $\in [0,1]$ by construction and that it equals zero if and only if $\sigma_{sv,t} = \sigma$ is constant. Strictly positive values imply $\sigma_{sv,t}$ is to some extent time-varying (not necessarily random, though). The H-index is therefore a natural measure of heteroscedasticity in our framework.[26]

The two-factor stochastic volatility process used in this paper has an average H-index value of about 0.20 (based on a large number of paths drawn from the model). It falls below 0.10 20% of the times, while it is rarely smaller than 0.05.

In Panel A of Figure 4, we report the outcome of the t-statistic both for the set of experiments with deterministic and stochastic volatility. We define an indicator variable I, which takes the value one if $\mathcal{H}_0$ was rejected (on the basis of z_{wb2}), and zero otherwise. The figure is a scatter plot of I versus the H-index. The fitted line originates from a logit regression of I on the H-index, which can be interpreted as the power of the test, conditional on H-index. As expected, the tendency to discard $\mathcal{H}_0$ is an increasing function of the H-index. When the deviation from the null is 0.15, the t-statistic is significant about two-thirds of the time for $\theta = 1/3$, which falls down to 50% for $\theta = 1$. Meanwhile, an H-index above 0.35 ($\theta = 1/3$) – 0.45 ($\theta = 1$) implies it more or less always lies in the rejection region. It thus requires rather convincing evidence against the null to firmly reject it, more so for $\theta = 1$.

In practice, we estimate the H-index based on

$$\hat{\text{H}}\text{-index} = 1 - \frac{\widehat{\text{IV}}^2}{\widehat{\text{IQ}}}, \tag{93}$$

where

$$\widehat{\text{IV}} = \frac{1}{\theta\psi_2^n} B\check{V}(1,1)^n - \frac{\psi_1^n}{\theta^2\psi_2^n}\hat{\rho}^2, \qquad \widehat{\text{IQ}} = \frac{1}{(\theta\psi_2^n)^2} B\check{V}(2,2)^n - 2\frac{\psi_1^n}{\theta^2\psi_2^n}\hat{\rho}^2\widehat{\text{IV}} - \left(\frac{\psi_1^n}{\theta^2\psi_2^n}\right)^2\hat{\rho}^4, \tag{94}$$

and

$$\hat{\rho}^2 = \hat{\rho}(0) + 2\sum_{k=1}^{q}\hat{\rho}(k) \tag{95}$$

estimates the asymptotic bias in $B\check{V}(l,r)^n$ in the presence of q-dependent measurement error (e.g., Hautsch and Podolskij, 2013, Lemma 2),

$$\hat{\rho}(k) = -\sum_{j=1}^{q-k+1} j\hat{\gamma}(k+j), \qquad \text{and} \qquad \hat{\gamma}(k) = \frac{1}{n-k}\sum_{i=1}^{n-k}\Delta_i^n Y^d \Delta_{i+k}^n Y^d, \tag{96}$$

[25] The statistic has been used in earlier work to test for the parametric form of volatility (e.g., Dette, Podolskij, and Vetter, 2006; Vetter and Dette, 2012). In contrast to our paper, the former operate with continuous X. Moreover, the ratio appears—sometimes in a different format—in other strands of the literature, for example asymptotic variance reduction (e.g., Clinet and Potiron, 2017), estimation of integrated variation (e.g., Andersen, Dobrev, and Schaumburg, 2014; Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008; Xiu, 2010), or in the context of jump-testing (e.g., Barndorff-Nielsen and Shephard, 2006; Kolokolov and Renò, 2016).

[26] While it is possible to base the t-statistic on the H-index by transforming the CLT in (66) via the delta rule, we refrain from doing so due to the severe amount of time it takes to run the code.

for $k = 0, \dots, q+1$.

In Panel B of Figure 4, we plot the distribution of the relative bias $\widehat{\text{H}}$-index/H-index as a function of θ across simulations under the alternative. Note that an unbiased estimator has the distribution centered at one. As apparent, $\widehat{\text{H}}$-index is slightly downward biased both for $\theta = 1/3$ and $\theta = 1$, which is mainly caused by a modest underestimation of IQ.[27] This leads to conservative statements about the true level of heteroscedasticity in the data, thereby reducing the rejection rate. The distribution is more dispersed and has a higher probability of being close to zero or even outright negative, when $\theta = 1$. This, in part, can help to explain why the simulated power is smaller for $\theta = 1$.

Overall, our noise- and jump-robust test of heteroscedasticity in diurnally-corrected diffusive volatility implemented via the bootstrap percentile-approach has good properties. In contrast to the CLT-based version of the test, it is almost unbiased, also for very small values of n, while it has decent—albeit not perfect—power under the presence of stochastic volatility.

6 Empirical application

We apply our framework to a large cross-sectional panel of US equity high-frequency data. It includes the 30 stocks of the Dow Jones Industrial Average—following the update of its constituent list on March 18, 2015—and the SPDR S&P 500 trust. The latter is an ETF with a price of about 1/10 the cash market value of the S&P 500 index. The sample period is January 4, 2010 through December 31, 2013 for a total of $T = 1,006$ official exchange trading days. Table 5 presents a list of ticker symbols along with a few summary statistics.[28]

The data were extracted from the TAQ database and comprise a complete transaction record for each stock. They were cleaned with the algorithm developed by Christensen, Oomen, and Podolskij (2014), who build on earlier work of Brownless and Gallo (2006) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). It is a standard way of preparing data for analysis in the high-frequency volatility literature.

To compute $\hat{\sigma}_{u,t}$, we create an equidistant log-price series for each asset pre-ticked to a 5-second resolution, i.e. $n = 4,680$. We then set $m = 780$—or $n/m = 6$—to retrieve a local estimate $\hat{\sigma}_{u,t}$ that covers a 30-second interval.[29] It ensures that we recover a detailed view of the diurnality in volatility, while still being able to purge the associated noise with decent accuracy. On each block, we bias-correct with the robust estimator in (39) using $q = 3$.[30] This is motivated by Panel A

[27] There is also an attenuation effect induced by the non-linear transformation of $\widehat{\text{IV}}$.

[28] Notice that the variance of the noise, as captured by $\hat{\xi}^2$, is generally smaller than what we assumed in the simulations. This is consistent with the notion that the noise has decreased over time.

[29] To estimate $\sigma_{u,t}$, we further delete a few outliers from the sample. Firstly, the Flash Crash of May 6, 2010. Secondly, for each 30-second interval we remove the top 1% of data for each stock—as measured by $|\Delta_i^m Y|$—to filter out observations typically associated with idiosyncratic news announcements. These events exert an unduly influence on the estimates due to our relatively small value of T. In a larger sample this should not be necessary.

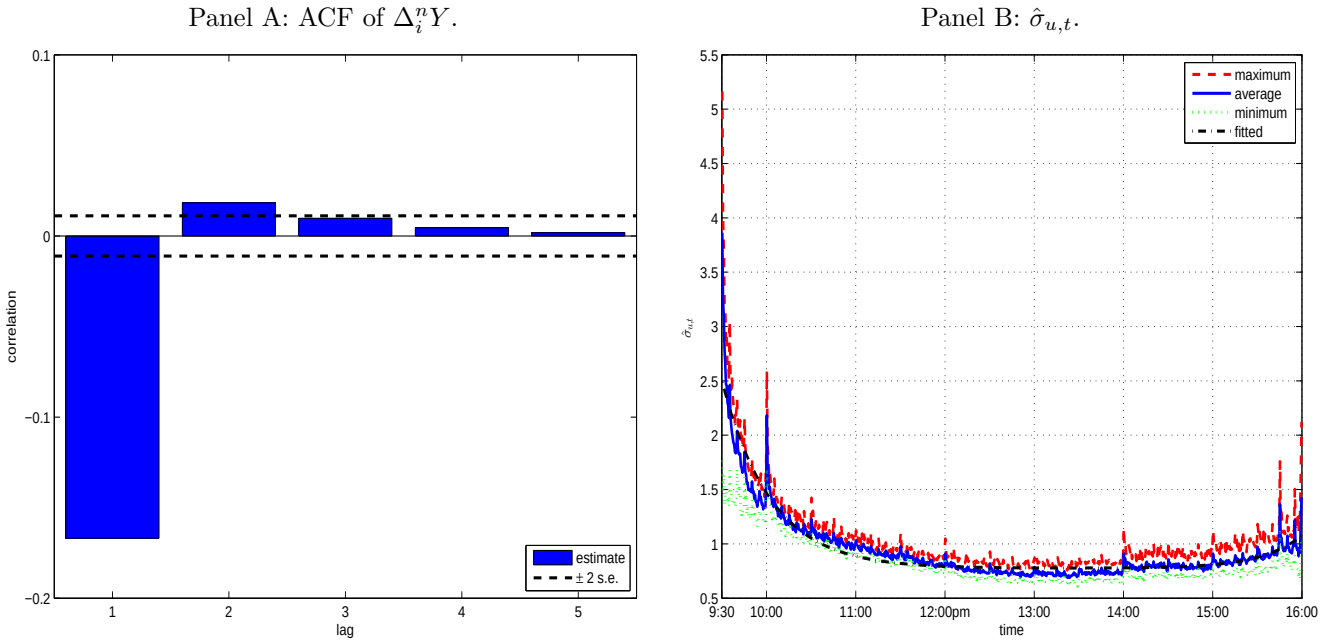
[30] ω^2 is sometimes estimated to be negative, due to the sampling distribution of $\hat{\omega}^2$. This cannot be true, of course,

Table 5: Descriptive statistics of TAQ high-frequency data.

					$\theta = 1/3$					$\theta = 1$				
						$\#\mathcal{Z}^n > p^*_{0.95}$					$\#\mathcal{Z}^n > p^*_{0.95}$			
ticker	n	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\xi}^2 \times 10^4$	k_n	z_{wb2}	z^d_{wb2}	Ĥ-index	Ĥ-index^d	k_n	z_{wb2}	z^d_{wb2}	Ĥ-index	Ĥ-index^d
AAPL	14,576	18.9	-0.12	0.12	40	0.89	0.57	0.23	0.12	120	0.58	0.17	0.18	0.06
AXP	6,843	19.5	-0.07	0.29	27	0.86	0.59	0.23	0.15	82	0.52	0.20	0.17	0.09
BA	6,486	18.7	-0.10	0.39	27	0.93	0.65	0.27	0.17	80	0.63	0.26	0.23	0.12
CAT	8,510	21.7	-0.09	0.18	30	0.85	0.42	0.22	0.11	91	0.55	0.14	0.18	0.05
CSCO	11,884	16.5	-0.38	1.74	36	0.99	0.96	0.41	0.35	108	0.81	0.61	0.29	0.21
CVX	8,568	16.1	-0.05	0.29	31	0.78	0.48	0.20	0.12	92	0.45	0.17	0.15	0.07
DD	6,521	18.4	-0.11	0.44	27	0.84	0.58	0.24	0.15	80	0.54	0.20	0.18	0.09
DIS	7,422	17.1	-0.14	0.34	29	0.90	0.68	0.26	0.18	86	0.53	0.24	0.18	0.09
GE	12,990	16.5	-0.41	2.02	38	0.99	0.98	0.42	0.36	114	0.83	0.66	0.30	0.22
GS	7,940	22.2	-0.10	0.22	29	0.91	0.57	0.24	0.14	88	0.58	0.19	0.19	0.08
HD	7,798	17.0	-0.16	0.27	29	0.92	0.76	0.28	0.19	88	0.62	0.28	0.21	0.11
IBM	7,193	13.9	-0.14	0.38	28	0.88	0.61	0.23	0.15	84	0.52	0.21	0.18	0.10
INTC	12,515	17.3	-0.37	1.24	37	1.00	0.97	0.39	0.32	112	0.82	0.54	0.29	0.18
JNJ	8,766	11.3	-0.18	0.32	31	0.93	0.81	0.28	0.21	93	0.61	0.29	0.20	0.12
JPM	12,141	21.9	-0.14	0.11	37	0.93	0.74	0.26	0.16	110	0.61	0.21	0.19	0.08
KO	7,991	12.2	-0.21	0.46	30	0.93	0.81	0.30	0.22	89	0.61	0.30	0.21	0.12
MCD	7,192	12.2	-0.13	0.28	28	0.89	0.66	0.26	0.16	84	0.57	0.26	0.20	0.11
MMM	5,466	14.9	-0.08	0.64	24	0.85	0.53	0.24	0.15	73	0.53	0.24	0.19	0.10
MRK	8,455	14.8	-0.25	0.49	31	0.97	0.88	0.33	0.26	92	0.68	0.37	0.23	0.14
MSFT	12,814	16.2	-0.34	0.72	38	0.99	0.96	0.37	0.31	113	0.79	0.56	0.27	0.18
NKE	4,847	18.0	-0.07	0.35	23	0.89	0.60	0.27	0.16	69	0.62	0.23	0.22	0.11
PFE	11,253	14.6	-0.38	2.12	35	0.99	0.98	0.43	0.38	106	0.85	0.67	0.31	0.23
PG	7,967	11.9	-0.16	0.28	30	0.92	0.77	0.27	0.20	89	0.55	0.30	0.19	0.12
TRV	4,779	14.9	-0.11	0.88	23	0.89	0.61	0.26	0.17	68	0.55	0.26	0.19	0.12
UNH	6,730	20.6	-0.09	0.28	27	0.96	0.74	0.31	0.19	81	0.69	0.28	0.25	0.12
UTX	5,938	16.1	-0.09	0.47	25	0.89	0.63	0.26	0.16	76	0.59	0.26	0.21	0.11
V	6,213	19.3	-0.12	0.53	26	0.94	0.73	0.28	0.19	78	0.68	0.31	0.24	0.14
VZ	9,191	13.4	-0.29	0.68	32	0.98	0.93	0.36	0.28	95	0.73	0.46	0.25	0.16
WMT	8,085	12.4	-0.17	0.25	30	0.95	0.84	0.29	0.21	89	0.62	0.26	0.21	0.11
XOM	10,693	14.7	-0.07	0.17	34	0.81	0.59	0.19	0.14	103	0.44	0.17	0.14	0.07
SPY	18,154	10.4	-0.05	0.10	45	0.92	0.80	0.21	0.16	135	0.49	0.23	0.14	0.08

Note. This table reports descriptive statistics for our TAQ high-frequency data computed daily and averaged over time. The sample covers January 4, 2010 through December 31, 2013 for a total of $T = 1,006$ days. n is the number of transaction data available after filtering, $\hat{\sigma} = \sqrt{256 \times \widehat{\text{IV}}}$ is an annualized jump-robust measure of volatility based on (94) (with $\theta = 1/3$), $\hat{\rho}_1$ is the first-order autocorrelation of $\Delta_i^n Y$, and $\hat{\xi}^2$ is the noise level $\times 10^4$. k_n is the pre-averaging window, while $\#\mathcal{Z}^n > p^*_{0.95}$ is the fraction of t-statistics for testing $\mathcal{H}_0$ larger than the 95%-quantile of the bootstrap distribution of $\mathcal{Z}^{n*}$ (based on z_{wb2}). Ĥ-index is the heteroscedasticity measure defined in (93). The latter three are computed for both $\theta = 1/3$ and $\theta = 1$. A superscript d refers to the average value of a statistic based on $\Delta_i^n Y^d$, where the rescaling is implemented via $\hat{\sigma}_{u,t}$ from (31).

