

Testing for heteroscedasticity in jumpy and noisy high-frequency data: A resampling approach*

Kim Christensen[†] Ulrich Hounyo[†] Mark Podolskij^{‡,†}

August, 2016

Abstract

In this paper, we propose a new way to measure and test the presence of time-varying volatility in a discretely sampled jump-diffusion process that is contaminated by microstructure noise. We use the concept of pre-averaged truncated bipower variation to construct our t -statistic, which diverges in the presence of a heteroscedastic volatility term (and has a standard normal distribution otherwise). The test is inspected in a general Monte Carlo simulation setting, where we note that in finite samples the asymptotic theory is severely distorted by infinite-activity price jumps. To improve inference, we suggest a bootstrap approach to test the null of homoscedasticity. We prove the first-order validity of this procedure, while in simulations the bootstrap leads to almost correctly sized tests. As an illustration, we apply the bootstrapped version of our t -statistic to a large cross-section of equity high-frequency data. We document the importance of jump-robustness, when measuring heteroscedasticity in practice. We also find that a large fraction of variation in intraday volatility is accounted for by seasonality. This suggests that, once we control for jumps and deflate asset returns by a non-parametric estimate of the conventional U-shaped diurnality profile, the variance of the rescaled return series is often close to constant within the day.

JEL Classification: C10; C80.

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

*The authors received research funding from the Danish Council for Independent Research (DFF – 4182-00050) and 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

Asset return volatility is central in financial economics with numerous applications including, for example, the construction of optimal portfolios, risk management, pricing and hedging of derivatives (see, e.g., Black and Scholes, 1973; Markowitz, 1952, among many others). Therefore, accurate modeling of the volatility dynamics and good estimates of this quantity are of particular importance in finance. For example, the pricing of options is crucially dependent on the functional form of volatility (e.g., Black and Scholes, 1973; Duffie and Harrison, 1993).

In this paper, we propose a procedure to measure the strength and test for the presence of time-varying volatility (based on a bootstrapping approach) in a *non-parametric* framework. We model the log-price of a financial asset as a general arbitrage-free Itô semimartingale, which does not necessarily display continuous sample paths. The process is contaminated by a market microstructure noise component and is observed on a fixed time interval with the distance between observations tending to zero.

In the literature, the *parametric* form of volatility tends to vary widely (see, e.g., Vasicek, 1977; Cox, Ingersoll, and Ross, 1985; Constantinides, 1992; Duffie and Kan, 1996, among others). Hence, researchers and practitioners typically rely on goodness-of-fit tests to check the postulated model (see, e.g., Aït-Sahalia, 1996; Corradi and White, 1999; Corradi and Distaso, 2006; Dette, Podolskij, and Vetter, 2006; Dette and Podolskij, 2008; Vetter and Dette, 2012). Of course, it is widely documented that time series of asset returns display time-varying volatility. At the high-frequency level, for instance, intraday seasonality in volatility is a well-known stylized feature, see Wood, McInish, and Ord (1985); Harris (1986); Andersen and Bollerslev (1997, 1998). Still, it is useful to be able to test whether volatility is constant. The main motivation, apart from the importance of the correct specification of the dynamics of volatility is that in practice, the volatility of asset returns has components which are highly persistent, especially over short horizons. Thus, volatility could be at least locally nearly constant. In particular, as we show in this paper, there is some support for this feature on a daily basis, after the returns series has been corrected for the above diurnal variation.

A visual inspection of financial time series does not provide clear evidence for either the presence or the absence of homoscedasticity over a short horizon, in contrast to the properties of asset returns over longer horizons. In concrete applications, the problem is complicated by the existence of market microstructure noise as well as the impact of jumps, i.e., the possible occurrence of discontinuous movements in prices. The so-called noise is due to the presence of, for example, bid-ask bounce effects or rounding errors, which contribute to a discrepancy between the latent efficient price process and the price observed by the econometrician. Hence, it is important for financial economists to have statistical tools that can shed some light on such issue. Recent papers dealing with the separation of the diffusion component from the jump part include Aït-Sahalia (2004); Barndorff-Nielsen and Shephard (2004, 2006); Huang and Tauchen (2005); Mancini (2009);

Todorov and Bollerslev (2010), while the issue of the effect of microstructure noise on estimators of volatility has been analyzed by, for example, Zhang, Mykland, and Ait-Sahalia (2005); Zhang (2006); Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008); Jacod, Li, Mykland, Podolskij, and Vetter (2009); Xiu (2010), while related work for the construction of jump- and noise-robust estimates of volatility or jump hypothesis testing include Podolskij and Vetter (2009a); Christensen, Oomen, and Podolskij (2010); Ait-Sahalia, Jacod, and Li (2012); Lee and Mykland (2012); Jing, Liu, and Kong (2014). In the literature, several tests for homoscedasticity in the context of financial high-frequency data have been proposed, based on different estimation procedures, e.g., Corradi and White (1999); Dette, Podolskij, and Vetter (2006); Dette and Podolskij (2008); Vetter and Dette (2012). To the best of our knowledge, none of the existing tests of homoscedasticity allow for joint effect of jumps and market microstructure noise in a completely non-parametric setting.

First, we use the concept of pre-averaged truncated bipower variation to construct our t -statistic, which diverges in the presence of a heteroscedastic volatility term (and has a standard normal distribution otherwise). Second, we assess by Monte Carlo simulation the accuracy of the proposed test, where we note that in finite samples the asymptotic theory is severely distorted by infinite-activity price jumps. Third, to improve inference, we suggest a bootstrap approach to test the null of homoscedasticity. Our proposed bootstrap method is new and is of independent interest. It can be viewed as the overlapping version of the wild blocks of blocks bootstrap studied recently by Hounyo, Gonçalves, and Meddahi (2015). Hence, we name this method “overlapping wild blocks of blocks bootstrap”. Our heteroscedasticity test statistics depends on the product of truncated bipower pre-averaged returns to which we apply the proposed overlapping wild blocks of blocks bootstrap. As in Hounyo, Gonçalves, and Meddahi (2015), our resampling method combines the wild bootstrap with the blocks of blocks bootstrap and then is able to capture elegantly the dependence and heterogeneity properties of the product of truncated bipower pre-averaged returns. We prove the first-order validity of this procedure, while in simulations the bootstrap leads to almost correctly sized tests. Finally, we complement the theory and Monte Carlo simulations results by conducting an empirical application. In particular, we apply bootstrapped version of our test of heteroscedasticity to a large cross-sectional panel of US equity data. It includes the 30 stocks of the Dow Jones Industrial Average index. We show that it is important to control for the jump part in such assessment. Otherwise researchers may often conclude by error that the volatility is heteroscedastic, whereas in reality it could not be necessarily the case over a short period horizon. We also find that a large fraction of variation in intraday volatility is accounted for by seasonality. This suggests that, once we control for jumps and deflate asset returns by a non-parametric estimate of the conventional U-shaped diurnality pattern, the volatility of the rescaled return series is often close to constant within the day.

The remainder of the paper is organized as follows. The next section briefly introduces the theoretical framework, and the main assumptions. We also review the existing asymptotic theory of the pre-averaged bipower-type estimator and state new results; in particular, we construct a jump-

and noise-robust test for absence of heteroscedasticity. In Section 3, we introduce the bootstrap method and show its consistency when testing homoscedasticity in noisy jump-diffusion setting. In Section 4, we present the Monte Carlo results, while an empirical illustration is conducted in Section 5. Section 6 concludes. All proofs and some auxiliary results are relegated to the Appendix.

2 Theoretical setup

Let X denote the latent efficient log-price defined on a probability space (Ω, \mathcal{F}, P) equipped with a filtration $(\mathcal{F}_t)_{t \geq 0}$. We model X as an Itô semimartingale process defined by the equation

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s + (\delta 1_{\{|\delta| \leq 1\}}) \star (\underline{\mu}_t - \underline{\nu}_t) + (\delta 1_{\{|\delta| > 1\}}) \star \underline{\mu}_t, \quad (1)$$

where $(a_t)_{t \geq 0}$ is a predictable, locally bounded drift process, $(\sigma_t)_{t \geq 0}$ is an adapted, càdlàg (i.e., a right-continuous process with limits from the left) volatility process, while $(W_t)_{t \geq 0}$ is a Brownian motion. Also, $\underline{\mu}$ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ and $\underline{\nu}$ is a predictable compensator of $\underline{\mu}$, such that $\underline{\nu}(ds, dx) = ds \otimes \lambda(dx)$, where λ is a σ -finite measure.

We further assume that:

Assumption (J): There exists a sequence of stopping times $(\tilde{\tau}_n)_{n=1}^\infty$ increasing to ∞ and a deterministic nonnegative function $\tilde{\gamma}_n$ such that $\int_{\mathbb{R}} \tilde{\gamma}_n(x)^\beta \lambda(dx) < \infty$ 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 $\beta \in [0, 2]$.

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

We rule out jumps in σ_t via the following assumption:

Assumption (V): σ_t is of the form:

$$\sigma_t = \sigma_0 + \int_0^t \tilde{a}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{\nu}_s dB_s, \quad (2)$$

where $(\tilde{a}_t)_{t \geq 0}$, $(\tilde{\sigma}_t)_{t \geq 0}$ and $(\tilde{\nu}_t)_{t \geq 0}$ 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., Equation (3) in Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008); Assumption 2 in Mykland and Zhang (2009), or Equation (3) in Gonçalves and Meddahi (2009)). It can be relaxed (see, for example, Assumption H1 in Barndorff-Nielsen, Graversen, Jacod, Podolskij, and Shephard (2006) for a weaker set of assumptions, which allow for jumps in σ_t).

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

Assumption (V’): $\sigma_t > 0$, for all $t \geq 0$.

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

$$Y_t = X_t + \epsilon_t, \quad (3)$$

where ϵ_t is a noise term that collects the market microstructure effects. In this paper, we assume that ϵ_t is i.i.d. and independent of X_t , such that

$$E(\epsilon_t) = 0 \quad \text{and} \quad E(\epsilon_t^2) = \omega^2, \quad (4)$$

for any t , where Y_t is observed.¹ Here, we follow Podolskij and Vetter (2009a) and assume that

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

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

2.2 Test of heteroscedasticity

As stated above, in this paper we develop a test of the “no heteroscedasticity” assumption. To achieve this, we partition the sample space Ω into the following two subsets:

$$\Omega_{\mathcal{H}_0} = \{\omega : \sigma_t \text{ is constant for } t \geq 0\}, \quad (5)$$

and $\Omega_{\mathcal{H}_a} = \Omega_{\mathcal{H}_0}^c$. 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 ω belongs to $\Omega_{\mathcal{H}_0}$. The key challenge we address is how to construct such a test, when X —apart from being driven by a Brownian component—is observed with measurement error and potentially discontinuous. The solution is based on computing a set of estimators, which reveal information about the presence of time-variation in the volatility process.

¹In some places, we can assume that the conditional variance of ϵ_t , i.e. $\omega_t^2 = E(\epsilon_t^2 | X)$ is càdlàg with a moment condition, as in Jacod, Podolskij, and Vetter (2010). Please refer to Remark 2.

In particular, we study the quadratic variation of X , which is defined by:

$$[X]_t = \int_0^t \sigma_s^2 ds + \sum_{s \leq t} |\Delta X_s|^2, \quad (6)$$

where $\int_0^t \sigma_s^2 ds$ —also known as the integrated variance—is the quadratic variation of the continuous part of X (i.e., when X follows (80)), while $\sum_{s \leq t} |\Delta X_s|^2$ is the sum of the squared jumps, where $\Delta X_s = X_s - X_{s-}$.

We note that if volatility is a constant, say σ , (1) reduces to

$$X_t = X_0 + \int_0^t a_s ds + \sigma(W_t - W_0) + (\delta 1_{\{|\delta| \leq 1\}}) \star (\underline{\mu}_t - \underline{\nu}_t) + (\delta 1_{\{|\delta| > 1\}}) \star \underline{\mu}_t, \quad (7)$$

while

$$[X]_t = \sigma^2 t + \sum_{s \leq t} |\Delta X_s|^2. \quad (8)$$

The construction of the t -statistic now progresses in two steps. Firstly, we account for microstructure noise by doing local pre-averaging of Y . Secondly, we tease out the continuous part of the quadratic variation by suitable removal of the jump component.

2.3 The pre-averaging approach

In the rest of this paper, without loss of generality, we confine the clock to the unit interval $t \in [0, 1]$. We assume that the noisy log-price Y_t is observed at the regular time points $t_i = i/n$, for $i = 0, \dots, n$. Then, the intraday log-returns (at frequency n) can be computed as:

$$\Delta_i^n Y \equiv Y_{i/n} - Y_{(i-1)/n}, \quad i = 1, \dots, n. \quad (9)$$

As $Y_t = X_t + \epsilon_t$, we can split $\Delta_i^n Y$ into

$$\Delta_i^n Y = \Delta_i^n X + \Delta_i^n \epsilon, \quad (10)$$

where $\Delta_i^n X = X_{i/n} - X_{(i-1)/n}$ denotes the n -frequency return of the efficient log-price, while $\Delta_i^n \epsilon = \epsilon_{i/n} - \epsilon_{(i-1)/n}$ 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 $g(0) = g(1) = 0$ and $\int_0^1 g(s)^2 ds > 0$. 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:

$$\phi_1(s) = \int_s^1 g'(u)g'(u-s)du \quad \text{and} \quad \phi_2(s) = \int_s^1 g(u)g(u-s)du, \quad (11)$$

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 $\psi_2 = 1/12$.

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

$$\Delta_i^n \bar{Y} = \sum_{j=1}^{k_n-1} g\left(\frac{j}{k_n}\right) \Delta_{i+j-1}^n Y, \quad i = 1, \dots, n - k_n + 2. \quad (12)$$

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 $(\Delta_i^n \bar{Y})_{i=1}^{n-k_n+2}$. It follows from the decomposition in (10) that $\Delta_i^n \bar{Y} = \Delta_i^n \bar{X} + \Delta_i^n \bar{\epsilon}$ and, as shown by Vetter (2008),

$$\Delta_i^n \bar{X} = O_p\left(\sqrt{\frac{k_n}{n}}\right) \quad \text{and} \quad \Delta_i^n \bar{\epsilon} = O_p\left(\frac{1}{\sqrt{k_n}}\right). \quad (13)$$

Thus, the noise is dampened, thereby reducing its influence on $\Delta_i^n \bar{Y}$. As an outcome, we retrieve a basically noise-free estimate, which can substitute the efficient log-return $\Delta_i^n X$ in subsequent computations, taking proper account of the dependence introduced in $(\Delta_i^n \bar{Y})_{i=1}^{n-k_n+2}$.² 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 $n^{-1/4}$ rate of convergence, Jacod, Li, Mykland, Podolskij, and Vetter (2009) propose to set:

$$k_n = \theta\sqrt{n} + o(n^{-1/4}), \quad (14)$$

for some $\theta \in (0, \infty)$. With this choice, the orders of the two terms $\Delta_i^n \bar{X}$ and $\Delta_i^n \bar{\epsilon}$ are balanced and equal to $O_p(n^{-1/4})$. An example of (14) used throughout this paper is $k_n = \lceil \theta\sqrt{n} \rceil$.

2.3.1 The pre-averaged bipower variation

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

$$BV(Y, l, r)^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{N_n} y(Y, l, r)_i^n, \quad (15)$$

where $l, r \geq 0$, $y(Y, l, r)_i^n = |\Delta_{i-1}^n \bar{Y}|^l |\Delta_{i-1+k_n}^n \bar{Y}|^r$, $N_n = n - 2k_n + 2$ and $\mu_p = E(|N(0, 1)|^p)$.³ 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 . Podolskij and Vetter (2009a) show that under suitable regularity conditions, in particular

²If k_n is even, it follows with the above definition of $g(x) = \min(x, 1 - x)$ that the pre-averaged returns in (12) can be rewritten as $\Delta_i^n \bar{Y} = \frac{1}{k_n} \sum_{j=1}^{k_n/2} Y_{i+k_n/2+j} - \frac{1}{k_n} \sum_{j=1}^{k_n/2} Y_{i+j}$. Thus, the sequence $(2\Delta_i^n \bar{Y})_{i=1}^{n-k_n+2}$ can be interpreted as a constituting a new set of increments from a price process that is constructed by simple averaging of the noisy log-price series, $(Y_{i/n})_{i=1}^n$, in a neighbourhood of i/n , thus making the use of the term pre-averaging and the associated notation transparent.

³In 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.

that X is a continuous Itô semimartingale (i.e., X follows (80)), then as $n \rightarrow \infty$

$$BV(l, r)^n \xrightarrow{p} BV(l, r) = \int_0^1 \left(\theta \psi_2 \sigma_s^2 + \frac{1}{\theta} \psi_1 \omega^2 \right)^{\frac{l+r}{2}} ds, \quad (16)$$

and

$$n^{1/4} \begin{pmatrix} BV(l_1, r_1)^n - BV(l_1, r_1) \\ BV(l_2, r_2)^n - BV(l_2, r_2) \end{pmatrix} \xrightarrow{st} MN(0, \Sigma), \quad (17)$$

with $l_1, r_1, l_2, r_2 \geq 0$, where “ \xrightarrow{st} ” denotes stable convergence in law, and $\Sigma = (\Sigma_{ij}^{l_1, r_1, l_2, r_2})_{1 \leq i, j \leq 2}$ is the conditional covariance matrix of the limiting process $n^{1/4}(BV(l_1, r_1)^n, BV(l_2, r_2)^n)'$.⁴

2.3.2 A truncated pre-averaged bipower variation

The estimator in (15) 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 (17) 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 (16), 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 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 (16) and (17) in the general setting, where X follows the Itô semimartingale in (1). Thus, we propose to set:

$$\check{B}V(l, r)^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{N_n} \check{y}(l, r)_i^n, \quad (18)$$

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

$$v_n = \alpha u_n^\varpi, \text{ for } \alpha > 0 \text{ and } \varpi \in (0, 1/2), \quad (19)$$

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 $\left(\frac{l_1+r_1-1}{2(l_1+r_1-\beta)} \vee \frac{l_2+r_2-1}{2(l_2+r_2-\beta)} \right) \leq$*

⁴The formal definition of Σ is given in Appendix A.

$\varpi < 1/2$. Furthermore, we assume (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 $k_n \rightarrow \infty$ as $n \rightarrow \infty$ such that (14) holds. Then, as $n \rightarrow \infty$,

$$n^{1/4} \begin{pmatrix} \check{B}\check{V}(l_1, r_1)^n - BV(l_1, r_1) \\ \check{B}\check{V}(l_2, r_2)^n - BV(l_2, r_2) \end{pmatrix} \xrightarrow{st} MN(0, \Sigma). \quad (20)$$

Theorem 2.1 shows that (18) is robust to the jump part in its limiting distribution. Note that Σ is identical to the matrix in (17). To our knowledge, the result is new with the main innovations being the statistic is (18) 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 (17), 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 $\check{B}\check{V}(2, 0)^n$, when X follows (1). Our paper generalizes the latter article to the bipower setting with—subject to the above constraint—arbitrary powers.

This enables extraction of an essentially noise-free and jump-robust estimate of the continuous piece of the quadratic variation in (6) and thus facilitates the construction of a test for the presence of time-variation in σ_t . An implication of (20) 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 $l_1 + r_1 \neq l_2 + r_2$, as $n \rightarrow \infty$,

$$\begin{aligned} & \check{B}\check{V}(l_1, r_1)^n - (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}} \xrightarrow{p} BV(l_1, r_1) - (BV(l_2, r_2))^{\frac{l_1+r_1}{l_2+r_2}} \\ & = \int_0^1 \left(\theta \psi_2 \sigma_s^2 + \frac{1}{\theta} \psi_1 \omega^2 \right)^{\frac{l_1+r_1}{2}} ds - \left[\int_0^1 \left(\theta \psi_2 \sigma_s^2 + \frac{1}{\theta} \psi_1 \omega^2 \right)^{\frac{l_2+r_2}{2}} ds \right]^{\frac{l_1+r_1}{l_2+r_2}} \geq 0, \end{aligned} \quad (21)$$

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

$$T_{\text{inf}}^n = \frac{n^{1/4} \left(\check{B}\check{V}(l_1, r_1)^n - (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}} \right)}{\sqrt{V}} \xrightarrow{d} N(0, 1), \quad (22)$$

where

$$V = \Sigma_{11} - 2 \left(\frac{l_1 + r_1}{l_2 + r_2} \right) (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}-1} \Sigma_{12} + \left(\frac{l_1 + r_1}{l_2 + r_2} \right)^2 (\check{B}\check{V}(l_2, r_2)^n)^{2 \left(\frac{l_1+r_1}{l_2+r_2}-1 \right)} \Sigma_{22}. \quad (23)$$

Note that the convergence in (22) holds only under \mathcal{H}_0 , while under \mathcal{H}_a it follows from (21) that $n^{1/4} \left(\check{B}\check{V}(l_1, r_1)^n - (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}} \right) \rightarrow \infty$. This way we can determine if X has homoscedastic or heteroscedastic volatility with asymptotically correct size and power tending to one, as $n \rightarrow \infty$. To render the test feasible in practice, we propose a consistent estimator of Σ below, which can be plugged into (23). It is both inherently robust to heteroscedasticity and positive semi-definite.

3 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 4, which reveals that in small samples, the feasible version of (22) (cf. (47)) is poorly approximated by the standard normal. Next, we propose a bootstrap estimator of the conditional covariance matrix of the limiting process $n^{1/4}(\check{B}\check{V}(l_1, r_1)^n, \check{B}\check{V}(l_2, r_2)^n)'$, 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. The first to utilize bootstrapping in this setting were Gonçalves and Meddahi (2009), who 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 (2015) extend their work to accommodate noise. The latter study 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 $(\Delta_i^n \bar{Y})_{i=1}^{n-k_n+2}$ appears natural.

Nevertheless, such a scheme is only consistent, if σ_t is constant. As shown by Hounyo, Gonçalves, and Meddahi (2015), the problem is that $|\Delta_i^n \bar{Y}|^2$ are heterogeneously distributed under time-varying volatility.⁵ 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 $|\Delta_i^n \bar{Y}|^2$, 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 (2015) generalizes Hounyo, Gonçalves, and Meddahi (2015) 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 $\check{B}\check{V}(l, r)^n$ is

$$\check{B}\check{V}(l, r)^{n*} = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{N_n} \check{y}(l, r)_i^{n*}, \quad (24)$$

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 (2015), we merge the wild bootstrap with block-based resampling. However, our bootstrap is new, and it can be viewed as an overlapping version

⁵This feature is highlighted by the asymptotic distribution of $\Delta_i^n \bar{Y}$ in (71) below.

of their algorithm. We name it “the overlapping wild blocks of blocks bootstrap.” We note that 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 $\delta_1 \in (0, 1)$:

$$b_n = O(n^{\delta_1}). \quad (25)$$

We divide the available data into overlapping blocks of size b_n , and the bootstrap is then based on $N_n - 2b_n + 2$ of these. In particular, we look at overlapping blocks of b_n consecutive observations 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 $\check{y}(l, r)_{N_n - b_n + 1}^n, \dots, \check{y}(l, r)_{N_n}^n$. The bootstrap sample is constructed by properly combining the first J_n blocks.

Let $u_1, \dots, u_{J_n + 1}$ be i.i.d. random variables, whose distribution is independent of the original sample. We denote by $\mu_q^* = E^*(u_j^q)$ its q th order moments.⁶ Then,

$$\bar{B}_j = \frac{1}{b_n} \sum_{i=1}^{b_n} \check{y}(l, r)_{i-1+j}^n, \quad j = 1, \dots, N_n - b_n + 1, \quad (26)$$

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

$$\check{y}(l, r)_m^{n*} - \bar{B}^{N_n} = \begin{cases} \frac{1}{\sqrt{b_n}} \sum_{j=1}^m (\check{y}(l, r)_m^n - \bar{B}_{b_n+j}) u_j, & \text{if } m \in I_1^n, \\ \frac{1}{\sqrt{b_n}} \sum_{j=1}^{b_n} (\check{y}(l, r)_m^n - \bar{B}_{m+j}) u_{m+j-b_n}, & \text{if } m \in I_2^n, \\ \frac{1}{\sqrt{b_n}} \sum_{j=1}^{N_n - b_n + 1 - m} (\check{y}(l, r)_m^n - \bar{B}_{J_n + 1 - j + b_n}) u_{J_n + 1 - j}, & \text{if } m \in I_3^n, \\ \frac{1}{\sqrt{b_n}} (\check{y}(l, r)_m^n - \bar{B}_{N_n - b_n + 1}) u_{J_n + 1}, & \text{if } m \in I_4^n, \end{cases} \quad (27)$$

where

$$\bar{B}^{N_n} = \frac{1}{N_n} \sum_{i=1}^{N_n} \check{y}(l, r)_i^n, \quad (28)$$

and

$$\begin{aligned} I_1^n &= \{1, \dots, b_n - 1\}, & I_2^n &= \{b_n, \dots, J_n\}, \\ I_3^n &= \{J_n + 1, \dots, N_n - b_n\}, & I_4^n &= \{N_n - b_n + 1, \dots, N_n\}. \end{aligned} \quad (29)$$

It is interesting to note that if we were to center $\check{y}(l, r)_m^n$ around the grand mean \bar{B}^{N_n} , instead of

⁶As 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 $n \rightarrow \infty$, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{n \rightarrow \infty} P[P^*(|Z_n^*| > \delta) > \varepsilon] = 0$, (ii) $Z_n^* = O_{P^*}(1)$ as $n \rightarrow \infty$, if for all $\varepsilon > 0$ there exists an $M_\varepsilon < \infty$ such that $\lim_{n \rightarrow \infty} P[P^*(|Z_n^*| > M_\varepsilon) > \varepsilon] = 0$, and (iii) $Z_n^* \xrightarrow{d^*} Z$ as $n \rightarrow \infty$, 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.