Figure 5: Properties of equity high-frequency data.



Note. In Panel A, we plot the ACF of our TAQ high-frequency equity data averaged across assets and over time. In Panel B, we present our estimator $\hat{\sigma}_{u,t}$ of within-day volatility. The cross-sectional average is reported, along with the highest and smallest point estimate. As a comparison, we also fit the parametric form of diurnal variation in (85) via non-linear least squares.

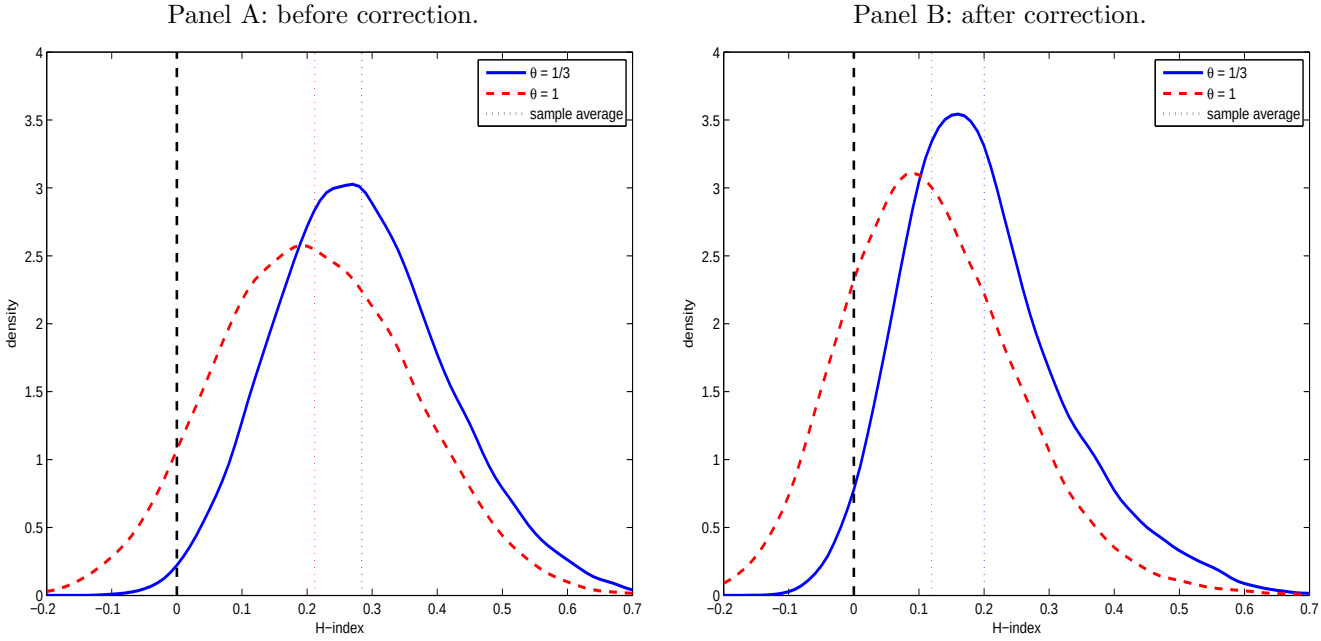
in Figure 5, which reports the autocorrelation function (ACF) of $\Delta_i^n Y$. As shown, while the first few autocorrelations are significant, the ACF dies out fast and is generally insignificant after lag three and negligible beyond lag five (not shown in figure), so this choice of q suffices to capture the observed serial dependence in the noise.

The cross-sectional average of $\hat{\sigma}_{u,t}$ is reported along with the minimum and maximum value in Panel B of Figure 5. We note $\hat{\sigma}_{u,t}$ features the reverted J-shape as reported in prior work, but also that it is very rough with sharp increases around pre-scheduled macroeconomic announcements (e.g., at 10:00am or 2:00pm). The latter is consistent with empirical findings in, e.g., Todorov and Tauchen (2011), who note that jumps in volatility are strongly correlated with large moves in market prices.[31] There are also some notable spikes in within-day volatility prior to the close of the exchange. Overall, diurnal variation is remarkably constant across assets, as gauged by the maximum and minimum value. To assess the parametric model used in the simulations, we estimate the parameters of (85) via non-linear least squares based on the cross-sectional average. The fitted equation $\sigma_{u,t} = 0.78 + 1.71e^{11.85t} + 0.30e^{14.17(1-t)}$ is a good approximation to our nonparametric estimates, although it does not track the sharp initial decay in early trading.

and we therefore truncate $\hat{\omega}^2$ at zero throughout the paper. This happens more often if the noise is really small relative to the underlying volatility of the asset, as demonstrated by the index-tracker SPY in Figure 1.

[31] It also suggests that $\sigma_{u,t}$ may not be as smooth as stipulated by Assumption (D3). Note, however, that although volatility peaks at the announcement, it does not necessarily jump. In our data it actually starts to increase around 1-minute to 30-seconds *prior* to the time, where the numbers are officially slated for release. This is consistent with the findings of Jiang, Lo, and Verdelhan (2011) from the U. S. Treasury market.

Figure 6: The distribution of the $\hat{\text{H}}$-index.



Note. We plot the cross-sectional distribution of the $\hat{\text{H}}$-index before and after diurnal correction with our nonparametric estimator from (31), also shown in Panel B of Figure 5. The dashed vertical line at zero indicates the theoretical lower bound.

We then construct the deflated log-returns and compute the test for each stock and day in the sample. In the right-hand side of Table 5, we report the average rejection rate of $\mathcal{H}_0$ and associated H-index measurement. As above, we base the investigation on $B\check{V}(2,2)^n$ and $B\check{V}(1,1)^n$ with $\theta = 1/3$ and $\theta = 1$. The bootstrap percentile approach is applied to evaluate the significance of our t-statistic, i.e. z_{wb2}. As a comparison, we also retrieve the corresponding results from the raw data prior to diurnal correction.

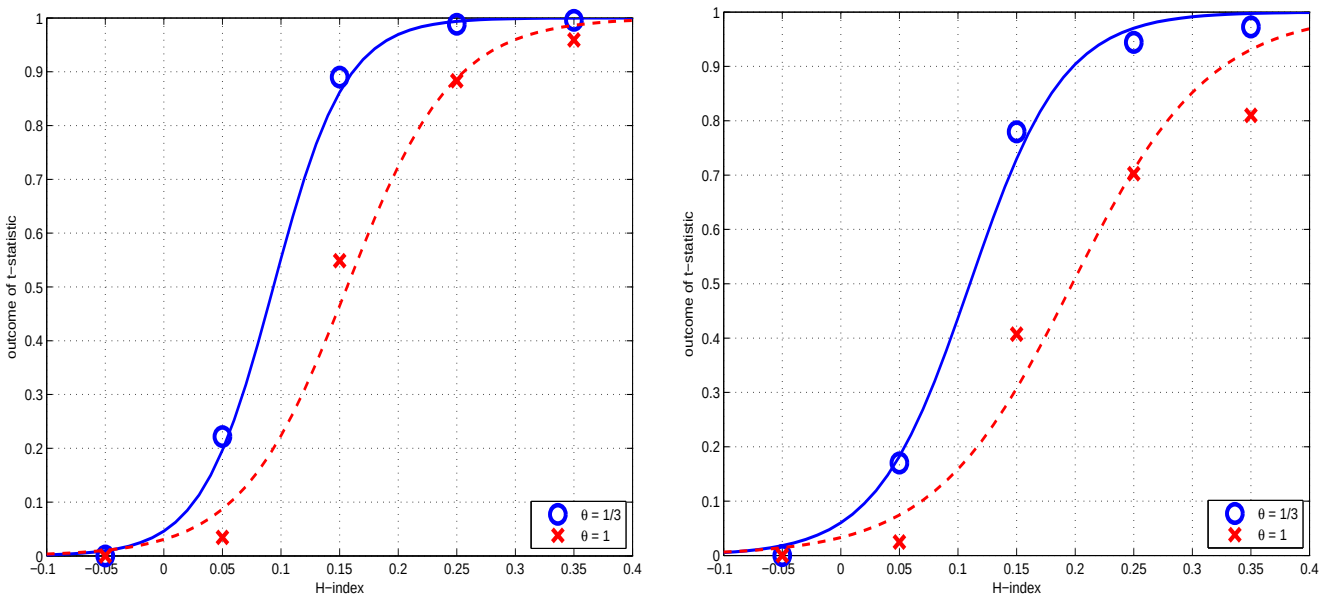
Looking at Table 5, we observe that $\mathcal{H}_0$ is discarded more than half of the times for almost every single stock, irrespective of θ, if there is no seasonal adjustment.[32] The levels are highest for $\theta = 1/3$, which is consistent with the increased power if θ is small, as uncovered in the simulation section. On the other hand, there is also some evidence of the caveat raised in Remark 3. That is, the pre-averaging estimator is affected harder by residual microstructure noise if θ is small, so that here the test is prone to discredit $\mathcal{H}_0$ in the presence of a general form of heteroscedasticity in the variance of the noise. Indeed, there is some tendency for stocks with very negative first-order return autocorrelation and large levels of noise—as measured by $\hat{\xi}^2$— to reject more frequently.

If we control for diurnal variation, the rejection rate drops about by 25% (for $\theta = 1/3$) to almost 50% (for $\theta = 1$). This implies the diurnal pattern is a first-order effect, which captures a large fraction of variation in intraday volatility. However, as readily seen, important and potent sources

[32]We add that due to the multiple testing, roughly five percentage points of these (or even a bit more as the test was found to be mildly oversized) can potentially be attributed to false positives, since we are testing at a significance level $\alpha = 0.05$. To control the family-wise error rate, a standard Bonferroni-type correction can be used (e.g., Andersen, Bollerslev, and Dobrev, 2007).

Figure 7: Empirical properties of t-statistic and $\hat{\text{H}}$-index.

Panel A: before correction. Panel B: after correction.



Note. We create an indicator variable I, which equals to one if the t-statistic is significant at the 5% nominal level, zero otherwise. We plot I against the $\hat{\text{H}}$-index before and after diurnal correction (i.e., based on ($\hat{\text{H}}$-index, z_{wb2}) or ($\hat{\text{H}}$-index^d, z^d_{wb2})) and as a function of θ. The curve is from a logistic regression between the two. We also show local averages of I.

of heteroscedasticity remain present in the data.

This story is corroborated by the $\hat{\text{H}}$-index, for which we also plot the cross-sectional distribution before and after diurnal correction in Figure 6. The distribution shifts to the left and displays less sampling variation after diurnal correction, while we again notice a slight increase in the dispersion by moving θ up, although the effect is weak. As measured by the cross-sectional average shown in Panel B of the figure, the strength of residual heteroscedasticity present in $\Delta_i^n Y^d$ is broadly comparable to that of the two-factor stochastic volatility model from Section 5 for $\theta = 1/3$, while it is somewhat less for $\theta = 1$.

At last, in Figure 7 we model the empirical rejection rate of the t-statistic as a function of the $\hat{\text{H}}$-index. The message of the logit fit is consistent with the simulations. It takes a relatively high reading of $\hat{\text{H}}$-index to confidently reject $\mathcal{H}_0$. Note that the rejection rates are (at least crudely) close if $\hat{\text{H}}$-index $\simeq$ $\hat{\text{H}}$-index^d, as can be gauged by comparing Panel A and B. The intuitive explanation is that the power of the test depends only on the level of heteroscedasticity, which is captured by the H-index, and not as such on whether one has diurnally-corrected or not.

7 Conclusion

In this paper, we study a new approach to determine if changes in intraday spot volatility of a discretely sampled noisy jump-diffusion model can be attributed solely to a deterministic cyclical component (i.e., the so-called U- or reverse J-shape) against an alternative of further variation

induced by a stochastic process.

We propose to construct a test of this hypothesis from an asset return series, which has been deflated by the diurnal component and, as such, is homoscedastic under the null. The t-statistic diverges to infinity, if the deflated return series is heteroscedastic, and it has a standard normal distribution otherwise. To get a feasible test, we develop a—surprisingly robust—nonparametric estimator of unobserved diurnal volatility, which (in contrast to the test itself) is computed directly from noisy high-frequency data without pre-averaging or jump-truncation. It requires only a trivial bias-correction to eliminate the noise variance. Our estimator is consistent and has a sampling error, which is of small enough order that replacing the true diurnal factor with it does not alter the asymptotic theory.

We inspect the properties of the test in a Monte Carlo simulation. We note the theory-based version has gross size distortions in the presence of infinite-activity price jumps, thus motivating a bootstrap. We validate the bootstrap and confirm it helps to improve inference by making the test almost correctly sized. The test also has acceptable power, but can fail to reject the null even in large samples, if a wide pre-averaging window is applied. The estimation of the diurnal factor has a limited impact, but it raises the rejection rate slightly.