the localized block average \bar{B}_{j+m} , it would yield a bootstrap observation

$$\check{y}(l, r)_m^{n*} - \bar{B}^{N_n} = \left(\check{y}(l, r)_m^n - \bar{B}^{N_n} \right) \eta_m, \quad (30)$$

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 $\text{var}(u_j) = 1$, we find that $E(\eta_m) = 0$, $\text{var}(\eta_m) = 1$, and $\text{cov}(\eta_m, \eta_{m-k}) = (1 - \frac{k}{b_n}) 1_{\{k \leq b_n\}}$. 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 (2015), 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 σ_t is time-varying. In this paper, we show that generating the bootstrap observations as in (27) does yield an asymptotically valid bootstrap for $(\check{B}\check{V}(l_1, r_1)^n, \check{B}\check{V}(l_2, r_2)^n)'$, even if σ_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 (2015), 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

$$\bar{B}^{N_n*} = \frac{1}{N_n} \sum_{i=1}^{N_n} \check{y}(l, r)_i^{n*}, \quad (31)$$

as the average value of the bootstrap observations. A closer inspection of \bar{B}^{N_n*} suggests that we can rewrite the centered bootstrap sample mean $\bar{B}^{N_n*} - \bar{B}^{N_n}$ as follows

$$N_n \left(\bar{B}^{N_n*} - \bar{B}^{N_n} \right) = \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} b_n (\bar{B}_j - \bar{B}_{j+b_n}) u_j. \quad (32)$$

Thus,

$$\begin{aligned} \check{B}\check{V}(l, r)^{n*} &= \check{B}\check{V}(l, r)^n + n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} b_n (\bar{B}_j - \bar{B}_{j+b_n}) u_j \\ &= \check{B}\check{V}(l, r)^n - \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} \check{\Delta} B(l, r)_j^n u_j, \end{aligned} \quad (33)$$

where

$$\check{\Delta} B(l, r)_j^n = \check{B}(l, r)_{j+b_n}^n - \check{B}(l, r)_j^n, \quad (34)$$

such that

$$\check{B}(l, r)_j^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{b_n} \check{y}(l, r)_{i-1+j}^n. \quad (35)$$

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

Lemma 3.1 *Assume that $\check{y}(l, r)_m^{n*}$ are generated as in (27). Then, it follows that*

$$E^*(\check{B}\check{V}(l, r)^{n*}) = \check{B}\check{V}(l, r)^n - \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} \check{\Delta} B(l, r)_j^n E^*(u_j), \quad (36)$$

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

$$\text{cov}^*(n^{1/4} \check{B}\check{V}(l_i, r_i)^{n*}, n^{1/4} \check{B}\check{V}(l_j, r_j)^{n*}) = \frac{\sqrt{n}}{b_n} \sum_{k=1}^{J_n} \check{\Delta} B(l_i, r_i)_k^n \check{\Delta} B(l_j, r_j)_k^n \text{var}^*(u_k). \quad (37)$$

Equation (36) of Lemma 3.1 implies that with $E^*(u_j) = 0$, $\check{B}\check{V}(l, r)^{n*}$ is an unbiased estimator of $\check{B}\check{V}(l, r)^n$, i.e. $E^*(\check{B}\check{V}(l, r)^{n*}) = \check{B}\check{V}(l, r)^n$. The second part shows that the bootstrap covariance of $n^{1/4} \check{B}\check{V}(l_i, r_i)^{n*}$ and $n^{1/4} \check{B}\check{V}(l_j, r_j)^{n*}$ depends on the variance of u . In particular, if we select $\text{var}^*(u) = 1/2$ as in Hounyo, Gonçalves, and Meddahi (2015):

$$\text{var}^* \left(n^{1/4} \begin{pmatrix} \check{B}\check{V}(l_1, r_1)^{n*} \\ \check{B}\check{V}(l_2, r_2)^{n*} \end{pmatrix} \right) = \check{\Sigma}^n, \quad (38)$$

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

$$\check{\Sigma}_{ij}^{l_i, r_i, l_j, r_j, n} = \frac{\sqrt{n}}{2b_n} \sum_{k=1}^{J_n} \check{\Delta} B(l_i, r_i)_k^n \check{\Delta} B(l_j, r_j)_k^n. \quad (39)$$

Note that based on (39), we can rewrite $\check{\Sigma}^n$ as

$$\check{\Sigma}^n = \frac{\sqrt{n}}{2b_n} \sum_{j=1}^{J_n} \check{\xi}_j \check{\xi}_j', \quad (40)$$

where $\check{\xi}_j \equiv \left(\Delta \check{B}(l_1, r_1)_j^n, \Delta \check{B}(l_2, r_2)_j^n \right)'$. 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 $n^{1/4}(\check{B}\check{V}(l_1, r_1)^n, \check{B}\check{V}(l_2, r_2)^n)'$ provided $\check{\Sigma}^n$ is a consistent estimator of Σ , as proved in Theorem 3.1 below. Note that $\check{\Sigma}^n$ is related to recent work on asymptotic variance estimation by Mykland and Zhang (2016); see also, e.g., Christensen, Podolskij, Thamrongrat, and Veliyev (2016); Jacod and Todorov (2009); Mancini and Gobbi (2012).

Remark 1 *Note that from (40), we can also rewrite $\check{\Sigma}^n$ as follows:*

$$\check{\Sigma}^n = \frac{1}{b_n} \sum_{m=1}^{b_n} \check{\Sigma}_m^n, \quad (41)$$

where

$$\check{\Sigma}_m^n = \frac{\sqrt{n}}{2} \sum_{j=0}^{\lfloor N_n/b_n \rfloor - 2} \check{\xi}_{jb_n+m} \check{\xi}'_{jb_n+m} = \left(\check{\Sigma}_{ij,m}^{l_1, r_1, l_2, r_2, n} \right)_{1 \leq i, j \leq 2}. \quad (42)$$

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 $n^{1/4}\check{B}\check{V}(l_1, r_1)^n$ and $n^{1/4}\check{B}\check{V}(l_2, r_2)^n$, as proposed by Hounyo (2015).

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 3.1 *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 (80)), and also if X has jumps (i.e., X follows (1)) with either*

$$l_1 + r_1 + l_2 + r_2 \leq 4(1 - \delta_1), \quad 0 \leq \beta < 4(1 - \delta_1), \quad (43)$$

or

$$l_1 + r_1 + l_2 + r_2 > 4(1 - \delta_1), \quad 0 \leq \beta < 4(1 - \delta_1), \quad \frac{l_1 + r_1 + l_2 + r_2 - 4(1 - \delta_1)}{2(l_1 + r_1 + l_2 + r_2 - \beta)} \leq \varpi < \frac{1}{2}. \quad (44)$$

Then, as $n \rightarrow \infty$, it holds that

$$\check{\Sigma}^n \xrightarrow{p} \Sigma, \quad (45)$$

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, (44) 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 is therefore sufficient that $\beta \in [0, 4/3]$.

Theorem 3.1 implies that in finite samples, we get a consistent and nonnegative estimator of V :

$$\check{V}^n = \check{\Sigma}_{11}^n - 2 \left(\frac{l_1 + r_1}{l_2 + r_2} \right) (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}-1} \check{\Sigma}_{12}^n + \left(\frac{l_1 + r_1}{l_2 + r_2} \right)^2 (\check{B}\check{V}(l_2, r_2)^n)^{2 \left(\frac{l_1+r_1}{l_2+r_2}-1 \right)} \check{\Sigma}_{22}^n. \quad (46)$$

Corollary 3.1 *Assume that the conditions from Theorem 3.1 hold true. If X is given by (7), 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 $l_1 + r_1 \neq l_2 + r_2$ and as $n \rightarrow \infty$,*

$$T^n \equiv \frac{n^{1/4} \left(\check{B}\check{V}(l_1, r_1)^n - (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}} \right)}{\sqrt{\check{V}^n}} \xrightarrow{d} N(0, 1). \quad (47)$$

Corollary 3.1 delivers the asymptotic normality of the studentized statistic T^n ; the feasible version of (22). Note that under the alternative presence of heteroscedasticity, $\check{B}\check{V}(l_1, r_1)^n - (\check{B}\check{V}(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}}$ converges to a strictly positive random variable. Moreover, as \check{V}^n was shown to be a robust estimator of V even in presence of jumps and noise, we can conclude that the statistic $T^n \rightarrow \infty$, if the realization of X has a heteroscedastic volatility path. Therefore, appealing to the properties of stable convergence, we deduce that

$$\lim_{n \rightarrow \infty} P(T^n > z_{1-\alpha} \mid \Omega_{\mathcal{H}_0}) = \alpha, \quad (48)$$

$$\lim_{n \rightarrow \infty} P(T^n > z_{1-\alpha} \mid \Omega_{\mathcal{H}_a}) = 1 \quad (49)$$

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 (47) does not require any resampling, it possesses inferior finite sample properties, as shown in Section 4.

Remark 2 *The results from Jacod, Podolskij, and Vetter (2010) and Podolskij and Vetter (2009a) indicate that some assumptions can be relaxed. In particular, in Corollary 3.1, 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. Here, the null is modified as*

$$\mathcal{H}_0 : \omega \in \Omega_{\mathcal{H}_0} \cap \left\{ \omega : t \mapsto \omega_t^2 = E(\epsilon_t^2 \mid X) \text{ is constant on } [0, 1] \right\}. \quad (50)$$

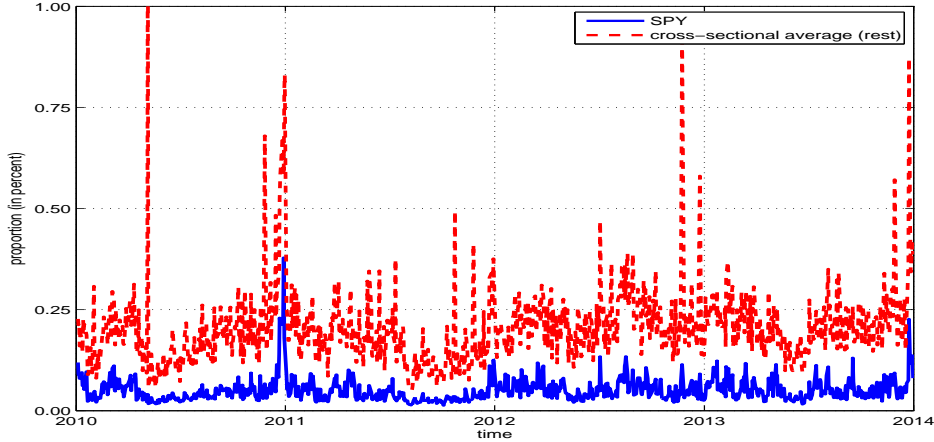
The null hypothesis is therefore a joint statement about the constancy of both diffusive and noise variance. Although such information should be useful in practice, because it delivers knowledge about the presence of heteroscedasticity irrespective of its origin, Figure 1 shows that for our empirical high-frequency data the pre-averaged truncated bipower variation is almost exclusively induced by diffusive volatility, so that very little residual noise is left in the data after pre-averaging. This suggests that any rejection of the null is probably due to genuine time-variation in σ_t .

Corollary 3.2 *Assume that the conditions of Theorem 3.1 hold true and the external random variable is chosen as $u_j \stackrel{\text{i.i.d.}}{\sim} (E^*(u_j), \text{var}^*(u_j))$, such that $\text{var}^*(u_j) = 1/2$. Then, as $n \rightarrow \infty$,*

$$\text{var}^* \left(n^{1/4} \begin{pmatrix} \check{B}\check{V}(l_1, r_1)^{n*} \\ \check{B}\check{V}(l_2, r_2)^{n*} \end{pmatrix} \right) = \check{\Sigma}^n \xrightarrow{P} \Sigma, \quad (51)$$

both in model (80) and (1), where Σ is defined in Appendix A.

Figure 1: Amount of microstructure noise left in $\check{B}\check{V}(1, 1)^n$.



Note. We plot the proportion of $\check{B}\check{V}(1, 1)^n$ that is due to residual variation (after pre-averaging) in the microstructure noise process. We rescale $\check{B}\check{V}(1, 1)^n$ by $\theta\psi_2^{kn}$ to provide an estimate of the integrated variance up to a bias term of order $\psi_1^{kn}\omega^2/(\theta^2\psi_2^{kn})$, see (16) and Theorem 2.1. The figure shows the ratio of this bias to the total variance estimate over time for the ticker symbols that are included in our empirical analysis. ω^2 is replaced by a consistent estimator $\hat{\omega}^2$, due to Oomen (2006).

Given the consistency of the bootstrap variance estimator, we now prove the associated convergence of the bootstrap distribution of $n^{1/4}(\check{B}\check{V}(l_1, r_1)^{n*}, \check{B}\check{V}(l_2, r_2)^{n*})'$.

Theorem 3.2 *Assume that all conditions from Corollary 3.2 hold true and that for any $\delta > 0$, $E^*(|u_j|)^{2+\delta} < \infty$. Then, as $n \rightarrow \infty$,*

$$(\check{\Sigma}^n)^{-1/2} n^{1/4} \begin{pmatrix} \check{B}\check{V}(l_1, r_1)^{n*} - E^*(\check{B}\check{V}(l_1, r_1)^{n*}) \\ \check{B}\check{V}(l_2, r_2)^{n*} - E^*(\check{B}\check{V}(l_2, r_2)^{n*}) \end{pmatrix} \xrightarrow{d^*} N(0, I_2), \quad (52)$$

in probability- P , both in model (80) and (1). Moreover, let

$$S^{n*} = \frac{n^{1/4} \left[\check{B}\check{V}(l_1, r_1)^{n*} - (\check{B}\check{V}(l_2, r_2)^{n*})^{\frac{l_1+r_1}{l_2+r_2}} - \left(E^*(\check{B}\check{V}(l_1, r_1)^{n*}) - \left(E^*(\check{B}\check{V}(l_2, r_2)^{n*}) \right)^{\frac{l_1+r_1}{l_2+r_2}} \right) \right]}{\sqrt{V}}, \quad (53)$$

where $l_1 + r_1 \neq l_2 + r_2$. It holds that

$$V^{n*} \equiv \text{var}^* \left[n^{1/4} \left(\check{B}\check{V}(l_1, r_1)^{n*} - (\check{B}\check{V}(l_2, r_2)^{n*})^{\frac{l_1+r_1}{l_2+r_2}} \right) \right] \xrightarrow{p} V, \quad (54)$$

and

$$S^{n*} \xrightarrow{d^*} N(0, 1), \quad (55)$$

in probability- P , both in model (7) and (1).

Theorem 3.2 shows that the normalized statistic S^{n*} is asymptotically normal both in model (7) 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

$$\mathcal{Z}^{n*} \equiv n^{1/4} \left[\check{B}V(l_1, r_1)^{n*} - (BV(l_2, r_2)^{n*})^{\frac{l_1+r_1}{l_2+r_2}} - \left(E^*(BV(l_1, r_1)^{n*}) - (E^*(BV(l_2, r_2)^{n*}))^{\frac{l_1+r_1}{l_2+r_2}} \right) \right] \quad (56)$$

and

$$\mathcal{Z}^n \equiv n^{1/4} \left(\check{B}V(l_1, r_1)^n - (BV(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}} \right). \quad (57)$$

Remark 3 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 3.2, the statistic $\mathcal{Z}^{n*} \xrightarrow{d^*} N(0, V)$, in probability- P . Note that as $\mathcal{Z}^n \xrightarrow{st} N(0, V)$ 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 $n \rightarrow \infty$. 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 achieve such improvement, we should bootstrap an asymptotically pivotal t -statistic. To this end, we propose a consistent bootstrap estimator of $\check{\Sigma}^n = \text{var}^* \left(n^{1/4} (\check{B}V(l_1, r_1)^{n*}, \check{B}V(l_2, r_2)^{n*})' \right)$. We look at the following adjusted bootstrap version of $\check{\Sigma}^n$ given by $\check{\Sigma}^{n*} = \left(\check{\Sigma}_{ij}^{l_1, r_1, l_2, r_2, n*} \right)_{1 \leq i, j \leq 2}$, where the individual entries of $\check{\Sigma}^{n*}$ are

$$\check{\Sigma}_{ij}^{l_i, r_i, l_j, r_j, n*} = \frac{\sqrt{n} \text{var}^*(u)}{b_n E^*(u^2)} \sum_{k=1}^{J_n} \Delta \check{B}(l_i, r_i)_k^{n*} \Delta \check{B}^*(l_j, r_j)_k^n, \quad (58)$$

with

$$\Delta \check{B}(l, r)_j^{n*} = \Delta \check{B}(l, r)_j^n u_j, \quad (59)$$

where $\Delta \check{B}(l, r)_j^n$ is from (34) and $(u_j)_{j=1}^{J_n}$ are the external random variables used to generate the bootstrap observations in (27). We can also write

$$\check{\Sigma}^{n*} = \frac{\sqrt{n} \text{var}^*(u)}{b_n E^*(u^2)} \sum_{j=1}^{J_n} \check{\xi}_j^* \check{\xi}_j^{*'}, \quad (60)$$

where $\check{\xi}_j^* \equiv u_j \left(\Delta \check{B}(l_1, r_1)_j^n, \Delta \check{B}(l_2, r_2)_j^n \right)'$. We can show that $\check{\Sigma}^{n*}$ consistently estimates $\check{\Sigma}^n$ for any choice of external random variable u with $E^*(|u_j|^4) < \infty$. Next, based on $\check{\Sigma}^{n*}$ we construct a bootstrap studentized variant of (47):

$$T^{n*} \equiv \frac{\mathcal{Z}^{n*}}{\sqrt{\check{V}^{n*}}}, \quad (61)$$

where

$$\check{V}^{n*} = \check{\Sigma}_{11}^{n*} - 2 \left(\frac{l_1 + r_1}{l_2 + r_2} \right) (BV(l_2, r_2)^n)^{\frac{l_1+r_1}{l_2+r_2}-1} \check{\Sigma}_{12}^{n*} + \left(\frac{l_1 + r_1}{l_2 + r_2} \right)^2 (BV(l_2, r_2)^n)^{2 \left(\frac{l_1+r_1}{l_2+r_2}-1 \right)} \check{\Sigma}_{22}^{n*}. \quad (62)$$

Theorem 3.3 *Assume that the conditions of Corollary 3.2 are true and the external random variable is chosen as $u_j \stackrel{\text{i.i.d.}}{\sim} (E^*(u_j), \text{var}^*(u_j))$, such that for any $\delta > 0$, $E^*(|u_j|^{4+\delta}) < \infty$. Then, as $n \rightarrow \infty$,*

$$\left(\check{\Sigma}^{n*} \right)^{-1/2} n^{1/4} \begin{pmatrix} \check{B}\check{V}(l_1, r_1)^{n*} - E^*(\check{B}\check{V}(l_1, r_1)^{n*}) \\ \check{B}\check{V}(l_2, r_2)^{n*} - E^*(\check{B}\check{V}(l_2, r_2)^{n*}) \end{pmatrix} \xrightarrow{d^*} N(0, I_2), \quad (63)$$

in probability- P , both in model (80) and (1). Also,

$$T^{n*} \xrightarrow{d^*} N(0, 1), \quad (64)$$

in probability- P , both in model (7) and (1).

Theorem 3.3 shows the asymptotic normality of the studentized statistic T^{n*} . An implication of results in Theorem 3.3 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*} .

4 Monte Carlo analysis

We here assess the properties of the non-parametric noise- and jump-robust test of heteroscedasticity in the diffusive 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 3, 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:

$$dX_t = a_t dt + \sigma_t dW_t + dJ_t, \quad (65)$$

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

We model diffusive volatility as $\sigma_t = \sigma_{sv,t} \sigma_{u,t}$, where $\sigma_{sv,t}$ and $\sigma_{u,t}$ represent two distinct features of time-varying volatility.

The first term, $\sigma_{sv,t}$, denotes a stochastic process, which allows for randomness in the evolution of σ_t over time. As commonly done in the literature, we assume that $\sigma_{sv,t}$ can be described by a stochastic volatility two-factor structure (SV2F):⁷

$$\sigma_{sv,t} = s\text{-exp}(\beta_0 + \beta_1 \tau_{1,t} + \beta_2 \tau_{2,t}), \quad (66)$$

where

$$d\tau_{1,t} = \alpha_1 \tau_{1,t} dt + dB_{1,t}, \quad d\tau_{2,t} = \alpha_2 \tau_{2,t} dt + (1 + \phi \tau_{2,t}) dB_{2,t}. \quad (67)$$

Here, $B_{1,t}$ and $B_{2,t}$ are two independent standard Brownian motions with $E(dW_t dB_{1,t}) = \rho_1 dt$ and $E(dW_t dB_{2,t}) = \rho_2 dt$.

⁷The 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 $x_0 = \ln(1.5)$.

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 $\rho_1 = \rho_2 = -0.3$.⁸ 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. $\tau_{1,0} \sim N(0, [2\alpha_1]^{-1})$. Meanwhile, τ_2 is started at $\tau_{2,0} = 0$ (e.g., Barndorff-Nielsen, Hansen, Lunde, and Shephard, 2008).

The second term, $\sigma_{u,t}$, is a deterministic trend that represents the diurnality pattern that has been reported to be an important determinant of intraday volatility in many financial return series (e.g., Andersen and Bollerslev, 1997). We follow earlier work of Hasbrouck (1999) and Andersen, Dobrev, and Schaumburg (2012) by using the specification:

$$\sigma_{u,t} = C + Ae^{-a_1 t} + Be^{-a_2(1-t)}. \quad (68)$$

We set $A = 0.75$, $B = 0.25$, $C = 0.88929198$ and $a_1 = a_2 = 10$, which produces a pronounced, asymmetric reverse J-shaped curvature in $\sigma_{u,t}^2$ with an average value of about 3 (1.5) times higher at the start (end) of each simulation compared to the observations in the middle.⁹ This is indeed also a good description of the actual intraday volatility pattern observed in our empirical high-frequency data (see Figure 5 below).

In absence of heteroscedasticity, i.e. under the null hypothesis of constant volatility, we set the variance equal to $E(\sigma_t^2)$, where $E(\sigma_t^2)$ denotes the unconditional expectation of σ_t^2 implied by the above SV2F model.

J_t is the jump component, which we model as a symmetric tempered stable process with Lévy measure given by:

$$\nu(dx) = c \frac{e^{-\lambda x}}{x^{1+\beta}} dx, \quad (69)$$

where $c > 0$, $\lambda > 0$, and $\beta \in [0, 2)$ measures the degree of jump activity (i.e. β is the Blumenthal-Gettoor index). We assume that $\lambda = 3$ and $\beta = 0.5$. The choice of β produces an infinite-activity, finite-variation process dominated by infinitely many, but absolutely summable, small jumps and finitely many large jumps. The idea is to subdue the price process to a stream of small jumps that, in contrast to the large ones, are typically difficult to filter via truncation, and which—set against a null hypothesis of constant volatility—can be confused by the t -statistic with time-varying diffusive volatility. We therefore anticipate that this setup introduces some size distortions in the test. To limit the proportion of the total quadratic variation due to jumps, we calibrate c so that jumps account for 20% of the overall return variation. This parameterization is consistent with extant papers (e.g., Aït-Sahalia, Jacod, and Li, 2012; Aït-Sahalia and Xiu, 2015; Huang and Tauchen, 2005; Todorov and Tauchen, 2010).

We approximate the continuous time representation of σ_t using an Euler scheme, while J_t is

⁸Note that these parameters are annualized. We assume there are 250 trading days in a year.

⁹The calibration of C ensures that $\int_0^1 \sigma_{u,t}^2 dt = 1$, leaving the average integrated variance unchanged.

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 Gryniv (2014).¹⁰ Note that the latter is exact, if $\beta < 1$, as is the case here.

We then 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).

A total of $T = 1,000$ Monte Carlo replica of this model 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 potential second-order effects, we follow Kalnina (2011) and model $\epsilon_{i/n}$ (for a given observation frequency n) as the realization of an MA(1) process:

$$\epsilon_{i/n} = \epsilon'_{i/n} + \varphi \epsilon'_{(i-1)/n}, \quad \text{where } \epsilon'_{i/n} \mid (\sigma_t)_{t \in [0,1]} \stackrel{\text{i.i.d.}}{\sim} N\left(0, \frac{\omega^2}{1 + \varphi^2}\right), \quad (70)$$

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 $\omega^2 = \xi^2 \sqrt{\int_0^1 \sigma_t^4 dt}$, 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., Ait-Sahalia and Yu, 2009).

With the observed price process Y available, we construct noisy returns at sampling frequency n as $\Delta_i^n Y \equiv Y_{i/n} - Y_{(i-1)/n}$. We take $n = 390, 780, 1560, 4680, 7800, 11700$ and 23400 , thereby varying the sample size across a broad range of selections. With the above interpretation of a time unit, the smallest (largest) value of n corresponds 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 5.