We implement our nonparametric estimator of diurnal variance and test of heteroscedasticity on real high-frequency data. The diurnal pattern explains a sizable portion of within-day variation in the volatility, as inferred by the notable drop in the rejection rate of the test and the reduction in the H-index—a descriptive statistic that measures the strength of time-varying volatility. It suggests that once we control for diurnal variation in practice, the rescaled log-returns are much closer to homoscedastic, but important sources of variation remain present in the data. So the definite answer to the title of the paper appears to be "no." The diurnal pattern does not explain *all* intraday variation in volatility, but it does capture a rather significant portion of it.

A The explicit form of Σ

In Section 2, we show that our proposed estimator $\check{\Sigma}^n$ is consistent for the asymptotic covariance matrix of $n^{1/4}\big(BV(Y^d, l_1, r_1)^n, BV(Y^d, l_2, r_2)^n\big)^\intercal$, i.e. Σ appearing in (22), Theorem B.1 and Theorem 4.1. We also prove a corresponding result for the bootstrap version, $\check{\Sigma}^{n*}$, in Section 4. In this short appendix, we derive an explicit expression for Σ, which was not put in the main text. We follow Podolskij and Vetter (2009a) by first defining:

$$h_{ij}(a, b, c) = \text{cov}\Big(|H_1|^{l_i}|H_2|^{r_i}, |H_3|^{l_j}|H_4|^{r_j}\Big),$$

where a is a real number, b and c are a two- and four-dimensional vector. Moreover, $(H_1, \ldots, H_4)$ follows a multivariate normal distribution with:

1. $E(H_l) = 0$ and $\text{var}(H_l) = b_1 a^2 + b_2 \omega^2$,
2. $H_1 \perp H_2$, $H_1 \perp H_4$, and $H_3 \perp H_4$,
3. $\text{cov}(H_1, H_3) = \text{cov}(H_2, H_4) = c_1 a^2 + c_2 omega^2$ and $\text{cov}(H_2, H_3) = c_3 a^2 + c_3 \omega^2$.

We set $t = \left(\frac{1}{\theta}\psi_1, \theta\psi_2\right)$ and define:

$$f_1(s) = \frac{1}{\theta}\phi_1(s), \quad f_2(s) = \theta\phi_2(s), \quad f_3(s) = \theta\phi_3(s), \quad f_4(s) = \frac{1}{\theta}\phi_4(s),$$

for $s \in [0, 2]$, where

$$\phi_1(s) = \int_0^{1-s} g'(u)g'(u+s)\text{d}u, \qquad \phi_2(s) = \int_0^{1-s} g(u)g(u+s)\text{d}u,$$

$$\phi_3(s) = \int_0^{2-s} g'(u)g'(u+s-1)\text{d}u \qquad \text{and} \qquad \phi_4(s) = \int_0^{2-s} g(u)g(u+s-1)\text{d}u.$$

We note that both f_1 and f_2 are 0 for $s \in [1, 2]$, according to the assumptions imposed on g. We next let $f(s) = \big(f_1(s), f_2(s), f_3(s), f_4(s)\big)^\intercal$. At last, we get that

$$\begin{aligned}\Sigma &= \left(\Sigma_{ij}^{l_1,r_1,l_2,r_2}\right)_{1\leq i,j\leq 2} \\ &= \int_0^1 \begin{pmatrix} w_{11}^{l_1,r_1,l_2,r_2} & w_{12}^{l_1,r_1,l_2,r_2} \\ w_{21}^{l_1,r_1,l_2,r_2} & w_{22}^{l_1,r_1,l_2,r_2} \end{pmatrix}(\sigma_{sv,u})\text{d}u,\end{aligned}$$

where

$$w_{ij}^{l_1,r_1,l_2,r_2}(\sigma_{sv,u}) = 2\theta\int_0^2 h_{ij}\big(\sigma_{sv,u}, t, f(s)\big)\text{d}s.$$

B Proofs

In this appendix, K denotes a generic constant, which changes from line to line. Also, as in Jacod and Protter (2012), we assume that a, σ, δ and X are bounded. As Jacod, Podolskij, and Vetter (2010) explain, this follows by a standard localization procedure, described in Jacod (2008), and does not lose generality. Formally, we derive our results under the assumption:

Assumption (G): X follows (1) with a and σ are adapted, càdlàg processes such that a, σ, δ and X are bounded, so that for some constant K and nonnegative deterministic function $\tilde{\gamma}$:

$$\|a_t(\omega)\| \leq K,\ \|\sigma_t(\omega)\| \leq K,\ \|X_t(\omega)\| \leq K,\ \|\delta(\omega,t,x)\| \leq \tilde{\gamma}(x) \leq K, \int_{\mathbb{R}} \tilde{\gamma}(x)^{\beta} \lambda(\mathrm{d}x) \leq K.$$

Throughout the appendix, it will be convenient to define the continuous part of X by X' and the discontinuous martingale part by X'', i.e.

$$X'_t = X_0 + \int_0^t a'_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s, \quad X''_t = X_t - X'_t, \tag{97}$$

where, according to the value of β, we set

$$a'_s = \begin{cases} a_s - \left(\delta 1_{\{|\delta| \leq 1\}}\right) \star \underline{\nu}_t, & \text{if} \quad \beta \leq 1 \\ a_s + \left(\delta 1_{\{|\delta| > 1\}}\right) \star \underline{\nu}_t, & \text{if} \quad \beta > 1 \end{cases}.$$

Then, we can write

$$Y_t = Y'_t + Y''_t, \tag{98}$$

where $Y'_t = X'_t + \epsilon_t$ and $Y''_t = X''_t$. As in the main text, if we write $BV(l,r)^n, \check{BV}(l,r)^n$, $B(l,r)^n_i$, $\Delta B(l,r)^n_i$, $\check{y}(l,r)^n_i$, $\check{B}(l,r)^n_i$ or $\check{\Delta} B(l,r)^n_i$, we assume they are defined with respect to Y^d.

Proof of Theorem 2.1. Here, we more or less follow the techniques applied in the proof of Theorem 4.1 in Hounyo (2017). Firstly, we introduce the pre-averaged return $\Delta^n_i \bar{Y}$ computed on the raw unscaled high-frequency returns:

$$\Delta^n_i \bar{Y} = \sum_{j=1}^{k_n - 1} g\left(\frac{j}{k_n}\right) \Delta^n_{i+j-1} Y, \quad i = 1, \ldots, n - k_n + 2.$$

Next, for the associated definitions of $\check{BV}(Y,l,r)^n, \check{BV}(Y',l,r)^n$, together with the maintained assumptions appearing in the main text and if $\sigma_{u,t} = 1$, the central limit theorem in Theorem 3 of Podolskij and Vetter (2009a) implies that, as $n \to \infty$,

$$n^{1/4} \begin{pmatrix} \check{BV}(Y', l_1, r_1)^n - BV(l_1, r_1) \\ \check{BV}(Y', l_2, r_2)^n - BV(l_2, r_2) \end{pmatrix} \xrightarrow{d_s} MN(0, \Sigma). \tag{99}$$

A careful inspection of the proof of this result shows that the stable convergence in (99) remains valid, when the pre-averaged return is given by $\Delta^n_i \bar{Y}^d$ in (17) and $\sigma_t = \sigma_{sv,t} \sigma_{u,t}$. Indeed, the main

ingredient is the weak convergence:

$$n^{1/4}\Delta_i^n\epsilon^d \xrightarrow{d} N\bigg(0, \frac{1}{\theta}\psi_1\omega^2\bigg),$$

and

$$n^{1/4}\Delta_i^n\bar{Y}'^d \overset{a}{\sim} N\bigg(0, \theta\psi_2\sigma^2_{sv,i/n} + \frac{1}{\theta}\psi_1\omega^2\bigg),$$

which follows from (5), as $\sigma_{u,t} > 0$ for all $t \geq 0$ and locally bounded. Thus, it suffices to prove that for any $l, r > 0$,

$$n^{1/4}\big(\check{BV}(Y^d, l, r)^n - \check{BV}(Y'^d, l, r)^n\big) \xrightarrow{p} 0. \tag{100}$$

To show (100), we let $F_u(x) = F(x)1_{\{|x_1|<u\}}1_{\{|x_2|<u\}}$, for some $u > 0$, where $F(x) = |x_1|^l|x_2|^r$ with $x = \begin{pmatrix} x_1 & x_2 \end{pmatrix}^\intercal$. As in the line of thought on page 385 in Jacod and Protter (2012), we can show that for $w_n = \upsilon_n/\sqrt{u_n}$ with $u_n = k_n/n$:

$$|F_{w_n}(x+y) - F_{w_n}(x)| \leq w_n^{-\frac{2}{1-2\omega}}\|x\|^{l+r+\frac{2}{1-2\omega}} + \Big(\big(1 + \|x\|^{l+r}\big)\big(\|y\| \wedge 1 + \|y\|^{l+r} \wedge w_n^{l+r}\big)\Big).$$

Next, let $x = \begin{pmatrix} \Delta_i^n\bar{Y}^d & \Delta_{i+k_n}^n\bar{Y}^d \end{pmatrix}^\intercal/\sqrt{u_n}$ and $y = \begin{pmatrix} \Delta_i^n\bar{Y}''^d & \Delta_{i+k_n}^n\bar{Y}''^d \end{pmatrix}^\intercal/\sqrt{u_n}$. According to (16.4.9) in Jacod and Protter (2012) in conjunction with results in part 3 in the proof of Lemma 16.4.5 in that book, for some $l + r > 0$:

$$E\big(\|x\|^{l+r}\big) \leq K,\ E\big(\|y\| \wedge 1\big) \leq Ku_n^{1-\beta/2}\phi_n \text{ and } E\big(\|y\|^2 \wedge w_n^2\big) \leq Ku_n^{\omega(2-\beta)}\phi_n, \tag{101}$$

where $\phi_n \to 0$ as $n \to 0$. In addition, from (101) and the inequality $(\|y\| \wedge w_n)^p \leq w_n^{p-m}(\|y\| \wedge w_n)^m$, for $0 < m < p$, it is found that

$$E\big(\|y\|^{l+r} \wedge w_n^{l+r}\big) \leq Kw_n^{l+r-2}E\big(\|y\| \wedge w_n\big)^2 \leq Ku_n^{\omega(l+r-\beta)-\frac{1}{2}(l+r-2)}\phi_n, \tag{102}$$

where again $\phi_n \to 0$ as $n \to 0$. Thus, from the above inequalities together with the definition

$$\begin{aligned}
&n^{1/4}\big(\check{BV}(Y^d, l, r)^n - \check{BV}(Y'^d, l, r)^n\big) \\
&= \frac{n^{\frac{l+r-3}{4}}}{\mu_l\mu_r}\sum_{i=1}^{n-2k_n+2}\Big(|\Delta_i^n\bar{Y}^d|^l|\Delta_{i+k_n}^n\bar{Y}^d|^r - |\Delta_i^n\bar{Y}'^d|^l|\Delta_{i+k_n}^n\bar{Y}'^d|^r\Big)1_{\{|\Delta_i^n\bar{Y}^d|<\upsilon_n\}}1_{\{|\Delta_{i+k_n}^n\bar{Y}^d|<\upsilon_n\}},
\end{aligned}$$

it follows that

$$\begin{aligned}
n^{\frac{l+r-3}{4}}\frac{1}{\mu_l\mu_r}\sum_{i=1}^{n-2k_n+2}&E\bigg(\Big|\Big(|\Delta_i^n\bar{Y}^d|^l|\Delta_{i+k_n}^n\bar{Y}^d|^r - |\Delta_i^n\bar{Y}'^d|^l|\Delta_{i+k_n}^n\bar{Y}'^d|^r\Big)1_{\{|\Delta_i^n\bar{Y}^d|<\upsilon_n\}}1_{\{|\Delta_{i+k_n}^n\bar{Y}^d|<\upsilon_n\}}\Big|\bigg) \\
&\leq Kn^{\frac{l+r-3}{4}}n \cdot u_n^{\frac{l+r}{2}}\Big(u_n + u_n^{1-r/2}\phi_n + u_n^{\omega(l+r-\beta)-\frac{1}{2}(l+r-2)}\phi_n\Big) \\
&\leq Kn^{\frac{1}{4}}\Big(n^{-1/2} + n^{\frac{\beta-2}{4}}\phi_n + n^{\frac{(l+r-2)-2\omega(l+r-\beta)}{4}}\phi_n\Big) \\
&\leq K\Big(n^{-1/4} + n^{\frac{(\beta-1)}{4}}\phi_n + n^{\frac{(l+r-1)-2\omega(l+r-\beta)}{4}}\phi_n\Big).
\end{aligned}$$

Thus, if $\beta < 1$ and $\frac{l+r-1}{2(l+r-\beta)} \leq \varpi < 1/2$, then $E\big(|n^{1/4}(\check{BV}(Y^d, l, r)^n - \check{BV}(Y'^d, l, r)^n)|\big) \to 0$ and

therefore $n^{1/4}(\check{BV}(Y^d,l,r)^n - \check{BV}(Y'^d,l,r)^n) \xrightarrow{p} 0$. This completes the proof of Theorem 2.1. ■

Next, we establish the following result (under no jumps) since it will be useful later in the proof of Theorem 4.1.

Theorem B.1 *Let l_1, r_1, l_2 and r_2 be four positive real numbers and X given by*

$$X_t = X_0 + \int_0^t a_s \mathrm{d}s + \int_0^t \sigma_s \mathrm{d}W_s. \tag{103}$$

We define:

$$\hat{\Sigma}^n = \frac{\sqrt{n}}{2b_n} \sum_{i=1}^{N_n - 2b_n + 1} \xi_i \xi_i^\intercal, \tag{104}$$

where $\xi_i \equiv (\Delta B(l_1,r_1)_i^n, \Delta B(l_2,r_2)_i^n)^\intercal$, such that

$$\Delta B(l,r)_j^n = B(Y^d,l,r)_{j+b_n}^n - B(Y^d,l,r)_j^n, \tag{105}$$

with

$$B(l,r)_j^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{b_n} y(l,r)_{i-1+j}^n. \tag{106}$$

Furthermore, we assume (V), (A), and impose the moment condition $E(|\epsilon_t|^s) < \infty$, for some $s > (3 \wedge 2(r_1+l_1) \wedge 2(r_2+l_2))$. If any l_i or r_i is in $(0,1]$, we postulate (V′), otherwise either (V′) or (A′). In addition, suppose that $k_n \to \infty$ as $n \to \infty$ such that (19) holds, and the block size b_n fulfills (44) for some $1/2 < \delta_1 < 2/3$. Then, as $n \to \infty$,

$$\hat{\Sigma}^n \xrightarrow{p} \Sigma, \tag{107}$$

where Σ is defined in Appendix A.