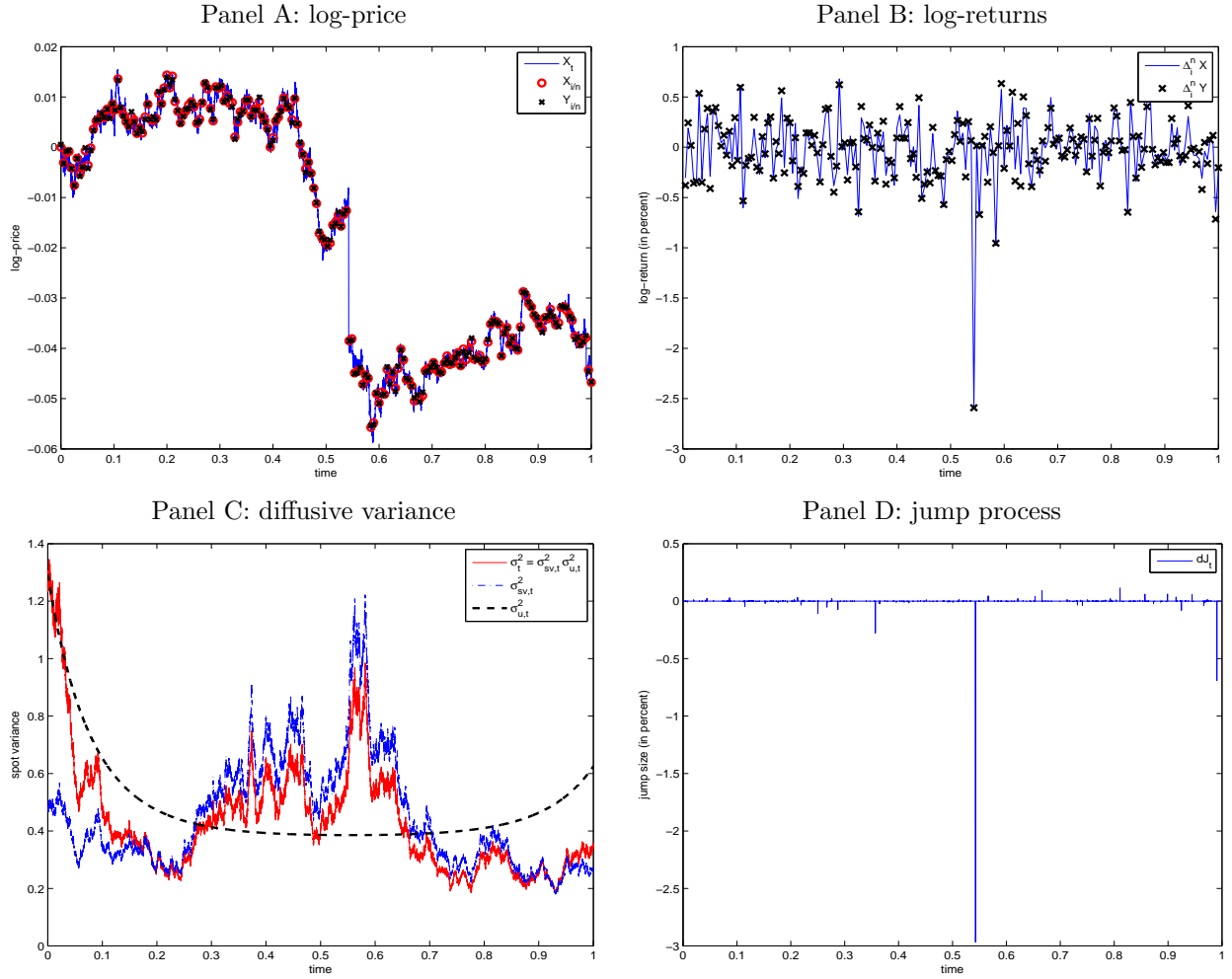
In Figure 2, we provide an illustrative example of a realization from the model for a specific choice of parameters.

We pre-average the sequence of simulated noisy high-frequency data using (12), which we do locally on a window of size $k_n = [\theta \sqrt{n}]$, where $[x]$ is nearest integer function and we settle on $\theta = 1/3$ and $\theta = 1$ (as also done in, e.g., Christensen, Kinnebrock, and Podolskij, 2010).¹¹ As standard in the literature, the weight function is $g(x) = \min(x, 1 - x)$.

¹⁰We thank Viktor Todorov for sharing Matlab code to simulate a tempered stable process.

¹¹As 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 2: Illustration of a simulation.



Note. The figure shows the sample path of the various ingredients in the noisy jump-diffusion model (the first of 10,000 replica in total). In Panel A – B, the sample size is $n = 390$, while the noise is based on $\xi^2 = 0.01$ and $\varphi = -0.5$ (large noise with medium negative first-order autocorrelation). In Panel C, the variance is measured relative to its unconditional average. The tempered stable jump process in Panel D has many small increments that are close, but not equal, to zero.

As stated above, we do inference by comparison of $\tilde{B}\tilde{V}(l, r)^n$ in (18) with $l_1 = r_1 = 2$ and $l_2 = r_2 = 1$. To apply truncation, we implement a threshold $v_n = cu_n^\varpi$ with $u_n = k_n/n$ in (19), which is adapted to an estimate of the local level of volatility. As in, e.g., Li, Todorov, and Tauchen (2013, 2015), we choose a fixed “rate” parameter of $\varpi = 0.49$, while we determine the “scale” dynamically in each simulation 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 (13). The intuition behind this construction is as follows. Assume that there are no jumps in the interval $[i/n, (i + k_n)/n]$. Then, under mild regularity conditions, the asymptotic distribution of the pre-averaged return $\Delta_i^n \bar{Y}$ in (12) is:

$$n^{1/4} \Delta_i^n \bar{Y} \stackrel{a}{\sim} N\left(0, \theta\psi_2\sigma_{i/n}^2 + \frac{1}{\theta}\psi_1\omega^2\right). \quad (71)$$

It follows from (16) that $BV(1, 1)^n \xrightarrow{p} \int_0^1 (\theta\psi_2\sigma_s^2 + \frac{1}{\theta}\psi_1\omega^2) ds$, so that $\sqrt{BV(1, 1)^n}$ is a (jump-robust) measure of the average dispersion (i.e., standard deviation) of the sequence $\Delta_i^n \bar{Y}$, while $\Phi(\cdot)$ controls how far out in the tails of the distribution truncation is enforced.¹² On the other hand, while $\Delta_i^n \bar{Y}$ is of order $O_p(n^{-1/4})$, $\varpi \in (0, 1/2)$ implies that u_n^ϖ shrinks at a slower pace than $\Delta_i^n \bar{Y}$. Therefore, purely “continuous” returns fall within the boundary of the threshold asymptotically. In contrast, if there are jumps in $[i/n, (i + k_n)/n]$, $\Delta_i^n \bar{Y}$ 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, 2015; Hounyo, Gonçalves, and Meddahi, 2015) we examine the robustness of our approach by adopting two candidate distributions:¹³

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

(2.)

$$u_j = \begin{cases} \frac{1}{\sqrt{2}} \left(\frac{1 - \sqrt{5}}{2} \right), & \text{with probability } p = \frac{\sqrt{5} + 1}{2\sqrt{5}} \\ \frac{1}{\sqrt{2}} \left(\frac{1 + \sqrt{5}}{2} \right), & \text{with probability } 1 - p = \frac{\sqrt{5} - 1}{2\sqrt{5}}. \end{cases} \quad (72)$$

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 (the constraint is only binding in the limit). We set $b_n^{\min} = \lceil 2n^\delta \rceil$ and $b_n^{\max} = \lceil \min(3n^\delta, N_n/2) \rceil$. 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

¹²While the pre-averaged (1,1)-bipower variation is robust to the presence of jumps in the p-lim, as $n \rightarrow \infty$, 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).

¹³We 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.

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 – 2, 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 3.1, or with the help of the bootstrap-based percentile and percentile- t approach—with the respective headings z_{wb} and t_{wb} —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 $\pm 1\%$ -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 the t -statistic, illustrating its robustness to market frictions. On the other hand, by comparing Table 1 with Table 2, 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. Also, as expected, the block size b_n increases monotonically with n .

Turning to the analysis of the rejection rates under \mathcal{H}_0 of constant 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 31.5% (22.5%) for $\theta = 1/3$ ($\theta = 1$). This is more than six (four) times larger than the nominal level. These rates improve and decline towards 5% as n increases, but remain 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% 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 weak preference for the one based on the discrete two-point distribution. In the empirical application below, we therefore base our investigation on z_{wb} .

Next, we look at the simulation results under \mathcal{H}_a with time-varying 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 drop in the rejection rates from Table 1 (with $\theta = 1/3$)

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

		\mathcal{H}_0 : constant volatility						\mathcal{H}_a : time-varying volatility													
		size			avg. block size			power			avg. block size										
		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: $\xi^2 = 0.0001$																					
$n =$	195	31.4	10.0	6.3	12.6	11.0	82	77	76	78	75	73.8	37.3	32.5	45.0	44.2	82	77	76	78	76
	390	23.4	5.9	6.1	9.9	9.6	136	131	130	132	131	75.7	44.9	43.5	51.5	51.1	136	131	129	131	131
	780	18.7	5.7	6.1	8.2	7.9	216	212	211	211	211	76.9	56.7	56.3	61.0	60.8	217	211	211	211	211
	1560	14.4	5.2	4.7	6.0	5.3	340	335	333	334	333	79.5	67.1	66.1	68.1	68.5	339	335	334	335	333
	4680	11.8	6.2	6.5	6.7	7.0	707	703	700	697	699	83.7	75.9	76.1	76.2	75.8	713	702	703	700	705
	7800	10.3	5.3	5.2	6.7	6.6	989	978	982	979	977	88.1	83.2	82.8	83.2	82.7	987	982	985	980	980
	11700	8.4	4.4	3.9	4.9	5.4	1,292	1,292	1,283	1,289	1,282	89.4	85.6	85.7	85.2	85.5	1,305	1,288	1,283	1,291	1,284
	23400	8.1	4.9	5.1	6.1	6.1	2,058	2,028	2,046	2,035	2,041	93.4	91.4	91.3	91.0	90.9	2,067	2,038	2,030	2,036	2,049
Panel B: $\xi^2 = 0.0010$																					
$n =$	195	31.6	10.1	6.3	11.9	11.0	82	77	76	78	75	73.8	37.2	31.7	45.0	44.1	82	77	76	78	76
	390	23.1	5.8	6.6	9.8	9.5	136	131	131	131	131	75.7	45.2	43.3	50.9	50.8	136	131	130	131	132
	780	18.7	5.6	6.1	8.5	8.4	216	211	210	211	210	76.9	56.9	56.1	61.0	60.7	217	211	210	210	210
	1560	14.3	5.2	4.7	6.0	5.5	340	336	334	335	334	79.6	66.8	66.2	68.1	68.9	339	336	335	337	332
	4680	11.8	6.2	6.2	6.6	7.2	707	696	699	697	695	83.6	75.3	76.1	75.7	76.1	714	696	702	696	699
	7800	10.1	5.4	5.2	6.6	6.8	991	988	983	982	989	88.2	82.8	82.7	83.5	82.8	985	980	984	983	981
	11700	8.3	4.1	3.8	5.0	5.2	1,292	1,288	1,287	1,288	1,289	89.5	85.3	85.7	85.4	85.7	1,306	1,292	1,274	1,290	1,285
	23400	8.2	4.9	5.1	5.7	6.0	2,058	2,049	2,047	2,046	2,047	93.2	91.4	91.2	90.9	90.9	2,067	2,032	2,053	2,037	2,045
Panel C: $\xi^2 = 0.0100$																					
$n =$	195	31.1	9.4	6.7	12.9	11.4	82	78	76	78	76	73.8	36.8	31.9	45.0	44.3	82	77	76	79	76
	390	22.9	5.5	6.6	10.5	9.8	135	131	131	131	131	74.1	45.2	43.2	50.4	51.2	135	130	130	131	132
	780	18.7	5.1	4.8	8.6	8.5	216	211	211	210	212	77.4	56.5	56.8	61.8	61.0	218	211	210	210	211
	1560	13.7	5.0	5.2	7.3	6.8	340	335	335	335	336	79.2	66.9	66.8	68.9	69.2	340	335	334	336	335
	4680	12.2	6.4	6.0	7.1	7.4	707	701	697	696	695	83.6	75.4	76.2	76.0	76.3	712	700	698	696	695
	7800	10.3	5.5	4.8	6.6	6.4	997	982	983	983	984	87.9	83.2	82.6	83.1	83.0	988	984	988	984	980
	11700	8.7	4.0	4.1	5.6	5.6	1,291	1,288	1,290	1,284	1,289	89.4	86.0	85.9	85.4	85.8	1,306	1,289	1,287	1,280	1,295
	23400	7.5	5.2	5.4	5.8	5.9	2,063	2,041	2,043	2,045	2,040	93.1	91.5	91.5	90.9	91.1	2,072	2,037	2,038	2,043	2,044

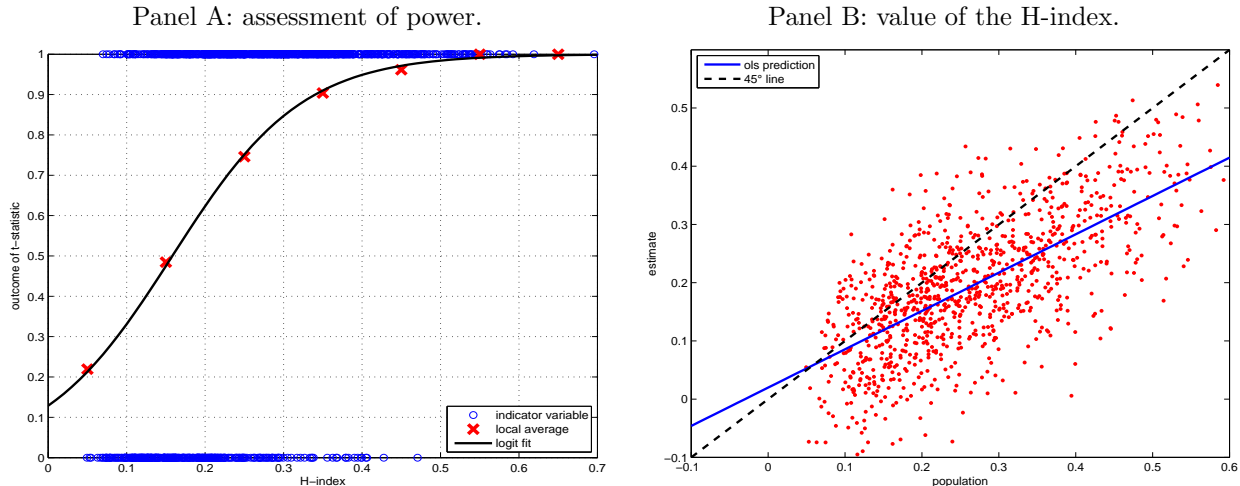
Note. We simulate from a model with drift, diffusive volatility, infinite-activity jumps and market microstructure noise. We test the hypothesis that σ_t is constant and report rejection rates both under \mathcal{H}_0 (size) and \mathcal{H}_a (power). In the latter, σ_t is time-varying due to diurnality and a two-factor SV structure. θ is a tuning parameter that is used to compute the pre-averaging window $k_n = \lfloor \theta \sqrt{n} \rfloor$, 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 (47), while z_{wb} and t_{wb} 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 4.