Proof of Theorem B.1. Here, recall that X follows (103) and note that given (104), we can rewrite $\hat{\Sigma}^n$ as follows:

$$\hat{\Sigma}^n = \frac{1}{b_n} \sum_{m=1}^{b_n} \hat{\Sigma}_m^n, \tag{108}$$

where

$$\hat{\Sigma}_m^n = \frac{\sqrt{n}}{2} \sum_{k=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \xi_{kb_n+m} \xi_{kb_n+m}^\intercal = \left(\hat{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n} \right)_{1 \le i,j \le 2}. \tag{109}$$

Thus, it suffices that $\hat{\Sigma}_m^n \xrightarrow{p} \Sigma$, uniformly in m. Thus, the proof is reduced to showing that

$$\operatorname*{p-lim}_{n \to \infty} \hat{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n} = \Sigma_{ij}^{l_1,r_1,l_2,r_2}, \quad 1 \le i,j \le 2, \tag{110}$$

uniformly in m. Note that we can rewrite $\hat{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}$ as

$$\hat{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n} = \frac{\sqrt{n}}{2} \sum_{k=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \Delta B(Y^d,l_i,r_i)_{kb_n+m}^n \Delta B(Y^d,l_j,r_j)_{kb_n+m}^n.$$

Then, given the definition of $\Delta B(Y^d, l, r)_m^n$ given in (105), by adding and subtracting appropriately, it follows that

$$\begin{aligned}
\hat{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n} &= \frac{\sqrt{n}}{2} \sum_{k=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \begin{pmatrix} 2B(Y^d, l_i, r_i)_{(k+1)b_n+m}^n B(Y^d, l_j, r_j)_{(k+1)b_n+m}^n \\ -B(Y^d, l_i, r_i)_{(k+1)b_n+m}^n B(Y^d, l_j, r_j)_{kb_n+m}^n \\ -B(Y^d, l_i, r_i)_{kb_n+m}^n B(Y^d, l_j, r_j)_{(k+1)b_n+m}^n \end{pmatrix} \\
&+ \frac{\sqrt{n}}{2} \begin{pmatrix} B(Y^d, l_i, r_i)_m^n B(Y^d, l_j, r_j)_m^n \\ +B(Y^d, l_i, r_i)_{(\lfloor \frac{N_n}{b_n} \rfloor - 1)b_n+m}^n B(Y^d, l_j, r_j)_{(\lfloor \frac{N_n}{b_n} \rfloor - 1)b_n+m}^n \\ -B(Y^d, l_i, r_i)_{(\lfloor \frac{N_n}{b_n} \rfloor - 2)b_n+m}^n B(Y^d, l_j, r_j)_{(\lfloor \frac{N_n}{b_n} \rfloor - 1)b_n+m}^n \\ -B(Y^d, l_i, r_i)_{(\lfloor \frac{N_n}{b_n} \rfloor - 1)b_n+m}^n B(Y^d, l_j, r_j)_{(\lfloor \frac{N_n}{b_n} \rfloor - 2)b_n+m}^n \end{pmatrix} \\
&= M_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) + R_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d),
\end{aligned}$$

where the remainder term

$$R_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) = O_p\Big(n^{-\frac{3}{2}} b_n^2\Big) = o_p(1),$$

uniformly in m, so long as $\delta_1 < 3/4$, where we apply the definition of $B(Y^d, l, r)_m^n$ in (106), the Cauchy-Schwartz inequality, and the fact that $E\big(|\Delta_i^n \bar{Y}^d|^l\big) \leq K n^{-l/4}$ (cf., Lemma 1 of Podolskij and Vetter, 2010). Next, we show the main term is such that

$$\underset{n\to\infty}{\text{p-lim}}\ M_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) = \Sigma_{ij}, \quad 1 \leq i, j \leq 2, \tag{111}$$

uniformly in m. We prove the result for the following unsymmetrized estimator:

$$\tilde{M}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) = \sqrt{n} \sum_{k=1}^{\lfloor \frac{N_n}{b_n} \rfloor - 1} \begin{pmatrix} B(Y^d, l_i, r_i)_{kb_n+m}^n B(Y^d, l_j, r_j)_{kb_n+m}^n \\ -B(Y^d, l_i, r_i)_{kb_n+m}^n B(Y^d, l_j, r_j)_{(k-1)b_n+m}^n \end{pmatrix}. \tag{112}$$

We introduce two approximations of $B(Y^d, l, r)_{jb_n+m}^n$:

$$\tilde{B}(Y^d, l, r)_{jb_n+m}^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{b_n} \tilde{y}(l, r)_{i-1+jb_n+m}^n,$$

$$\bar{B}(Y^d, l, r)_{jb_n+m}^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{b_n} \tilde{y}(l, r)_{i-1+(j-1)b_n+m}^n,$$

where $\tilde{y}(Y^d, l, r)_i = \big|\Delta_i^n \tilde{Y}^d\big|^l \big|\Delta_{i+k_n}^n \tilde{Y}^d\big|^r$ with $\Delta_i^n \tilde{Y}^d = \Delta_i^n \bar{\epsilon}^d + \sigma_{sv, \frac{jb_n}{N_n}} \Delta_i^n \bar{W}$, for $jb_n + m \leq i \leq (j+1)b_n + m - 1$. We then show that the error due to replacing $\Delta_i^n \bar{Y}^d$ by $\Delta_i^n \tilde{Y}^d$ is small enough to be ignored and, hence, does not affect our theoretical results. This is true, because σ_{sv} is assumed

to be an Itô semimartingale itself, so that

$$E\Big(\big|\Delta_i^n\bar{Y}^d-\Delta_i^n\tilde{Y}^d\big|\Big)=E\Bigg(\Big|\sum_{j=1}^{k_n}g\Big(\frac{j}{k_n}\Big)\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}a_s^d\mathrm{d}s+\sum_{j=1}^{k_n}g\Big(\frac{j}{k_n}\Big)\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\sigma_{sv,s}-\sigma_{sv,\frac{jb_n}{N_n}}\Big)\mathrm{d}W_s\Big|\Bigg)$$
$$\leq K\left(\frac{k_n}{n}+\left(\sum_{j=1}^{k_n}g^2\Big(\frac{j}{k_n}\Big)E\Big(\Big|\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}}\Big(\sigma_{sv,s}-\sigma_{sv,\frac{jb_n}{N_n}}\Big)\mathrm{d}W_s\Big|\Big)^2\right)^{1/2}\right)$$
$$\leq K\left(\frac{k_n}{n}+\Big(\frac{k_n}{n}\frac{b_n}{n}\Big)^{1/2}\right)\leq K\frac{(k_nb_n)^{1/2}}{n}.$$

Note that $E\big(|B(Y^d,l,r)_m^n|\big)\leq K\frac{b_n}{n}$ uniformly in m, and so

$$E\Big(\big|B(Y^d,l,r)_{jb_n+m}^n-\tilde{B}(Y^d,l,r)_{jb_n+m}^n\big|\Big)\leq Kb_n\left(\frac{(k_nb_n)^{1/2}}{n}\Big(\frac{1}{\sqrt{k_n}}\Big)^{\frac{(l+r)}{4}-1}\right)$$
$$\leq K\Big(\frac{b_n}{n}\Big)^{3/2}.$$

As for $\bar{B}\left(Y^d,l,r\right)_{jb_n+m}^n$, we find that $E\Big(\big|B(Y^d,l,r)_{jb_n+m}^n-\bar{B}(Y^d,l,r)_{jb_n+m}^n\big|\Big)\leq K\Big(\frac{b_n}{n}\Big)^{3/2}$. And because $\delta<2/3$, we deduce that $\tilde{M}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d)-\bar{M}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d)=o_p(1)$, uniformly in m, where

$$\bar{M}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d)=\sqrt{n}\sum_{k=1}^{\lfloor\frac{N_n}{b_n}\rfloor-1}\Big(\mathcal{B}_{kb_n+m}^n-\hat{\mathcal{B}}_{kb_n+m}^n\Big),$$

such that

$\mathcal{B}_{kb_n+m}^n=\bar{B}(Y^d,l_1,r_1)_{kb_n+m}^n\bar{B}(Y^d,l_2,r_2)_{kb_n+m}^n$ and $\hat{\mathcal{B}}_{kb_n+m}^n=\bar{B}(Y^d,l_1,r_1)_{kb_n+m}^n\tilde{B}(Y^d,l_2,r_2)_{(k-1)b_n+m}^n$.

Then,

$$\sqrt{n}\left|\sum_{k=1}^{\lfloor\frac{N_n}{b_n}\rfloor-1}E\Big(\mathcal{B}_{kb_n+m}^n-E\Big(\mathcal{B}_{kb_n+m}^n\mid\mathcal{F}_{\frac{(k-1)b_n+m}{N_n}}^n\Big)\Big)\right|\leq K\frac{b_n^{3/2}}{n},$$
$$\sqrt{n}\left|\sum_{k=1}^{\lfloor\frac{N_n}{b_n}\rfloor-1}E\Big(\hat{\mathcal{B}}_{kb_n+m}^n-E\Big(\hat{\mathcal{B}}_{kb_n+m}^n\mid\mathcal{F}_{\frac{(k-1)b_n+m}{N_n}}^n\Big)\Big)\right|\leq K\frac{b_n^{3/2}}{n},$$

by conditional independence, and now we are left with

$$\bar{M}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d)=\sqrt{n}\sum_{k=1}^{\lfloor\frac{N_n}{b_n}\rfloor-1}E\Big(\mathcal{B}_{kb_n+m}^n-\hat{\mathcal{B}}_{kb_n+m}^n\mid\mathcal{F}_{\frac{(k-1)b_n+m}{N_n}}^n\Big)+o_p(1),$$

uniformly in m. As in Podolskij and Vetter (2010) and using $\delta>1/2$, we note that

$$\sqrt{n}E\Big(\mathcal{B}_{kb_n+m}^n-\hat{\mathcal{B}}_{kb_n+m}^n\mid\mathcal{F}_{\frac{(k-1)b_n+m}{N_n}}^n\Big)=2\theta\int_{\frac{(k-1)b_n}{N_n}}^{\frac{kb_n}{N_n}}\int_0^2h_{ij}\big(\sigma_{sv,u},t,f(s)\big)\mathrm{d}s\mathrm{d}u+o\Big(\frac{b_n}{N_n}\Big),$$

uniformly in k and m, and thus

$$\begin{aligned}\bar{M}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) &= 2\theta \int_0^1 \int_0^2 h_{ij}\big(\sigma_{sv,u}, t, f(s)\big)\mathrm{d}s\mathrm{d}u + o_p(1) \\ &= \int_0^1 w_{ij}^{l_1,r_1,l_2,r_2}\big(\sigma_{sv,u}\big)\mathrm{d}u + o_p(1),\end{aligned}$$

uniformly in m, and the proof is complete. ■

Proof of Theorem 4.1. We prove (64) solely in model (1), which is enough, as it is the most general and nests (103). Now, under the stated assumptions, the definitions of $\check{\Sigma}_{ij}^{l_1,r_1,l_2,r_2,n}(Y^d)$, $\check{\Sigma}_{ij}^{l_1,r_1,l_2,r_2,n}(Y'^d)$, and the limiting result in Theorem B.1, we deduce that, as $n \to \infty$,

$$\underset{n\to\infty}{\text{p-lim}}\ \hat{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y'^d) = \Sigma_{ij}^{l_1,r_1,l_2,r_2}, \quad \text{for } 1 \le i,j \le 2,$$

uniformly in m. Thus, to get the desired result, it suffices to show that

$$\underset{n\to\infty}{\text{p-lim}}\ \Big(\check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) - \check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y'^d)\Big) = 0, \quad \text{for } 1 \le i,j \le 2, \tag{113}$$

uniformly in m. Inserting the definition of $\check{\Sigma}_{ij}^{l_1,r_1,l_2,r_2,n}(Y^d)$ and $\check{\Sigma}_{ij}^{l_1,r_1,l_2,r_2,n}(Y'^d)$, it holds that

$$\begin{aligned}&\frac{2}{\sqrt{n}}\Big(\check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) - \check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y'^d)\Big) \\ &= \sum_{k=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \Big(\check{\Delta B}(Y^d, l_i, r_i)_{kb_n+m}^n \check{\Delta B}(Y^d, l_j, r_j)_{kb_n+m}^n - \check{\Delta B}(Y'^d, l_i, r_i)_{kb_n+m}^n \check{\Delta B}(Y'^d, l_j, r_j)_{kb_n+m}^n\Big) \\ &= \sum_{k=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \Big(\Big(\check{B}(Y^d, l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y^d, l_j, r_j)_{(k+1)b_n+m}^n - \check{B}(Y'^d, l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y'^d, l_j, r_j)_{(k+1)b_n+m}^n\Big) \\ &- \Big(\check{B}(Y^d, l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y^d, l_j, r_j)_{kb_n+m}^n - \check{B}(Y'^d, l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y'^d, l_j, r_j)_{kb_n+m}^n\Big) \\ &- \Big(\check{B}(Y^d, l_i, r_i)_{kb_n+m}^n \check{B}(Y^d, l_j, r_j)_{(k+1)b_n+m}^n - \check{B}(Y'^d, l_i, r_i)_{kb_n+m}^n \check{B}(Y'^d, l_j, r_j)_{(k+1)b_n+m}^n\Big) \\ &+ \Big(\check{B}(Y^d, l_i, r_i)_{kb_n+m}^n \check{B}(Y^d, l_j, r_j)_{kb_n+m}^n - \check{B}(Y'^d, l_i, r_i)_{kb_n+m}^n \check{B}(Y'^d, l_j, r_j)_{kb_n+m}^n\Big)\Big)\end{aligned} \tag{114}$$

where

$$\check{B}(Y^d, l, r)_j^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{b_n} \check{y}(Y^d, l, r)_{i-1+j}^n.$$

In the following, we define:

$$\begin{aligned}\pi_{k,k'}^{l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) &= \check{y}(Y^d, l_i, r_i)_k^n \check{y}(Y^d, l_j, r_j)_{k'}^n - \check{y}(Y'^d, l_i, r_i)_k^n \check{y}(Y'^d, l_j, r_j)_{k'}^n \\ &= \Big(|\Delta_k^n \bar{Y}^d|^{l_1} |\Delta_{k+k_n}^n \bar{Y}^d|^{r_1} |\Delta_{k'}^n \bar{Y}^d|^{l_2} |\Delta_{k'+k_n}^n \bar{Y}^d|^{r_2} \\ &\quad - |\Delta_k^n \bar{Y}'^d|^{l_1} |\Delta_{k+k_n}^n \bar{Y}'^d|^{r_1} |\Delta_{k'}^n \bar{Y}'^d|^{l_2} |\Delta_{k'+k_n}^n \bar{Y}'^d|^{r_2}\Big) 1_{C_{k,k'}},\end{aligned}$$

where $C_{k,k'} = \big\{|\Delta_k^n \bar{Y}^d| < \upsilon_n\big\} \cap \big\{|\Delta_{k+k_n}^n \bar{Y}^d| < \upsilon_n\big\} \cap \big\{|\Delta_{k'}^n \bar{Y}'^d| < \upsilon_n\big\} \cap \big\{|\Delta_{k'+k_n}^n \bar{Y}'^d| < \upsilon_n\big\}$. Then, from (114) it follows that

$$
\begin{aligned}
&\check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y^d) - \check{\Sigma}_{ij,m}^{l_1,r_1,l_2,r_2,n}(Y'^d) \\
&= \frac{n^{\frac{l_1+r_1+l_2+r_2-6}{4}}}{2\mu_{l_1}\mu_{r_1}\mu_{l_1}\mu_{r_2}} \sum_{j=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \sum_{k=1}^{b_n}\sum_{k'=1}^{b_n} \Big(\pi_{k-1+(j+1)b_n+m,k'-1+(j+1)b_n+m}^{l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) - \pi_{k-1+(j+1)b_n+m,k'-1+jb_n+m}^{l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) \\
&- \pi_{k-1+jb_n+m,k'-1+(j+1)b_n+m}^{l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) + \pi_{k-1+jb_n+m,k'-1+jb_n+m}^{l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) \Big) \\
&\equiv \check{\Sigma}_{ij,m}^{(1),l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) - \check{\Sigma}_{ij,m}^{(2),l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) - \check{\Sigma}_{ij,m}^{(3),l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) + \check{\Sigma}_{ij,m}^{(4),l_1,r_1,l_2,r_2,n}(Y^d, Y'^d).
\end{aligned}
$$

The statement in (113) is therefore reduced to showing that

$$
\check{\Sigma}_{ij,m}^{(k),l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) \xrightarrow{p} 0, \tag{115}
$$

for $k = 1, \ldots, 4$. The convergence in probability to zero of the four terms is proven with identical techniques. It is therefore sufficient to show it for a single k, so we do it with $k = 1$. To this end, let $F_u(x) = F(x)1_{\{|x_1|<u\}}1_{\{|x_2|<u\}}1_{\{|x_3|<u\}}1_{\{|x_4|<u\}}$, for $u > 0$, where $F(x) = |x_1|^{l_1}|x_2|^{r_1}|x_3|^{l_2}|x_4|^{r_2}$ with $x = \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix}^\intercal$. Following the line of thought used also in the proof of Theorem 2.1, we can show that for $w_n = \upsilon_n/\sqrt{u_n}$ with $u_n = k_n/n$:

$$
|F_{w_n}(x+y) - F_{w_n}(x)| \leq w_n^{\frac{-2}{1-2\omega}} \|x\|^{p+\frac{2}{1-2\omega}} + \Big((1 + \|x\|^p)\big(\|y\| \wedge 1 + (\|y\| \wedge w_n)^p\big) \Big),
$$

where $p = l_1 + r_1 + l_2 + r_2$. Next, set $x = \begin{pmatrix} \Delta_k^n \bar{Y}^d & \Delta_{k+k_n}^n \bar{Y}^d & \Delta_{k'}^n \bar{Y}^d & \Delta_{k'+k_n}^n \bar{Y}^d \end{pmatrix}^\intercal / \sqrt{u_n}$, $y = \begin{pmatrix} \Delta_k^n \bar{Y}''^d & \Delta_{k+k_n}^n \bar{Y}''^d & \Delta_{k'}^n \bar{Y}''^d & \Delta_{k'+k_n}^n \bar{Y}''^d \end{pmatrix}^\intercal / \sqrt{u_n}$. As in (101) – (102), it holds true that

$$
E\big(\|x\|^p\big) \leq K, \quad E\big(\|y\| \wedge 1\big) \leq K u_n^{1-\beta/2}\phi_n \quad \text{and} \quad E\big((\|y\| \wedge w_n)^p\big) \leq K u_n^{\omega(p-\beta)-\frac{(p-2)}{2}}\phi_n, \tag{116}
$$

where $\phi_n \to 0$ as $n \to 0$. Therefore,

$$
\begin{aligned}
&\frac{n^{\frac{l_1+r_1+l_2+r_2-6}{4}}}{2\mu_{l_1}\mu_{r_1}\mu_{l_1}\mu_{r_2}} \sum_{j=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \sum_{k=1}^{b_n}\sum_{k'=1}^{b_n} \underbrace{E\Big(\big| \pi_{k-1+(j+1)b_n+m,k'-1+(j+1)b_n+m}^{l_1,r_1,l_2,r_2,n}(Y^d, Y'^d) \big| \Big)}_{= O\left(u_n^{\frac{l_1+r_1+l_2+r_2}{2}} \left(u_n + u_n^{1-r/2}\phi_n + u_n^{\omega(4-r)-1}\phi_n \right) \right)} \\
&\leq K n^{\frac{4\delta_1-2}{4}} \Big(n^{-\frac{1}{2}} + n^{\frac{\beta-2}{4}}\phi_n + n^{\frac{l_1+r_1+l_2+r_2-2-2\omega(l_1+r_1+l_2+r_2-\beta)}{4}}\phi_n \Big) \\
&\leq K \Big(n^{\delta_1-1} + n^{\frac{4\delta_1-4+\beta}{4}}\phi_n + n^{\frac{4\delta_1-4+l_1+r_1+l_2+r_2-2\omega(l_1+r_1+l_2+r_2-\beta)}{4}}\phi_n \Big) \to 0,
\end{aligned}
$$

which concludes the proof of (113) and, hence, Theorem 4.1. ■

Proof of Lemma 4.1. The linearity of the expectation operator implies that

$$E^*\big(\check{BV}(l,r)^{n*}\big) = E^*\left[\check{BV}(l,r)^n - \frac{1}{\sqrt{b_n}}\sum_{j=1}^{J_n}\check{\Delta B}(l,r)_j^n u_j\right]$$

$$= \check{BV}(l,r)^n - \frac{1}{\sqrt{b_n}}\sum_{j=1}^{J_n}\check{\Delta B}(l,r)_j^n E^*(u_j).$$

Then, if $E^*(u_j) = 0$, it follows that $E^*\big(\check{BV}(l,r)^{n*}\big) = \check{BV}(l,r)^n$. The second part of the lemma follows from (46) and (52), as for $1 \leq i,j \leq 2$,

$$\begin{aligned}
&\mathrm{cov}^*\big(n^{1/4}\check{BV}(l_i,r_i)^{n*}, n^{1/4}\check{BV}(l_j,r_j)^{n*}\big) \\
&= \sqrt{n}\,\mathrm{cov}^*\left(\check{BV}(l_i,r_i)^n - \frac{1}{\sqrt{b_n}}\sum_{k=1}^{J_n}\check{\Delta B}(l_i,r_i)_k^n u_k, \check{BV}(l_j,r_j)^n - \frac{1}{\sqrt{b_n}}\sum_{k=1}^{J_n}\check{\Delta B}(l_j,r_j)_k^n u_k\right) \\
&= \frac{\sqrt{n}}{b_n}\sum_{k=1}^{J_n}\check{\Delta B}(l_i,r_i)_k^n\check{\Delta B}(l_j,r_j)_k^n \mathrm{var}^*(u_k).
\end{aligned}$$

Thus, if $\mathrm{var}^*(u_k) = 1/2$, we find that

$$\mathrm{cov}^*\big(n^{1/4}\check{BV}(l_i,r_i)^{n*}, n^{1/4}\check{BV}(l_j,r_j)^{n*}\big) = \check{\Sigma}_{ij}^{l_1,r_1,l_2,r_2,n},$$

where

$$\check{\Sigma}_{ij}^{l_1,r_1,l_2,r_2,n} = \frac{\sqrt{n}}{2b_n}\sum_{k=1}^{J_n}\check{\Delta B}(l_i,r_i)_k^n\check{\Delta B}(l_j,r_j)_k^n.$$

■

Proof of Corollary 4.1. Given (25), (27), and (65) the results follows from the properties of stable convergence. ■

Proof of Corollary 4.2. The result follows directly given (57) and the consistency result of $\check{\Sigma}^{l_1,r_1,l_2,r_2}$ in Theorem 4.1. ■

Proof of Theorem 4.2. We again prove the theorem in model (1) only, noting that this is enough, as it nests both (103) and (10). Write

$$Z^{n*} = \big(\check{\Sigma}^n\big)^{-1/2} n^{1/4}\sum_{j=1}^{J_n} D_j e_j^* \equiv n^{1/4}\sum_{j=1}^{J_n} z_j^*,$$

with $z_j^* \equiv \big(\check{\Sigma}^n\big)^{-1/2} D_j e_j^*$,

$$D_j = \begin{pmatrix} \check{\Delta B}(l_1,r_1)_j^n & 0 \\ 0 & \check{\Delta B}(l_2,r_2)_j^n, \end{pmatrix} \quad \text{and} \quad e_j^* = \begin{pmatrix} u_j - E^*(u_j) \\ u_j - E^*(u_j) \end{pmatrix}$$

where u_j are i.i.d. with $\mathrm{var}^*(u_j) = 1/2$. Note that e_j^* is an i.i.d. zero mean vector. We follow Pauly (2011) and use a modified Cramer-Wold device to establish the bootstrap CLT. Let $D = \{\lambda_k : k \in$

$N\}$ be a countable dense subset of the unit circle on $\mathbb{R}^2$. The proof follows by showing that for any $\lambda \in D, \lambda^\intercal Z_n^* \xrightarrow{d^*} N(0,1)$, in probability-$P$, as $n \to \infty$. We note that

$$\lambda^\intercal Z_n^* = n^{1/4} \sum_{j=1}^{J_n} \lambda^\intercal z_j^*.$$

It follows from Lemma 4.1 and Corollary 4.2 that $E^*(\lambda^\intercal Z_n^*) = 0$ and $\text{var}^*(\lambda^\intercal Z_n^*) = 1$ for all n. To conclude, it thus remains to prove that $\lambda^\intercal Z_n^*$ is asymptotically normally distributed, conditionally on the original sample and with probability P approaching one. As $(z_j^*)_{j=1}^{J_n}$ forms an independent array—conditionally on the sample—by the Berry-Esseen bound (e.g., Katz (1963)), for some small $\varepsilon > 0$ and a constant $K > 0$, $\sup_{x\in\mathbb{R}} \left| P^*\big(\sum_{j=1}^{J_n} n^{1/4} \lambda^\intercal z_j^* \leq x \big) - \Phi(x) \right| \leq K \sum_{j=1}^{J_n} E^* \left| n^{1/4} \lambda^\intercal z_j^* \right|^{2+\varepsilon}$. Next, we show that $\sum_{j=1}^{J_n} E^* \left| n^{1/4} \lambda^\intercal z_j^* \right|^{2+\varepsilon} = o_p(1)$. First, for a constant K independent of n (note that the moments of e_j^* do not depend on n) and any $1 \leq j \leq J_n$ by the c_r-inequality:

$$\left| \lambda^\intercal z_j^* \right|^{2+\varepsilon} \leq \|\lambda\|^{2+\varepsilon} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \|D_j\|^{2+\varepsilon} \|e_j^*\|^{2+\varepsilon}.$$