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

		\mathcal{H}_0 : constant volatility						\mathcal{H}_a : time-varying volatility													
		size			avg. block size			power			avg. block size										
		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: $\xi^2 = 0.0001$																					
$n =$	195	22.3	11.9	7.9	14.6	13.3	77	74	72	74	72	47.8	24.7	20.0	29.2	28.9	77	74	72	75	73
	390	21.8	10.2	10.3	13.5	13.3	135	129	130	131	129	52.7	27.4	26.9	31.4	31.7	135	129	130	131	130
	780	15.9	7.4	7.7	10.4	10.4	215	210	209	211	210	52.8	32.8	31.8	36.2	35.7	217	211	211	211	209
	1560	15.0	6.4	6.7	8.8	8.9	342	334	334	334	334	53.6	38.1	37.5	40.7	40.8	341	333	335	336	336
	4680	10.5	6.5	6.5	7.6	7.7	704	694	698	695	696	64.9	54.9	55.7	56.7	57.0	707	703	698	698	695
	7800	8.7	5.3	5.7	6.8	7.2	991	981	987	984	977	66.9	61.3	60.8	60.7	60.8	1,002	986	979	980	982
	11700	8.1	5.7	5.6	6.7	6.6	1,299	1,296	1,287	1,294	1,279	69.6	64.3	64.6	64.0	64.1	1,311	1,290	1,292	1,284	1,287
	23400	7.7	5.6	5.8	6.1	5.9	2,072	2,041	2,036	2,046	2,036	76.9	72.6	71.9	72.2	71.9	2,066	2,044	2,037	2,042	2,044
Panel B: $\xi^2 = 0.0010$																					
$n =$	195	22.3	12.0	7.9	14.6	13.4	77	74	73	74	73	47.9	24.5	19.8	29.0	29.1	77	74	72	75	73
	390	22.0	10.1	10.4	13.5	13.3	135	130	130	131	130	52.8	27.2	26.7	31.7	31.6	135	129	130	131	130
	780	15.8	7.4	7.5	10.5	10.3	215	209	209	211	210	52.8	32.6	31.9	36.4	35.5	217	210	211	211	210
	1560	15.1	6.3	6.6	9.0	8.9	342	336	334	337	332	53.6	37.8	37.6	40.9	40.9	341	336	336	336	335
	4680	10.5	6.4	6.3	7.6	7.7	704	698	699	700	696	64.9	54.9	55.7	56.8	57.3	707	699	697	696	697
	7800	8.7	5.4	5.8	6.8	7.3	991	984	989	986	990	66.9	61.3	61.1	60.5	60.8	1,001	982	985	978	981
	11700	8.1	5.7	5.6	6.8	6.7	1,299	1,287	1,293	1,287	1,284	69.6	64.2	64.7	63.8	64.1	1,312	1,293	1,291	1,289	1,285
	23400	7.8	5.6	5.8	6.2	5.9	2,073	2,039	2,046	2,046	2,046	76.9	72.5	71.8	72.3	71.9	2,067	2,044	2,058	2,037	2,041
Panel C: $\xi^2 = 0.0100$																					
$n =$	195	21.9	11.3	8.3	14.1	13.9	77	74	72	75	73	48.3	24.4	18.8	29.1	28.5	77	74	72	75	73
	390	21.7	10.3	10.3	13.3	13.5	135	131	131	131	130	53.0	26.8	26.3	32.3	31.8	135	130	130	131	131
	780	16.1	7.4	7.7	10.4	10.6	215	211	210	212	210	52.6	32.7	31.7	36.1	36.2	217	210	209	211	211
	1560	15.1	6.3	6.9	9.3	9.3	342	334	336	335	335	53.6	38.3	37.4	40.9	41.2	341	334	334	334	334
	4680	10.7	5.8	6.5	7.6	7.8	705	696	699	697	700	64.9	55.1	55.3	56.6	56.9	706	700	699	697	698
	7800	8.6	5.3	5.9	6.8	7.1	990	985	988	987	983	67.0	61.2	61.0	60.6	61.1	999	991	986	989	986
	11700	8.3	5.5	5.3	6.5	6.4	1,300	1,288	1,280	1,289	1,290	69.7	63.9	64.2	63.9	63.9	1,313	1,288	1,296	1,292	1,287
	23400	8.1	5.5	5.8	5.9	6.0	2,071	2,044	2,042	2,031	2,041	76.6	72.6	72.6	72.2	72.1	2,059	2,046	2,042	2,041	2,042

Note. We simulate from a model with drift, diffusive volatility, infinite-activity jumps and market microstructure noise. We test the hypothesis that σ_t is constant and report rejection rates both under \mathcal{H}_0 (size) and \mathcal{H}_a (power). In the latter, σ_t is time-varying due to diurnality and a two-factor SV structure. θ is a tuning parameter that is used to compute the pre-averaging window $k_n = \lfloor \theta \sqrt{n} \rfloor$, 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 (47), while z_{wb} and t_{wb} 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 4.

Figure 3: Properties of H-index and t -statistic.



Note. We report the H-index and the outcome of the t -statistic (based on z_{wb2}) for testing \mathcal{H}_0 . H-index from (73) is a measure of heteroscedasticity in σ_t , while its empirical counterpart $\widehat{\text{H-index}}$ is defined in (74). In Panel A, we create an indicator variable I , which tracks whether the t -statistic is significant ($I=1$) or not ($I=0$) at the 5% nominal level. We plot I against the H-index. The crosses are local averages of I around the H-index, while the curve is based on a logistic regression of between the two. The latter can be interpreted as the power of the test conditional on the amount of heteroscedasticity. In Panel B, we regress $\widehat{\text{H-index}}$ on H-index. The fit is $\widehat{\text{H-index}} = 0.0198 + 0.6582 \cdot \text{H-index}$. The 45-degree line offers a reference point for an unbiased estimator. Throughout, the graphs are based on the setting with $n = 23,400$, $\xi^2 = 0.01$, $\varphi = -0.5$ and $\theta = 1$.

vis-à-vis Table 2 (with $\theta = 1$). This is because the increased amount of smoothing diminishes the ability to uncover heteroscedastic volatility, and it highlights a crucial trade-off in practice, as higher values of θ generally render our testing procedures more resilient to the effects of microstructure noise.

This suggests our test is not always powerful enough to pick up variation in σ_t . There are several possible explanations of this finding. First, the problem is not trivial. It may just be hard to detect fluctuations in σ_t from noisy high-frequency data, leaving the jump distortion aside. Second, even if σ_t is time-varying, its sample path—which differ between simulations—can move so little it appears virtually homoscedastic, making it impossible to separate \mathcal{H}_a from natural sampling variation, at least for the sample sizes simulated here.

To shed light on this, we compute an H-index:

$$\text{H-index} = 1 - \frac{\left(\int_0^1 \sigma_s^2 ds\right)^2}{\int_0^1 \sigma_s^4 ds}. \quad (73)$$

The H-index in (73) compares the square of integrated variance to the integrated quarticity. It has the intuitive interpretation that it describes how much σ_t deviates from \mathcal{H}_0 of constant volatility (in percent), see, e.g., Podolskij and Wasmuth (2013).¹⁴ We note that $\text{H-index} \in [0, 1]$ by construction, where $\text{H-index} = 0$ if and only if σ_t is constant, while values greater than zero imply σ_t is to

¹⁴This statistic has indeed 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 also appears in a slightly different form and context in the jump-testing literature (e.g., Barndorff-Nielsen and Shephard, 2006; Kolokolov and Renò, 2016).

some extent time-varying. The H-index is therefore a natural measure of heteroscedasticity in our framework. The average value of the H-index for the two-factor stochastic volatility model with diurnality used in this paper is about 0.25. It falls below 0.10 only 4% of the time, while it is never smaller than 0.05.

In Panel A of Figure 3, we report the outcome of the t -statistic for the set of experiments with time-varying 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 H-index. The fitted line originates from a logit regression of I on H-index, which can be interpreted as the power of the test, conditional on the H-index. As expected, the propensity to discard \mathcal{H}_0 is an increasing function of the H-index. When the deviation from the null is about 0.15, the t -statistic is significant about half of the times, while an H-index above 0.3 – 0.4 implies it more or less always lies in the rejection region.

In practice, we estimate the H-index with an empirical counterpart based on the truncated pre-averaged bipower variation:

$$\widehat{\text{H-index}} = 1 - \frac{(\tilde{B}\tilde{V}(1, 1)^n)^2}{\tilde{B}\tilde{V}(2, 2)^n}. \quad (74)$$

In Panel B of Figure 3, we plot the $\widehat{\text{H-index}}$ against the H-index. We note that $\widehat{\text{H-index}}$ is a consistent estimator of the H-index (up to a bias term, which can be made arbitrarily small with large choices of θ and appears negligible in practice, if $\theta = 1$, please revisit Figure 1). As apparent, $\widehat{\text{H-index}}$ is also downward biased, leading to an understatement of the true level of heteroscedasticity in the diffusive process. The slope coefficient in a regression between the two is about two-thirds, and this in part helps to explain, why it requires a fairly high reading of the H-index to confidently reject \mathcal{H}_0 in our simulations.

Overall, our noise- and jump-robust test of heteroscedasticity in 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.

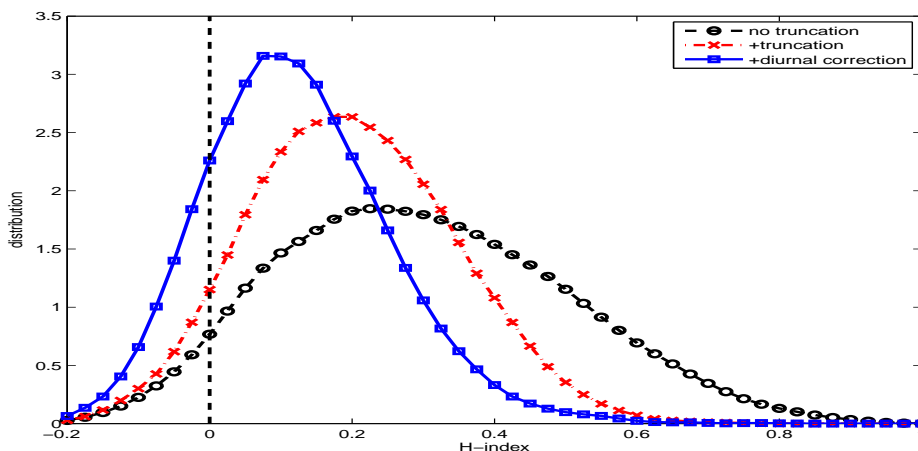
5 Empirical application

In this section, we apply our test of heteroscedasticity to a large cross-sectional panel of US equity high-frequency data. It includes the 30 stocks of the Dow Jones Industrial Average index—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. Our data are extracted from the TAQ database and comprise a complete transaction record for the ticker symbols associated with these stocks. The sample period is January 4, 2010 through December 31, 2013 for a total of 1,006 official exchange trading days. We cleaned the high-frequency data according the filter developed by Christensen, Oomen, and Podolskij (2014), building on earlier work of Brownless

and Gallo (2006); Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). It is a standard way of preparing high-frequency data for analysis in the volatility literature. In the left-hand part of Table 3, we provide a list of the companies included, along with a few summary statistics of their associated transaction data.