Thus,

$$\begin{aligned} E^*\big(|\lambda^\intercal z_j^*|^{2+\varepsilon} \big) &\leq \|\lambda\|^{2+\varepsilon} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \|D_j\|^{2+\varepsilon} E^*\big(\|e_j^*\|^{2+\varepsilon} \big) \\ &\leq K \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \|D_j\|^{2+\varepsilon}, \end{aligned}$$

implying that

$$\begin{aligned} \sum_{j=1}^{J_n} E^* |n^{1/4} \lambda^\intercal z_j^*|^{2+\varepsilon} &\leq K n^{\frac{2+\varepsilon}{4}} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \|D_j\|^{2+\varepsilon} \\ &\leq K n^{\frac{2+\varepsilon}{4}} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \Big(\big(\check{\Delta B}(l_1, r_1)_j^n\big)^{2+\varepsilon} + \big(\check{\Delta B}(l_2, r_2)_j^n\big)^{2+\varepsilon} \Big) \\ &\leq K n^{\frac{2+\varepsilon}{4}} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \big(\check{B}(l_1, r_1)_{j+b_n}^n - \check{B}(l_1, r_1)_j^n \big)^{2+\varepsilon} \\ &+ K n^{\frac{2+\varepsilon}{4}} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \big(\check{B}(l_2, r_2)_{j+b_n}^n - \check{B}(l_2, r_2)_j^n \big)^{2+\varepsilon} \\ &\leq K n^{\frac{2+\varepsilon}{4}} \left\| \big(\check{\Sigma}^n\big)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \Big(\big(\check{B}(l_1, r_1)_j^n\big)^{2+\varepsilon} + \big(\check{B}(l_2, r_2)_j^n\big)^{2+\varepsilon} \Big), \end{aligned} \tag{117}$$

where the second inequality is due to that, for any j, $\|D_j\|^2 = \big(\check{\Delta B}(l_1, r_1)_j^n\big)^2 + \big(\check{\Delta B}(l_2, r_2)_j^n\big)^2$,

while the third is by expression of $\Delta\check{B}(l,r)^n_j$. Next, note that by definition of $\check{B}(l,r)^n_j$:

$$\begin{aligned}\sum_{j=1}^{J_n}\big(\check{B}(l,r)^n_j\big)^{2+\varepsilon} &= \sum_{j=1}^{J_n}\left(n^{\frac{l+r}{4}-1}\frac{1}{\mu_l\mu_r}\sum_{i=1}^{b_n}\check{y}(l,r)^n_{i-1+j}\right)^{2+\varepsilon}\\ &\leq Kn^{\left(\frac{l+r}{4}-1\right)(2+\varepsilon)}b_n^{1+\varepsilon}\sum_{j=1}^{J_n}\sum_{i=1}^{b_n}\big(\check{y}(l,r)^n_{i-1+j}\big)^{2+\varepsilon}\\ &= O_p\big(n^{(\delta_1-1)(1+\varepsilon)}\big).\end{aligned}$$

We can therefore write (117) as follows:

$$\begin{aligned}\sum_{j=1}^{J_n}E^*\big(|n^{1/4}\lambda^\intercal z^*_j|^{2+\varepsilon}\big) &= O_p\Big(n^{\frac{2+\varepsilon}{4}}n^{(\delta_1-1)(1+\varepsilon)}\Big)\\ &= o_p(1),\end{aligned}$$

where the last equality follows as for $\varepsilon > 2$, so long as $1/2 < \delta_1 < 2/3$, $(\delta_1-1)(1+\varepsilon)+\frac{2+\varepsilon}{4} < 0$. This completes the proof of (71). The last results then follow by application of the delta rule. ■

Proof of Theorem 4.3. First, we define:

$$\begin{aligned}H^{n*} &= \big(\check{\Sigma}^{n*}\big)^{-1/2}n^{1/4}\begin{pmatrix}\check{BV}(l_1,r_1)^{n*}-E^*\big(\check{BV}(l_1,r_1)^{n*}\big)\\ \check{BV}(l_2,r_2)^{n*}-E^*\big(\check{BV}(l_2,r_2)^{n*}\big)\end{pmatrix}\\ &\equiv \big(\check{\Sigma}^{n*}\big)^{-1/2}\big(\check{\Sigma}^{n}\big)^{1/2}Z^{n*},\end{aligned}$$

where

$$Z^{n*} = \big(\check{\Sigma}^{n}\big)^{-1/2}n^{1/4}\begin{pmatrix}\check{BV}(l_1,r_1)^{n*}-E^*\big(\check{BV}(l_1,r_1)^{n*}\big)\\ \check{BV}(l_2,r_2)^{n*}-E^*\big(\check{BV}(l_2,r_2)^{n*}\big)\end{pmatrix}.$$

It follows from Theorem 4.2 that $Z^{n*} \xrightarrow{d^*} N(0,I_2)$. Thus, the central limit theory for H^{n*} is established, if we can show that $\big(\check{\Sigma}^{n*}\big)^{-1}\check{\Sigma}^{n} = \big(\check{\Sigma}^{n}\big)^{-1}\check{\Sigma}^{n*} \xrightarrow{p^*} I_2$. To do this, we prove that

$$E^*\Big[\big(\check{\Sigma}^{n}\big)^{-1}\check{\Sigma}^{n*}\Big] \xrightarrow{p^*} I_2 \qquad \text{and} \qquad \mathrm{var}^*\Big[\big(\check{\Sigma}^{n}\big)^{-1}\check{\Sigma}^{n*}\Big] \xrightarrow{p^*} 0. \tag{118}$$

The first equation in (118) holds by the definition of $\check{\Sigma}^n$ and $\check{\Sigma}^{n*}$. Next, again by definition:

$$\begin{aligned}\mathrm{var}^*\Big[\big(\check{\Sigma}^{n}\big)^{-1}\big(\check{\Sigma}^{n*}\big)\Big] &= \Big[\big(\check{\Sigma}^{n}\big)^{-1}\oplus\big(\check{\Sigma}^{n}\big)^{-1}\Big]\ \mathrm{var}^*\left(\frac{\sqrt{n}}{b_n}\frac{\mathrm{var}^*(u)}{E^*(u^2)}\sum_{j=1}^{J_n}\check{\xi}_j\check{\xi}_j^\intercal u_j^2\right)\\ &= \left(\frac{\sqrt{n}}{b_n}\frac{\mathrm{var}^*(u)}{E^*(u^2)}\right)^2\Big[\big(\check{\Sigma}^{n}\big)^{-1}\oplus\big(\check{\Sigma}^{n}\big)^{-1}\Big]\sum_{j=1}^{J_n}\mathrm{var}^*\big(\check{\xi}_j\check{\xi}_j^\intercal u_j^2\big)\\ &= \mathrm{var}^*(u^2)\left(\frac{\mathrm{var}^*(u)}{E^*(u^2)}\right)^2\Big[\big(\check{\Sigma}^{n}\big)^{-1}\oplus\big(\check{\Sigma}^{n}\big)^{-1}\Big]\frac{n}{b_n^2}\sum_{j=1}^{J_n}\big(\check{\xi}_j\check{\xi}_j^\intercal\big)\oplus\big(\check{\xi}_j\check{\xi}_j^\intercal\big).\end{aligned}$$

As in the proof of Theorem 4.2 in Hounyo (2017):

$$E\left(\left\|\frac{n}{b_n^2}\sum_{j=1}^{J_n}\left(\check{\xi}_j\check{\xi}_j^{\intercal}\right)\oplus\left(\check{\xi}_j\check{\xi}_j^{\intercal}\right)\right\|\right)\leq K\frac{n}{b_n^2}\sum_{j=1}^{J_n}\Bigg(\sqrt{E\Big(\big|\check{B}(l_1,r_1)_j^n\big|^4\Big)}\sqrt{E\Big(\big|\check{B}(l_2,r_2)_j^n\big|^4\Big)}$$
$$+\sqrt{E\Big(\big|\check{B}(l_1,r_1)_{j+b_n}^n\big|^4\Big)}\sqrt{E\Big(\big|\check{B}(l_2,r_2)_j^n\big|^4\Big)}$$
$$+\sqrt{E\Big(\big|\check{B}(l_1,r_1)_j^n\big|^4\Big)}\sqrt{E\Big(\big|\check{B}(l_2,r_2)_{j+b_n}^n\big|^4\Big)}$$
$$+\sqrt{E\Big(\big|\check{B}(l_1,r_1)_{j+b_n}^n\big|^4\Big)}\sqrt{E\Big(\big|\check{B}(l_2,r_2)_{j+b_n}^n\big|^4\Big)}\Bigg)\leq K\frac{b_n^2}{n^2}\to 0,$$

where the last inequality follows, because $\dfrac{l+r-1}{2(l+r-\beta)}\leq\varpi<1/2$ means that $\sqrt{E\Big(\big|\check{B}(l,r)_j^n\big|^4\Big)}\leq K\dfrac{b_n^2}{n^2}$. As $J_n=O(n)$ and $b_n=O\big(n^{\delta_1}\big)$ such that $1/2<\delta_1<2/3$ from (44), it follows that

$$\mathrm{var}^*\Big[\big(\check{\Sigma}^n\big)^{-1}\check{\Sigma}^{n*}\Big]\xrightarrow{p^*}0.$$

This finishes the proof of the first part Theorem 4.3. The last result again follows by a direct application of the delta rule. ■

Proof of Proposition 3.1. To begin with, notice that

$$\big(\sqrt{m}\Delta_{(t-1)m+i}^m Y\big)^2=\big(\sqrt{m}\Delta_{(t-1)m+i}^m X\big)^2+\sqrt{m}\big(\sqrt{m}\Delta_{(t-1)m+i}^m X\big)\big(\Delta_{(t-1)m+i}^m\epsilon\big)+m\big(\Delta_{(t-1)m+i}^m\epsilon\big)^2.$$

Thus, for $s\in[t-1+(j-1)/m,t-1+j/m)$, where $j=1,\ldots,m$ and $t=1,\ldots,T$,

$$\begin{aligned}\hat{\sigma}_{u,s}^2=&\frac{1}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta_{(t-1)m+j}^m X\big)^2+\frac{\sqrt{m}}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta_{(t-1)m+j}^m X\big)\big(\Delta_{(t-1)m+j}^m\epsilon\big)\\&+\frac{m}{T}\sum_{t=1}^{T}\big(\Delta_{(t-1)m+j}^m\epsilon\big)^2-\frac{m}{T}\sum_{t=1}^{T}\big[\hat{\mathrm{var}}\big(\epsilon_{(t-1)+(j-1)/m}\big)+\hat{\mathrm{var}}\big(\epsilon_{(t-1)+j/m}\big)\big].\end{aligned}\tag{119}$$

The proof now proceeds in three steps, where we show that:

$$\frac{1}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta_{(t-1)m+j}^m X\big)^2=\sigma_{u,s}^2+O_{L^2}\big(T^{-1/2}m^{1/4}\big),\tag{120}$$

$$\frac{\sqrt{m}}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta_{(t-1)m+j}^m X\big)\big(\Delta_{(t-1)m+j}^m\epsilon\big)=O_P\Big(\sqrt{m/T}\Big),\tag{121}$$

and

$$\frac{m}{T}\sum_{t=1}^{T}\big(\Delta_{(t-1)m+j}^m\epsilon\big)^2=\frac{m}{T}\sum_{t=1}^{T}\big(\mathrm{var}\big(\epsilon_{t-1+(j-1)/m}\big)+\mathrm{var}\big(\epsilon_{t-1+j/m}\big)\big)+O_P\big(mT^{-1/2}\big).\tag{122}$$

To show step 1, we define:

$$\alpha^m_{(t-1)m+j} \equiv \sqrt{m}\sigma_{t-1+(j-1)/m}\Delta^m_{(t-1)m+j}W \quad \text{and} \quad \chi^m_{(t-1)m+j} \equiv \sqrt{m}\Delta^m_{(t-1)m+j}X - \alpha^m_{(t-1)m+j}.$$

We also set

$$\tilde{\sigma}^2_{u,s} \equiv \frac{1}{T}\sum_{t=1}^{T}\sigma^2_{t-1+(j-1)/m}.$$

Now, the proof is complete, if we can show that:

$$\tilde{\sigma}^2_{u,s} - \sigma^2_{u,s} = O_{L^2}\big(T^{-1/2}\big), \tag{123}$$

$$\frac{1}{T}\sum_{t=1}^{T}|\alpha^m_{(t-1)m+j}|^2 - \tilde{\sigma}^2_{u,s} = O_{L^2}\big(T^{-1/2}\big), \tag{124}$$

and

$$\frac{1}{T}\sum_{t=1}^{T}\Big(\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)^2 - |\alpha^m_{(t-1)m+j}|^2\Big) = O_{L^2}\big(T^{-1/2}m^{1/4}\big). \tag{125}$$

As for (123), note that

$$\tilde{\sigma}^2_{u,s} \equiv \frac{1}{T}\sum_{t=1}^{T}\sigma^2_{t-1+(j-1)/m} = \frac{\sigma^2_{u,(j-1)/m}}{T}\sum_{t=1}^{T}\sigma^2_{sv,t-1+(j-1)/m}.$$

Now, by Assumption (D2) we deduce that

$$\mathrm{var}\big(\tilde{\sigma}^2_{u,s}\big) \leq \frac{K}{T}\bigg(1 + 2\sum_{k=0}^{\infty}\mathrm{cov}\big(\sigma^2_{sv,1}, \sigma^2_{sv,1+k}\big)\bigg).$$

Hence (123) follows.

Next, (124) can be verified by martingale techniques. First, we write

$$\frac{1}{T}\sum_{t=1}^{T}|\alpha^m_{(t-1)m+j}|^2 - \tilde{\sigma}^2_{u,s} = \frac{1}{T}\sum_{t=1}^{T}\bigg(|\alpha^m_{(t-1)m+j}|^2 - E\Big(|\alpha^m_{(t-1)m+j}|^2 \mid \mathcal{F}_{t-1+\frac{j-1}{m}}\Big)\bigg).$$

Then,

$$\begin{aligned} E\left[\frac{1}{T}\sum_{t=1}^{T}|\alpha^m_{(t-1)m+j}|^2 - \tilde{\sigma}^2_{u,s}\right]^2 &= \frac{1}{T^2}\sum_{t=1}^{T}E\bigg(|\alpha^m_{(t-1)m+j}|^2 - E\Big(|\alpha^m_{(t-1)m+j}|^2 \mid \mathcal{F}_{t-1+\frac{j-1}{m}}\Big)\bigg)^2 \\ &= \frac{2}{T^2}\sum_{t=1}^{T}\sigma^4_{t+(j-1)/m}. \end{aligned}$$

Treating the error $T^{-1}\sum_{t=1}^{T}\Big(\big(\sqrt{m}\Delta^m_{(t-1)+j}X\big)^2 - |\alpha^m_{(t-1)m+j}|^2\Big)$ from (125)) is the hardest one. In the following, we denote

$$X'_t = X_0 + \int_0^t a''_s\mathrm{d}s + \int_0^t \sigma_s\mathrm{d}W_s \quad \text{and} \quad X''_t = X_t - X'_t = \big(\delta 1_{\{|\delta|\leq 1\}}\big) \star \big(\underline{\mu}_t - \underline{\nu}_t\big),$$

where

$$a_t'' = a_s + \big(\delta 1_{\{|\delta|>1\}}\big) \star \underline{\nu}_t$$

is bounded.