In the right-hand side of Table 3, we report the rejection rate of \mathcal{H}_0 and average H-index measurement across the sample for each stock. As above, we base our investigation on $\tilde{B}\tilde{V}(2, 2)^n$ and $\tilde{B}\tilde{V}(1, 1)^n$ and apply the bootstrap percentile approach to evaluate our t -statistic, i.e. z_{wb2} . \mathcal{H}_0 is discarded often, most of the time is excess of 50% – 60%. Moreover, the H-index is far away from zero and typically exceeds 20%, so that the deviation from \mathcal{H}_0 is material. This suggests there is a lot of variation in σ_t . These findings are supported by Figure 4, where we plot the cross-sectional distribution of the H-index measure across stocks and over time. The “no truncation” curve is a kernel smoothed density estimate of the H-index based on the non-truncated estimator in (13), which is not sufficiently jump-robust, as detailed above, or has a large finite sample distortion. The “+ truncation” adds the truncation approach, implemented as advocated in the simulation section. In comparison, the density estimate of the H-index of the non-truncated estimator is shifted much farther to the right. Taken together, this forcefully suggests intraday volatility varies in ways that cannot be ascribed to sampling variation and therefore is not compatible with a constant diffusive coefficient. This is not too surprising, of course, given the vast literature on time-varying volatility. Still, our results show it is important to control for the jump part in this assessment, as otherwise the degree of heteroscedasticity in σ_t is severely overstated.

Figure 4: Cross-sectional distribution of the empirical H-index.



Note. We report a kernel smoothed estimate of the cross-sectional distribution of the H-index: “no truncation” is for the non-truncated estimator, “+truncation” adds jump-truncation, while “+diurnal correction” is based on a jump-robust version of our statistic after the return series has been deflated with a non-parametric jump-robust estimate of intraday diurnality, cf. Figure 5.

Table 3: Descriptive statistics of TAQ high-frequency data.

Name	Exchange	Ticker	n	k_n	$\hat{\sigma}$	$\hat{\rho}_1$	$\hat{\gamma}$	$\#t > \Phi(0.95)$		H-index	
								z_{wb2}	z_{wb2}^d	HI	HI ^d
3M	NYSE	MMM	5,466	73	13.7	-0.10	0.91	0.527	0.249	0.179	0.109
American Express	NYSE	AXP	6,844	82	18.3	-0.08	0.86	0.519	0.211	0.158	0.092
Apple	NASDAQ	AAPL	14,577	120	18.0	-0.14	0.91	0.585	0.193	0.166	0.071
Boeing	NYSE	BA	6,489	80	17.3	-0.11	0.90	0.636	0.267	0.217	0.120
Caterpillar	NYSE	CAT	8,511	91	20.5	-0.10	0.85	0.556	0.146	0.168	0.057
Chevron	NYSE	CVX	8,569	92	15.1	-0.07	0.86	0.441	0.183	0.136	0.076
Cisco Systems	NASDAQ	CSCO	11,884	108	17.1	-0.38	1.81	0.809	0.624	0.275	0.210
Coca-Cola	NYSE	KO	7,992	89	11.7	-0.23	1.11	0.609	0.317	0.196	0.123
DuPont	NYSE	DD	6,522	80	17.3	-0.12	0.92	0.539	0.207	0.173	0.088
ExxonMobil	NYSE	XOM	10,695	103	14.0	-0.10	0.87	0.441	0.186	0.126	0.077
General Electric	NYSE	GE	12,991	114	17.1	-0.41	2.04	0.832	0.686	0.288	0.228
Goldman Sachs	NYSE	GS	7,941	88	20.8	-0.11	0.88	0.578	0.207	0.181	0.081
The Home Depot	NYSE	HD	7,799	88	16.3	-0.17	0.97	0.625	0.296	0.201	0.114
Intel	NASDAQ	INTC	12,516	112	17.5	-0.37	1.67	0.822	0.575	0.274	0.185
IBM	NYSE	IBM	7,194	84	13.0	-0.15	0.95	0.521	0.215	0.165	0.100
Johnson & Johnson	NYSE	JNJ	8,767	93	10.6	-0.20	1.05	0.601	0.330	0.187	0.126
JPMorgan Chase	NYSE	JPM	12,143	110	21.0	-0.16	0.94	0.613	0.224	0.178	0.089
McDonald's	NYSE	MCD	7,193	84	11.3	-0.14	0.93	0.567	0.260	0.188	0.108
Merck	NYSE	MRK	8,457	92	14.2	-0.26	1.18	0.670	0.384	0.218	0.143
Microsoft	NASDAQ	MSFT	12,815	113	16.2	-0.34	1.46	0.788	0.583	0.262	0.188
Nike	NYSE	NKE	4,847	69	16.3	-0.08	0.86	0.613	0.229	0.213	0.106
Pfizer	NYSE	PFE	11,254	106	15.0	-0.39	1.93	0.849	0.685	0.294	0.234
Procter & Gamble	NYSE	PG	7,968	89	11.2	-0.17	1.00	0.563	0.305	0.178	0.122
Travelers	NYSE	TRV	4,779	68	13.5	-0.12	0.93	0.552	0.263	0.184	0.117
UnitedHealth Group	NYSE	UNH	6,730	81	18.9	-0.11	0.90	0.689	0.298	0.236	0.121
United Technologies	NYSE	UTX	5,939	76	14.7	-0.11	0.92	0.595	0.271	0.197	0.111
Verizon	NYSE	VZ	9,192	95	13.1	-0.30	1.30	0.738	0.491	0.241	0.164
Visa	NYSE	V	6,213	78	17.7	-0.14	0.94	0.677	0.321	0.229	0.144
Wal-Mart	NYSE	WMT	8,086	89	11.7	-0.18	0.99	0.621	0.281	0.200	0.116
Walt Disney	NYSE	DIS	7,422	86	16.1	-0.16	0.97	0.524	0.248	0.168	0.097
SPDR S&P 500 ETF	NYSE	SPY	18,154	135	10.3	-0.05	0.83	0.500	0.227	0.126	0.080

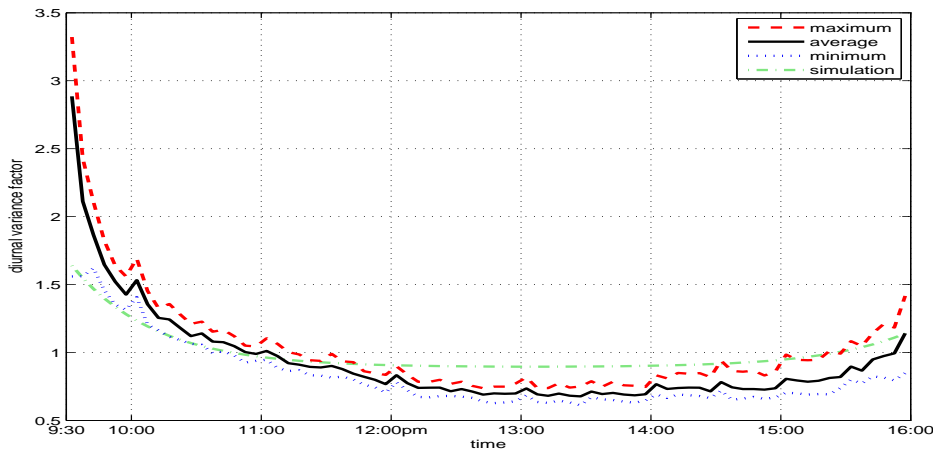
Note. This table reports descriptive statistics for our TAQ high-frequency data. The numbers are computed daily and averaged by ticker across the sample, which covers January 4, 2010 through December 31, 2013, for a total of $T = 1,006$ days. n is number of transaction data available after filtering, k_n is the length of the pre-averaging window, $\hat{\sigma}$ is an annualized jump-robust realized measure of volatility based on a bias-corrected version of $BV(1,1)^n$ (cf. (16)), $\hat{\rho}_1$ is the first-order autocorrelation of $\Delta_1^n Y$, $\hat{\gamma}$ is the noise-ratio parameter of Oomen (2006), while $\#t > \Phi(0.95)$ is the fraction of t -statistics (based on z_{wb2}) for testing \mathcal{H}_0 of homoscedasticity larger than the 95%-quantile from the standard normal distribution. H-index is the heteroscedasticity index defined in (73). A superscript d refers to the value of the statistic, after the log-return series has been corrected for diurnal variation in σ_t .

5.1 Heteroscedasticity and intraday periodicity

It has been noticed that changes in intraday financial volatility is largely predictable, reflecting a natural periodicity in the trading environment (e.g., Andersen and Bollerslev, 1997, 1998). In the equity market, for instance, time-of-day volatility resembles a reverse J-shape with high volatility in the morning, as overnight information gets incorporated into prices, followed by a cooling down around noon, and then rising volatility prior to the close of the exchange as traders adjust their holdings. In this light, we can ask whether the rejection of \mathcal{H}_0 is driven mostly by the (more or less) deterministic variation of σ_t due to intraday periodicity?

Our procedure enables such an assessment and can help to shed light on this question. We therefore apply our test and the H-index measure to a modified log-return series that has been deflated by a non-parametric jump-robust measure of diurnal variation in σ_t . We follow Boudt, Croux, and Laurent (2011), who propose a time-of-the-day volatility factor as a jump-robust weighted standard deviation (WSD) estimator. We first compute a previous-tick interpolated 5-minute log-return series across trading days. The WSD measure is then computed stock-by-stock as advocated in their paper. The corresponding estimate is reported in Figure 5. The shape aligns with that found in the previous work cited above, as noticed by the good correspondence with the setting in the simulation section, which is inserted as a reference point.

Figure 5: Intraday variance factor in equity high-frequency data.



Note. The figure presents the relative variation in the intraday returns of our TAQ high-frequency equity data. We construct a series of 5-minute returns for each stock and then estimate the variance factors with the non-parametric jump-robust weighted standard deviation estimator in Boudt, Croux, and Laurent (2011). The average value across the 31 companies we analyze is reported, along with the highest and smallest estimate. As a comparison, we also plot the curve used in our simulations.

In Table 3, we recompute the t -statistic and H-index measure on the diurnal-corrected log-return series. The two columns related to these numbers are highlighted by a superscript d . As seen, the rejection of \mathcal{H}_0 drops substantially and now fails on average only about every fourth day. The H-index often hovers around 0.1, which suggests that large portion of the time-variation is sourced from the diurnal component, although residual volatility is still not completely constant. The implication is that most of the days, but not always, if we get a good measure of the initial

level of σ_t , for example via a kernel-based estimate, we can largely predict the intraday sample path for volatility on a majority of days. This is corroborated by Figure 4, where we include the cross-sectional distribution of the H-index measure after the diurnal correction has been switched on (the “+ diurnal correction” curve). As evident, the distribution is less skewed to the right and more concentrated around zero, although some deviation from \mathcal{H}_0 remains.

6 Conclusion

The main contribution of this paper is to propose an easy-to-compute statistical procedure to measure the strength and test the presence of time-varying volatility in a discretely sampled jump-diffusion process that is contaminated by microstructure noise. We use the concept of truncated pre-averaged bipower variation to determine whether there is a presence of heteroscedasticity in a discretely sampled process or not in noisy jump-diffusion models. The test statistic diverges to infinity if heteroscedasticity are present and have a normal distribution otherwise. To improve inference, we suggest a bootstrap approach, and theoretically justify the use of the proposing bootstrap method for testing the null hypothesis of “absence of heteroscedasticity” in noisy high frequency data. Our Monte Carlo simulations show that the bootstrap improves the finite sample properties of the asymptotic theory based test in presence of infinite-activity price jumps. To complement the theory and Monte Carlo simulations results, we conduct an empirical application, which documents the importance of jump-robustness, when measuring heteroscedasticity in practice. We also find that a large fraction of variation in intraday volatility is accounted for by seasonality. This suggests that, once we control for jumps and deflate asset returns by a non-parametric estimate of the conventional U-shaped diurnality profile, the variance of the rescaled return series is often close to constant within the day.