Also, note that an error $O_{L^2}\big(T^{-1/2}m^{1/4}\big)$ appears, when we are dealing with terms that do not have a martingale structure (i.e., the jump component of X). Here, Corollary 2.1.9 of Jacod and Protter (2012) plays a key role in the proof, because it turns out the discontinuous part is asymptotically negligible in front of its Brownian part on a vanishing interval of the form $[t-1+(i-1)/m, t-1+i/m)$.

The definition of $\chi^m_{(t-1)m+j}$ and Assumption (V) yields the decomposition:

$$\begin{aligned}
\chi^m_{(t-1)m+j} &= \sqrt{m}\Delta^m_{(t-1)m+j}X - \alpha^m_{(t-1)m+j} \\
&= \sqrt{m}\Delta^m_{(t-1)m+j}X' + \sqrt{m}\Delta^m_{(t-1)m+j}X'' - \alpha^m_{(t-1)m+j} \\
&= \sqrt{m}\Bigg(\int_{t-1+\frac{j-1}{m}}^{t-1+\frac{j}{m}} a_s'' \mathrm{d}s + \int_{t-1+\frac{j-1}{m}}^{t-1+\frac{j}{m}} \Big(\sigma_s - \sigma_{t-1+\frac{j-1}{m}}\Big)\mathrm{d}W_s + \Delta^m_{(t-1)m+j}X''\Bigg) \\
&\equiv \chi^m_{(t-1)m+j}(1) + \chi^m_{(t-1)m+j}(2) + \chi^m_{(t-1)m+j}(3),
\end{aligned}$$

where

$$\chi^m_{(t-1)m+j}(1) = \sqrt{m}\Bigg(\frac{1}{m}a''_{t-1+\frac{j-1}{m}} + \int_{t-1+\frac{j-1}{m}}^{t-1+\frac{j}{m}} \Big[\tilde{\sigma}_{t-1+\frac{j-1}{m}}\Big(W_s - W_{t-1+\frac{j-1}{m}}\Big) + \tilde{v}_s\Big(B_s - B_{t-1+\frac{j-1}{m}}\Big)\Big]\mathrm{d}W_s\Bigg),$$

$$\begin{aligned}
\chi^m_{(t-1)m+j}(2) &= \sqrt{m}\Bigg(\int_{t-1+\frac{j-1}{m}}^{t-1+\frac{j}{m}} \Big(a_s'' - a''_{t-1+\frac{j-1}{m}}\Big)\mathrm{d}s + \int_{t-1+\frac{j-1}{m}}^{t-1+\frac{j}{m}} \Bigg[\int_{t-1+\frac{j-1}{m}}^{s} \tilde{a}_s \mathrm{d}u\Bigg]\mathrm{d}W_s\Bigg) \\
&\quad + \sqrt{m}\Bigg(\int_{t-1+\frac{j-1}{m}}^{t-1+\frac{j}{m}} \Big(\tilde{\sigma}_u - \tilde{\sigma}_{t-1+\frac{j-1}{m}}\Big)\mathrm{d}W_s + \Bigg[\int_{t-1+\frac{j-1}{m}}^{s} \Big(\tilde{v}_u - \tilde{v}_{t-1+\frac{j-1}{m}}\Big)\mathrm{d}B_s\Bigg]\mathrm{d}W_s\Bigg),
\end{aligned}$$

$$\chi^m_{(t-1)m+j}(3) = \sqrt{m}\Delta^m_{(t-1)m+j}X''.$$

Together with the assumptions behind Proposition 3.1, we can then appeal to the Burkholder and Cauchy-Schwarz inequality (first and second expression) and Lemma 2.1.5 in Jacod and Protter (2012) with $p=4$ (last estimate) to deduce that:

$$\chi^m_{(t-1)m+j}(1) = O_{L^4}\big(m^{-1/2}\big), \quad \chi^m_{(t-1)m+j}(2) = O_{L^4}\big(m^{-1}\big), \quad \chi^m_{(t-1)m+j}(3) = O_{L^4}\big(m^{1/4}\big). \tag{126}$$

Next, let $f(x) = x^2$ so that $f'(x) = 2x$. Then, by Taylor expansion:

$$\frac{1}{T}\sum_{t=1}^{T}\Big(\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)^2 - |\alpha^m_{(t-1)m+j}|^2\Big) = A^m_{(t-1)m+j}(1) + A^m_{(t-1)m+j}(2) + A^m_{(t-1)m+j}(3) + O_{L^2}\big(T^{-1/2}\big),$$

where

$$A^m_{(t-1)m+j}(k) = \frac{2}{T}\sum_{t=1}^{T}\big(\alpha^m_{(t-1)m+j}\chi^m_{(t-1)m+j}(k)\big),$$

for $k = 1, 2$ and 3. Note the following martingale difference property:

$$A^m_{(t-1)m+j}(k) = E\Big(\alpha^m_{(t-1)m+j}\chi^m_{(t-1)m+j}(k) \mid \mathcal{F}_{t-1+\frac{j-1}{m}}\Big) = 0,$$

for $k = 1, 2$ and 3. Thus, from the Cauchy-Schwarz inequality

$$\begin{aligned} E\big[|A^m_{(t-1)m+j}(1)|^2\big] &= \frac{4}{T^2}\sum_{t=1}^{T} E\Big[|\big(\alpha^m_{(t-1)m+j}\chi^m_{(t-1)m+j}(1)\big)|^2\Big] \\ &\leq \frac{4}{T^2}\sum_{t=1}^{T}\big(E|\left(\alpha^m_{j+tm}\right)|^4\big)^{1/2}\big(E|\left(\chi^m_{j+tm}\left(1\right)\right)|^4\big)^{1/2}. \end{aligned}$$

Thus, given (126) and the fact that $\alpha^m_{(t-1)m+j} = O_{L^4}(1)$, it follows that

$$A^m_{(t-1)m+j}(1) = O_{L^2}\big(T^{-1/2}m^{-1/2}\big), \quad A^m_{(t-1)m+j}(2) = O_{L^2}\big(T^{-1/2}m^{-1}\big), \quad \text{and} \quad A^m_{(t-1)m+j}(3) = O_{L^2}\big(T^{-1/2}m^{1/4}\big).$$

This shows (125), and then we conclude that:

$$\begin{aligned} \frac{1}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)^2 &= \sigma^2_{u,s} + O_{L^2}\big(T^{-1/2}\big) + O_{L^2}\big(T^{-1/2}m^{-1/2}\big) + O_{L^2}\big(T^{-1/2}m^{-1}\big) + O_{L^2}\big(T^{-1/2}m^{1/4}\big) \\ &= \sigma^2_{u,s} + O_{L^2}\big(T^{-1/2}m^{1/4}\big), \end{aligned}$$

which completes the entire proof of step 1. We move forward to step 2. To deduce (121), we write:

$$\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\big(\Delta^m_{(t-1)m+j}\epsilon\big) = \sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\epsilon_{t-1+j/m} - \sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\epsilon_{t-1+(j-1)/m}.$$

Note that from (5) and as ϵ_t is independently distributed with $X \perp\!\!\!\perp \epsilon$:

$$E\left[\frac{\sqrt{m}}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\epsilon_{t-1+j/m} \mid X\right] = 0 \quad \text{and} \quad E\left[\frac{\sqrt{m}}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\epsilon_{t-1+(j-1)/m} \mid X\right] = 0.$$

Then, we get

$$E\left[\left(\frac{\sqrt{m}}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\epsilon_{t-1+k/m}\right)^2 \mid X\right] = \frac{m}{T}\omega^2\left[\frac{1}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)^2\sigma^2_{u,t-1+k/m}\right],$$

for $k = j$ and $j-1$. And from $T^{-1}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)^2 = O_P(1)$ and $\sigma^2_{u,t}$ being bounded:

$$\frac{\sqrt{m}}{T}\sum_{t=1}^{T}\big(\sqrt{m}\Delta^m_{(t-1)m+j}X\big)\epsilon_{t-1+k/m} = O_P\big(T^{-1/2}m^{1/2}\big),$$

for $k = j$ and $j-1$. This establishes (121). To show (122), note that because $\big(\Delta^m_{(t-1)m+j}\epsilon\big)^2$ is a 1-dependent sequence: $E\Big[\big(\Delta^m_{(t-1)m+j}\epsilon\big)^2\Big] = \mathrm{var}\big(\epsilon_{t-1+(j-1)/m}\big) + \mathrm{var}\big(\epsilon_{t-1+j/m}\big)$, and

$$\mathrm{var}\left[\frac{1}{T}\sum_{t=1}^{T}\Big(\big(\Delta^m_{(t-1)m+j}\epsilon\big)^2 - \big(\mathrm{var}(\epsilon_{t-1+(j-1)/m}) + \mathrm{var}(\epsilon_{t-1+j/m})\big)\Big)\right] \leq \frac{K}{T}.$$

Hence,

$$\frac{1}{T}\sum_{t=1}^{T}\big(\Delta^m_{(t-1)m+j}\epsilon\big)^2 = \frac{1}{T}\sum_{t=1}^{T}\big(\mathrm{var}(\epsilon_{t-1+(j-1)/m}) + \mathrm{var}(\epsilon_{t-1+j/m})\big) + O_P\big(T^{-1/2}\big).$$

It follows that

$$\begin{aligned}
&\frac{m}{T}\sum_{t=1}^{T}\Big(\big(\Delta^m_{(t-1)m+j}\epsilon\big)^2 - \big(\hat{\mathrm{var}}(\epsilon_{t-1+(j-1)/m}) + \hat{\mathrm{var}}(\epsilon_{t-1+j/m})\big)\Big) \\
&\qquad = -\frac{m}{T}\sum_{t=1}^{T}\Big[\big(\hat{\mathrm{var}}(\epsilon_{t-1+(j-1)/m}) - \mathrm{var}(\epsilon_{t-1+(j-1)/m})\big) + \big(\hat{\mathrm{var}}(\epsilon_{t-1+j/m}) - \mathrm{var}(\epsilon_{t-1+j/m})\big)\Big] \\
&\qquad + O_P\big(mT^{-1/2}\big),
\end{aligned}$$

so (122) holds. Inserting (120) – (122) into (119):

$$\hat{\sigma}^2_{u,s} = \sigma^2_{u,s} + O_P\big(mT^{-1/2}\big) - \frac{m}{T}\sum_{t=1}^{T}\Big[\big(\hat{\mathrm{var}}(\epsilon_{t-1+(j-1)/m}) - \mathrm{var}(\epsilon_{t-1+(j-1)/m})\big) + \big(\hat{\mathrm{var}}(\epsilon_{t-1+j/m}) - \mathrm{var}(\epsilon_{t-1+j/m})\big)\Big].$$

At last, we note that when $\hat{\mathrm{var}}\big(\epsilon_{t-1+(j-1)/m}\big)$ is given by (32), it holds that

$$\frac{1}{T}\sum_{t=1}^{T}\Big[\hat{\mathrm{var}}\big(\epsilon_{t-1+(j-1)/m}\big) + \hat{\mathrm{var}}\big(\epsilon_{t-1+j/m}\big)\Big] = O_P\Big(\big(n/m\big)^{-1/2}T^{-1/2}\Big).$$

Hence, (31) reduces to:

$$\hat{\sigma}^2_{u,s} = \sigma^2_{u,s} + O_P\big(mT^{-1/2}\big) + O_P\big(m^{3/2}n^{-1/2}T^{-1/2}\big),$$

which completes the proof. ■

References

Aït-Sahalia, Y., and J. Jacod, 2012a, "Analyzing the spectrum of asset returns: Jump and volatility components in high frequency data," *Journal of Economic Literature*, 50(4), 1007–1050.

——— , 2012b, "Is Brownian motion necessary to model high-frequency data?," *Annals of Statistics*, 38(5), 3093–3128.

Aït-Sahalia, Y., J. Jacod, and J. Li, 2012, "Testing for jumps in noisy high frequency data," *Journal of Econometrics*, 168(2), 207–222.

Aït-Sahalia, Y., and D. Xiu, 2015, "Increased correlation among asset classes: Are volatility or jumps to blame, or both?," *Journal of Econometrics*, 194(2), 205–219.

Aït-Sahalia, Y., and J. Yu, 2009, "High frequency market microstructure noise estimates and liquidity measures," *Annals of Applied Statistics*, 3(1), 422–457.

Andersen, T. G., and T. Bollerslev, 1997, "Intraday periodicity and volatility persistence in financial markets," *Journal of Empirical Finance*, 4(2), 115–158.

——— , 1998, "Deutsche Mark-Dollar volatility: Intraday activity patterns, macroeconomic announcements, and longer run dependencies," *Journal of Finance*, 53(1), 219–265.

Andersen, T. G., T. Bollerslev, and A. Das, 2001, "Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns," *Journal of Finance*, 56(1), 305–327.

Andersen, T. G., T. Bollerslev, and D. Dobrev, 2007, "No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications," *Journal of Econometrics*, 138(1), 125–180.

Andersen, T. G., D. Dobrev, and E. Schaumburg, 2012, "Jump-robust volatility estimation using nearest neighbour truncation," *Journal of Econometrics*, 169(1), 75–93.

——— , 2014, "A robust neighborhood truncation approach to estimation of integrated quarticity," *Econometric Theory*, 30(1), 3–59.

Baeumer, B., and M. M. Meerschaert, 2010, "Tempered stable Lévy motion and transient super-diffusion," *Journal of Computational and Applied Mathematics*, 233(10), 2438–2448.

Bandi, F. M., and J. R. Russell, 2006, "Separating microstructure noise from volatility," *Journal of Financial Economics*, 79(3), 655–692.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard, 2006, "A central limit theorem for realized power and bipower variations of continuous semimartingales," in *From Stochastic Calculus to Mathematical Finance: The Shiryaev Festschrift*, ed. by Y. Kabanov, R. Lipster, and J. Stoyanov. Springer-Verlag, Heidelberg, pp. 33–68.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard, 2008, "Designing realized kernels to measure the ex post variation of equity prices in the presence of noise," *Econometrica*, 76(6), 1481–1536.

——— , 2009, "Realized kernels in practice: trades and quotes," *Econometrics Journal*, 12(3), 1–32.

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

——— , 2006, "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics*, 4(1), 1–30.

Boudt, K., C. Croux, and S. Laurent, 2011, "Robust estimation of intraweek periodicity in volatility and jump detection," *Journal of Empirical Finance*, 18(2), 353–367.

Brownless, C. T., and G. M. Gallo, 2006, "Financial econometric analysis at ultra-high frequency: Data handling concerns," *Computational Statistics and Data Analysis*, 51(4), 2232–2245.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen, 2003, "Alternative models for stock price dynamics," *Journal of Econometrics*, 116(1-2), 225–257.

Christensen, K., S. Kinnebrock, and M. Podolskij, 2010, "Pre-averaging estimators of the ex-post covariance matrix in noisy diffusion models with non-synchronous data," *Journal of Econometrics*, 159(1), 116–133.

Christensen, K., R. C. A. Oomen, and M. Podolskij, 2014, "Fact or friction: Jumps at ultra high frequency," *Journal of Financial Economics*, 114(3), 576–599.

Christensen, K., M. Podolskij, N. Thamrongrat, and B. Veliyev, 2016, "Inference from high-frequency data: A subsampling approach," *Journal of Econometrics*, 197(2), 245–272.

Christensen, K., M. Podolskij, and M. Vetter, 2013, "On covariation estimation for multivariate continuous Itô semimartingales with noise in non-synchronous observation schemes," *Journal of Multivariate Analysis*, 120(1), 59–84.

Clinet, S., and Y. Potiron, 2017, "Efficient asymptotic variance reduction when estimating volatility in high frequency data," Working paper, University of Tokyo.

Corsi, F., D. Pirino, and R. Renò, 2010, "Threshold bipower variation and the impact of jumps on volatility forecasting," *Journal of Econometrics*, 159(2), 276–288.

Delbaen, F., and W. Schachermayer, 1994, "A general version of the fundamental theorem of asset pricing," *Mathematische Annalen*, 300(1), 463–520.

Dette, H., V. Golosnoy, and J. Kellermann, 2016, "The effect of intraday periodicity on realized volatility measures," Working paper, Ruhr-Universität Bochum.

Dette, H., and M. Podolskij, 2008, "Testing the parametric form of the volatility in continuous time diffusion models–A stochastic process approach," *Journal of Econometrics*, 143(1), 56–73.

Dette, H., M. Podolskij, and M. Vetter, 2006, "Estimation of integrated volatility in continuous-time financial models with applications to goodness-of-fit testing," *Scandinavian Journal of Statistics*, 33(2), 259–278.

Dovonon, P., S. Gonçalves, U. Hounyo, and N. Meddahi, 2014, "Bootstrapping high-frequency jump tests," Working paper, CREATES, Aarhus University.

Dudek, A. E., J. Leśkow, E. Paparoditis, and D. N. Politis, 2014, "A generalized block bootstrap for seasonal time series," *Journal of Time Series Analysis*, 35(2), 89–114.

Engle, R. F., and M. E. Sokalska, 2012, "Forecasting intraday volatility in the US equity market. Multiplicative component GARCH," *Journal of Financial Econometrics*, 10(1), 54–83.

Fama, E. F., 1965, "The behavior of stock-market prices," *Journal of Business*, 38(1), 34–105.

Gonçalves, S., U. Hounyo, and N. Meddahi, 2014, "Bootstrap inference for pre-averaged realized volatility based on non-overlapping returns," *Journal of Financial Econometrics*, 12(4), 679–707.

Gonçalves, S., and N. Meddahi, 2009, "Bootstrapping realized volatility," *Econometrica*, 77(1), 283–306.

Hansen, P. R., and A. Lunde, 2006, "Realized variance and market microstructure noise," *Journal of Business and Economic Statistics*, 24(2), 127–161.

Harris, L., 1986, "A transaction data study of weekly and intradaily patterns in stock returns," *Journal of Financial Economics*, 16(1), 99–117.

Hasbrouck, J., 1999, "The dynamics of discrete bid and ask quotes," *Journal of Finance*, 54(6), 2109–2142.

Hautsch, N., and M. Podolskij, 2013, "Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence," *Journal of Business and Economic Statistics*, 31(2), 165–183.

Hecq, A., S. Laurent, and F. C. Palm, 2012, "Common intraday periodicity," *Journal of Financial Econometrics*, 10(2), 325–353.

Hounyo, U., 2014, "The wild tapered block bootstrap," Working paper, CREATES, Aarhus University.

——— , 2017, "Bootstrapping integrated covariance matrix estimators in noisy jump-diffusion models with non-synchronous trading," *Journal of Econometrics*, 197(1), 130–152.

Hounyo, U., S. Gonçalves, and N. Meddahi, 2017, "Bootstrapping pre-averaged realized volatility under market microstructure noise," *Econometric Theory*, 33(4), 791–838.

Huang, X., and G. Tauchen, 2005, "The relative contribution of jumps to total price variance," *Journal of Financial Econometrics*, 3(4), 456–499.

Jacod, J., 2008, "Asymptotic properties of realized power variations and related functionals of semimartingales," *Stochastic Processes and their Applications*, 118(4), 517–559.

Jacod, J., Y. Li, P. A. Mykland, M. Podolskij, and M. Vetter, 2009, "Microstructure noise in the continuous case: The pre-averaging approach," *Stochastic Processes and their Applications*, 119(7), 2249–2276.

Jacod, J., M. Podolskij, and M. Vetter, 2010, "Limit theorems for moving averages of discretized processes plus noise," *Annals of Statistics*, 38(3), 1478–1545.

Jacod, J., and P. E. Protter, 2012, *Discretization of Processes*. Springer-Verlag, Berlin, 2nd edn.

Jacod, J., and V. Todorov, 2009, "Testing for common arrival of jumps in discretely-observed multidimensional processes," *Annals of Statistics*, 37(4), 1792–1838.

Jiang, G. J., I. Lo, and A. Verdelhan, 2011, "Information shocks, liquidity shocks, jumps, and price discovery: Evidence from the U.S. Treasury market," *Journal of Financial and Quantitative Analysis*, 46(2), 527–551.

Jing, B.-Y., Z. Liu, and X.-B. Kong, 2014, "On the estimation of integrated volatility with jumps and microstructure noise," *Journal of Business and Economic Statistics*, 32(3), 457–467.

Kalnina, I., 2011, "Subsampling high frequency data," *Journal of Econometrics*, 161(2), 262–283.

Kalnina, I., and O. Linton, 2008, "Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error," *Journal of Econometrics*, 147(1), 47–59.

Katz, M. L., 1963, "Note on the Berry-Esseen theorem," *Annals of Mathematical Statistics*, 34(3), 1107–1108.

Kolokolov, A., and R. Renò, 2016, "Efficient multipowers," Working paper, University of Verona.

——— , 2017, "Jumps or flatness?," Working paper, University of Verona.

Lee, S. S., and P. A. Mykland, 2008, "Jumps in financial markets: A new nonparametric test and jump dynamics," *Review of Financial Studies*, 21(6), 2535–2563.

Li, J., V. Todorov, and G. Tauchen, 2013, "Volatility occupation times," *Annals of Statistics*, 41(4), 1865–1891.

——— , 2016, "Estimating the volatility occupation time via regularized Laplace inversion," *Econometric Theory*, 32(5), 1253–1288.

Liu, R. Y., 1988, "Bootstrap procedures under some non-i.i.d. models," *Annals of Statistics*, 16(4), 1696–1708.

Mammen, E., 1993, "Bootstrap and wild bootstrap for high dimensional linear models," *Annals of Statistics*, 21(1), 255–285.

Mancini, C., 2009, "Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps," *Scandinavian Journal of Statistics*, 36(2), 270–296.

Mancini, C., and F. Gobbi, 2012, "Identifying the Brownian covariation from the co-jumps given discrete observations," *Econometric Theory*, 28(2), 249–273.

Mandelbrot, B. B., 1963, "The variation of certain speculative prices," *Journal of Business*, 36(4), 394–419.

Mykland, P. A., and L. Zhang, 2009, "Inference for continuous semimartingales observed at high frequency: A general approach," *Econometrica*, 77(5), 1403–1445.

——— , 2017, "Assessment of uncertainty in high frequency data: The observed asymptotic variance," *Econometrica*, 85(1), 197–231.

Pauly, M., 2011, "Weighted resampling of martingale difference arrays with applications," *Electronic Journal of Statistics*, 5(1), 41–52.

Podolskij, M., and M. Vetter, 2009a, "Bipower-type estimation in a noisy diffusion setting," *Stochastic Processes and their Applications*, 119(9), 2803–2831.

——— , 2009b, "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," *Bernoulli*, 15(3), 634–658.

——— , 2010, "Understanding limit theorems for semimartingales: A short survey," *Statistica Neerlandica*, 64(3), 329–351.

Podolskij, M., and K. Wasmuth, 2013, "Goodness-of-fit testing for fractional diffusions," *Statistical Inference for Stochastic Processes*, 16(2), 147–159.

Politis, D. N., J. P. Romano, and M. Wolf, 1999, *Subsampling*, vol. 1. Springer-Verlag.

Shao, X., 2010, "The dependent wild bootstrap," *Journal of the American Statistical Association*, 105(489), 218–235.

Taylor, S. J., and X. Xu, 1997, "The incremental volatility information in one million foreign exchange quotations," *Journal of Empirical Finance*, 4(4), 317–340.

Todorov, V., and T. Bollerslev, 2010, "Jumps and betas: A new framework for disentangling and estimating systematic risks," *Journal of Econometrics*, 157(2), 220–235.

Todorov, V., and G. Tauchen, 2010, "Activity signature functions for high-frequency data analysis," *Journal of Econometrics*, 154(2), 125–138.

——— , 2011, "Volatility jumps," *Journal of Business and Economic Statistics*, 29(3), 356–371.

——— , 2012, "The realized Laplace transform of volatility," *Econometrica*, 80(3), 11051127.

Todorov, V., G. Tauchen, and I. Grynkiv, 2014, "Volatility activity: Specification and estimation," *Journal of Econometrics*, 178(1), 180–193.

Vetter, M., 2008, "Estimation methods in noisy diffusion models," Ph.D. thesis, Ruhr-Universität Bochum.

——— , 2010, "Limit theorems for bipower variation of semimartingales," *Stochastic Processes and their Applications*, 120(1), 22–38.

Vetter, M., and H. Dette, 2012, "Model checks for the volatility under microstructure noise," *Bernoulli*, 18(4), 1421–1447.

Wood, R. A., T. H. McInish, and J. K. Ord, 1985, "An investigation of transactions data for NYSE stocks," *Journal of Finance*, 40(3), 723–739.

Wu, C. F. J., 1986, "Jackknife, bootstrap and other resampling methods in regression analysis," *Annals of Statistics*, 14(4), 1261–1295.

Xiu, D., 2010, "Quasi-maximum likelihood estimation of volatility with high frequency data," *Journal of Econometrics*, 159(1), 235250.

Zhang, L., 2006, "Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach," *Bernoulli*, 12(6), 1019–1043.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia, 2005, "A tale of two time scales: determining integrated volatility with noisy high-frequency data," *Journal of the American Statistical Association*, 100(472), 1394–1411.

Research Papers 2017

2017-13: Niels S. Grønborg, Asger Lunde, Allan Timmermann and Russ Wermers: Picking Funds with Confidence

2017-14: Martin M. Andreasen and Anders Kronborg: The Extended Perturbation Method: New Insights on the New Keynesian Model

2017-15: Andrea Barletta, Paolo Santucci de Magistris and Francesco Violante: A Non-Structural Investigation of VIX Risk Neutral Density

2017-16: Davide Delle Monache, Stefano Grassi and Paolo Santucci de Magistris: Does the ARFIMA really shift?

2017-17: Massimo Franchi and Søren Johansen: Improved inference on cointegrating vectors in the presence of a near unit root using adjusted quantiles

2017-18: Matias D. Cattaneo, Michael Jansson and Kenichi Nagasawa: Bootstrap-Based Inference for Cube Root Consistent Estimators

2017-19: Daniel Borup and Martin Thyrsgaard: Statistical tests for equal predictive ability across multiple forecasting methods

2017-20: Tommaso Proietti and Alessandro Giovannelli: A Durbin-Levinson Regularized Estimator of High Dimensional Autocovariance Matrices

2017-21: Jeroen V.K. Rombouts, Lars Stentoft and Francesco Violante: Variance swap payoffs, risk premia and extreme market conditions

2017-22: Jakob Guldbæk Mikkelsen: Testing for time-varying loadings in dynamic factor models

2017-23: Roman Frydman, Søren Johansen, Anders Rahbek and Morten Nyboe Tabor: The Qualitative Expectations Hypothesis: Model Ambiguity, Concistent Representations of Market Forecasts, and Sentiment

2017-24: Giorgio Mirone: Inference from the futures: ranking the noise cancelling accuracy of realized measures

2017-25: Massimiliano Caporin, Gisle J. Natvik, Francesco Ravazzolo and Paolo Santucci de Magistris: The Bank-Sovereign Nexus: Evidence from a non-Bailout Episode

2017-26: Mikkel Bennedsen, Asger Lunde and Mikko S. Pakkanen: Decoupling the short- and long-term behavior of stochastic volatility

2017-27: Martin M. Andreasen, Jens H.E. Christensen and Simon Riddell: The TIPS Liquidity Premium

2017-28: Annastiina Silvennoinen and Timo Teräsvirta: Consistency and asymptotic normality of maximum likelihood estimators of a multiplicative time-varying smooth transition correlation GARCH model

2017-29: Cristina Amado, Annastiina Silvennoinen and Timo Teräsvirta: Modelling and forecasting WIG20 daily returns

2017-30: Kim Christensen, Ulrich Hounyo and Mark Podolskij: Is the diurnal pattern sufficient to explain the intraday variation in volatility? A nonparametric assessmentfafa