Appendix

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 $n^{1/4}(BV(Y, l_1, r_1)^n, BV(Y, l_2, r_2)^n)'$, i.e. Σ appearing in (17), Theorem B.1 and Theorem 3.1. We also prove a corresponding result for the bootstrap version, $\check{\Sigma}^{n*}$, in Section 3. 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:

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

where a is a real number, b and c are a two- and four-dimensional vector. Moreover, (H_1, \dots, 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

$$\begin{aligned} \phi_1(s) &= \int_0^{1-s} g'(u)g'(u+s)du, & \phi_2(s) &= \int_0^{1-s} g(u)g(u+s)du, \\ \phi_3(s) &= \int_0^{2-s} g'(u)g'(u+s-1)du & \text{and} & \quad \phi_4(s) = \int_0^{2-s} g(u)g(u+s-1)du. \end{aligned}$$

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 $f(s) = (f_1(s), f_2(s), f_3(s), f_4(s))'$. At last, we get that

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

where

$$w_{ij}^{l_1, r_1, l_2, r_2}(\sigma_u) = 2\theta \int_0^2 h_{ij}(\sigma_u, t, f(s)) ds.$$

B Proofs

In the 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 $\tilde{\gamma}$:

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

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.

$$X'_t = X_0 + \int_0^t a'_s ds + \int_0^t \sigma_s dW_s, \quad X''_t = X_t - X'_t, \quad (75)$$

where, according to the value of β , we set

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

Then, we can write

$$Y_t = Y'_t + Y''_t, \quad (76)$$

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

Proof of Theorem 2.1. Here, we more or less follow the techniques applied in the proof of Theorem 4.1 in Hounyo (2015). Under the stated assumptions, the definitions of $\check{B}V(Y, l, r)^n, \check{B}V(Y', l, r)^n$, and the central limit theorem in Theorem 3 of Podolskij and Vetter (2009a), it holds that, as $n \rightarrow \infty$,

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

Thus, it suffices to prove that for any $l, r > 0$,

$$n^{1/4} (\check{B}V(Y, l, r)^n - \check{B}V(Y', l, r)^n) \xrightarrow{p} 0. \quad (77)$$

To show (77), 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 $x = (x_1 \ x_2)'$. As in the line of thought on page 385 in Jacod and Protter (2012), we can show that for $w_n = v_n / \sqrt{u_n}$ with $u_n = k_n / n$:

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

Next, let $x = (\Delta_{i-1}^n \bar{Y} \ \Delta_{i-1+k_n}^n \bar{Y})' / \sqrt{u_n}$ and $y = (\Delta_{i-1}^n \bar{Y}'' \ \Delta_{i-1+k_n}^n \bar{Y}'')' / \sqrt{u_n}$. 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 $l + r > 0$:

$$E(\|x\|^{l+r}) \leq K, \quad E(\|y\| \wedge 1) \leq K u_n^{1-\beta/2} \phi_n \quad \text{and} \quad E(\|y\|^2 \wedge w_n^2) \leq K u_n^{\omega(2-\beta)} \phi_n, \quad (78)$$

where $\phi_n \rightarrow 0$ as $n \rightarrow 0$. In addition, from (78) 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

$$E(\|y\|^{l+r} \wedge w_n^{l+r}) \leq K w_n^{l+r-2} E(\|y\| \wedge w_n)^2 \leq K u_n^{\omega(l+r-\beta) - \frac{1}{2}(l+r-2)} \phi_n, \quad (79)$$

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

$$\begin{aligned} & n^{1/4} (B\check{V}(Y, l, r)^n - B\check{V}(Y', l, r)^n) \\ &= \frac{n^{\frac{l+r-3}{4}}}{\mu_l \mu_r} \sum_{i=1}^{n-2k_n+2} \left(|\Delta_{i-1}^n \bar{Y}|^l |\Delta_{i-1+k_n}^n \bar{Y}|^r - |\Delta_{i-1}^n \bar{Y}'|^l |\Delta_{i-1+k_n}^n \bar{Y}'|^r \right) 1_{\{|\Delta_{i-1}^n \bar{Y}| < v_n\}} 1_{\{|\Delta_{i-1+k_n}^n \bar{Y}| < v_n\}}, \end{aligned}$$

it follows that

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

Thus, if $\beta < 1$ and $\frac{l+r-1}{2(l+r-\beta)} \leq \varpi < 1/2$, then $E(|n^{1/4}(B\check{V}(Y, l, r)^n - B\check{V}(Y', l, r)^n)|) \rightarrow 0$ and therefore $n^{1/4}(B\check{V}(Y, l, r)^n - B\check{V}(Y', l, r)^n) \xrightarrow{p} 0$. This completes the proof of Theorem 2.1. \blacksquare

Next, we establish the following result (under no jumps) since it will be useful later in the proof of Theorem 3.1.

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

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s. \quad (80)$$

We define:

$$\hat{\Sigma}^n = \frac{\sqrt{n}}{2b_n} \sum_{i=1}^{N_n-2b_n+1} \xi_i \xi_i', \quad (81)$$

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

$$\Delta B(l, r)_j^n = B(Y, l, r)_{j+b_n}^n - B(Y, l, r)_j^n, \quad (82)$$

with

$$B(l, r)_j^n = n^{\frac{l+r}{4}-1} \frac{1}{\mu_l \mu_r} \sum_{i=1}^{b_n} y(l, r)_{i-1+j}^n. \quad (83)$$

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 $k_n \rightarrow \infty$ as $n \rightarrow \infty$ such that (14) holds, and the block size b_n fulfills (25) for some $1/2 < \delta_1 < 2/3$. Then, as $n \rightarrow \infty$,

$$\hat{\Sigma}^n \xrightarrow{p} \Sigma, \quad (84)$$

where Σ is defined in Appendix A.

Proof of Theorem B.1. Here, recall that X follows (80) and note that given (81), we can rewrite $\hat{\Sigma}^n$ as follows:

$$\hat{\Sigma}^n = \frac{1}{b_n} \sum_{m=1}^{b_n} \hat{\Sigma}_m^n, \quad (85)$$

where

$$\hat{\Sigma}_m^n = \frac{\sqrt{n}}{2} \sum_{k=0}^{\lfloor \frac{N_n}{b_n} \rfloor - 2} \xi_{kb_n+m} \xi'_{kb_n+m} = \left(\hat{\Sigma}_{ij,m}^{l_1, r_1, l_2, r_2, n} \right)_{1 \leq i, j \leq 2}. \quad (86)$$

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

$$\text{p-lim}_{n \rightarrow \infty} \hat{\Sigma}_{ij,m}^{l_1, r_1, l_2, r_2, n} = \Sigma_{ij}^{l_1, r_1, l_2, r_2}, \quad 1 \leq i, j \leq 2, \quad (87)$$

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

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

Then, given the definition of $\Delta B(Y, l, r)_m^n$ given in (82), by adding and subtracting appropriately, it follows that

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

where the remainder term is

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

uniformly in m , so long as $\delta_1 < 3/4$, where we apply the definition of $B(Y, l, r)_m^n$ in (83), the Cauchy-Schwartz inequality, and the fact that $E(|\Delta_i^n \bar{Y}^l|) \leq Kn^{-l/4}$ (cf., Lemma 1 of Podolskij

and Vetter, 2010). Next, we show the main term is such that

$$\text{p-lim}_{n \rightarrow \infty} M_{ij,m}^{l_1, r_1, l_2, r_2, n}(Y) = \Sigma_{ij}, \quad 1 \leq i, j \leq 2, \quad (88)$$

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

$$\tilde{M}_{ij,m}^{l_1, r_1, l_2, r_2, n}(Y) = \sqrt{n} \sum_{k=1}^{\lfloor \frac{N_n}{b_n} \rfloor - 1} \begin{pmatrix} B(Y, l_i, r_i)_{kb_n+m}^n B(Y, l_j, r_j)_{kb_n+m}^n \\ -B(Y, l_i, r_i)_{kb_n+m}^n B(Y, l_j, r_j)_{(k-1)b_n+m}^n \end{pmatrix}. \quad (89)$$

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

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

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

$$\begin{aligned} E\left(|\Delta_i^n \bar{Y} - \Delta_i^n \tilde{Y}|\right) &= E\left(\left|\sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) \int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} a_s ds + \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) \int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} (\sigma_s - \sigma_{\frac{jb_n}{N_n}}) dW_s\right|\right) \\ &\leq K \left(\frac{k_n}{n} + \left(\sum_{j=1}^{k_n} g^2\left(\frac{j}{k_n}\right) E\left(\left|\int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} (\sigma_s - \sigma_{\frac{jb_n}{N_n}}) dW_s\right|\right)^2\right)^{1/2}\right) \\ &\leq K \left(\frac{k_n}{n} + \left(\frac{k_n b_n}{n}\right)^{1/2}\right) \leq K \frac{(k_n b_n)^{1/2}}{n}. \end{aligned}$$

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

$$\begin{aligned} E\left(|B(Y, l, r)_{jb_n+m}^n - \tilde{B}(Y, l, r)_{jb_n+m}^n|\right) &\leq K b_n \left(\frac{(k_n b_n)^{1/2}}{n} \left(\frac{1}{\sqrt{k_n}}\right)^{\frac{(l+r)}{4}-1}\right) \\ &\leq K \left(\frac{b_n}{n}\right)^{3/2}. \end{aligned}$$

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

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

such that

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

Then,

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

by conditional independence, and now we are left with

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

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

$$\sqrt{n} E \left(\mathcal{B}_{kb_n+m}^n - \hat{\mathcal{B}}_{kb_n+m}^n \mid \mathcal{F}_{\frac{(k-1)b_n+m}{N_n}}^n \right) = 2\theta \int_{\frac{(k-1)b_n}{N_n}}^{\frac{kb_n}{N_n}} \int_0^2 h_{ij}(\sigma_u, t, f(s)) ds du + o \left(\frac{b_n}{N_n} \right),$$

uniformly in k and m , and thus

$$\begin{aligned} \bar{M}_{ij,m}^{l_1, r_1, l_2, r_2, n}(Y) &= 2\theta \int_0^1 \int_0^2 h_{ij}(\sigma_u, t, f(s)) ds du + o_p(1) \\ &= \int_0^1 w_{ij}^{l_1, r_1, l_2, r_2}(\sigma_u) du + o_p(1), \end{aligned}$$

uniformly in m , and the proof is complete. \blacksquare

Proof of Theorem 3.1. We prove (45) solely in model (1), which is enough, as it is the most general and nests (80). Now, under the stated assumptions, the definitions of $\check{\Sigma}_{ij}^{l_1, r_1, l_2, r_2, n}(Y)$, $\check{\Sigma}_{ij}^{l_1, r_1, l_2, r_2, n}(Y')$, and the limiting result in Theorem B.1, we deduce that, as $n \rightarrow \infty$,

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

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

$$\text{p-lim}_{n \rightarrow \infty} \left(\check{\Sigma}_{ij,m}^{l_1, r_1, l_2, r_2, n}(Y) - \check{\Sigma}_{ij,m}^{l_1, r_1, l_2, r_2, n}(Y') \right) = 0, \quad \text{for } 1 \leq i, j \leq 2, \quad (90)$$

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

$$\begin{aligned}
& \frac{2}{\sqrt{n}} \left(\check{\Sigma}_{ij, m}^{l_1, r_1, l_2, r_2, n}(Y) - \check{\Sigma}_{ij, m}^{l_1, r_1, l_2, r_2, n}(Y') \right) \\
&= \sum_{k=0}^{\lfloor \frac{Nn}{b_n} \rfloor - 2} \left(\Delta \check{B}(Y, l_i, r_i)_{kb_n+m}^n \Delta \check{B}(Y, l_j, r_j)_{kb_n+m}^n - \Delta \check{B}(Y', l_i, r_i)_{kb_n+m}^n \Delta \check{B}(Y', l_j, r_j)_{kb_n+m}^n \right) \\
&= \sum_{k=0}^{\lfloor \frac{Nn}{b_n} \rfloor - 2} \left(\left(\check{B}(Y, l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y, l_j, r_j)_{(k+1)b_n+m}^n - \check{B}(Y', l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y', l_j, r_j)_{(k+1)b_n+m}^n \right) \right. \\
&\quad - \left(\check{B}(Y, l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y, l_j, r_j)_{kb_n+m}^n - \check{B}(Y', l_i, r_i)_{(k+1)b_n+m}^n \check{B}(Y', l_j, r_j)_{kb_n+m}^n \right) \\
&\quad - \left(\check{B}(Y, l_i, r_i)_{kb_n+m}^n \check{B}(Y, l_j, r_j)_{(k+1)b_n+m}^n - \check{B}(Y', l_i, r_i)_{kb_n+m}^n \check{B}(Y', l_j, r_j)_{(k+1)b_n+m}^n \right) \\
&\quad \left. + \left(\check{B}(Y, l_i, r_i)_{kb_n+m}^n \check{B}(Y, l_j, r_j)_{kb_n+m}^n - \check{B}(Y', l_i, r_i)_{kb_n+m}^n \check{B}(Y', l_j, r_j)_{kb_n+m}^n \right) \right) \quad (91)
\end{aligned}$$

where

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

In the following, we define:

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

where $C_{k, k'} = \{|\Delta_{k-1}^n \bar{Y}| < v_n\} \cap \{|\Delta_{k-1+k_n}^n \bar{Y}| < v_n\} \cap \{|\Delta_{k'-1}^n \bar{Y}'| < v_n\} \cap \{|\Delta_{k'-1+k_n}^n \bar{Y}'| < v_n\}$.

Then, from (91) it follows that

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

The statement in (90) is therefore reduced to show that

$$\check{\Sigma}_{ij, m}^{(k), l_1, r_1, l_2, r_2, n}(Y, Y') \xrightarrow{p} 0, \quad (92)$$

for $k = 1, \dots, 4$. 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

$x = (x_1 \ x_2 \ x_3 \ x_4)'$. Following the line of thought used also in the proof of Theorem 2.1, we can show that for $w_n = v_n/\sqrt{u_n}$ with $u_n = k_n/n$:

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

where $p = l_1 + r_1 + l_2 + r_2$. Next, set $x = (\Delta_{k-1}^n \bar{Y} \ \Delta_{k-1+k_n}^n \bar{Y} \ \Delta_{k'-1}^n \bar{Y} \ \Delta_{k'-1+k_n}^n \bar{Y})' / \sqrt{u_n}$, $y = (\Delta_{k-1}^n \bar{Y}'' \ \Delta_{k-1+k_n}^n \bar{Y}'' \ \Delta_{k'-1}^n \bar{Y}'' \ \Delta_{k'-1+k_n}^n \bar{Y}'')' / \sqrt{u_n}$. As in (78) — (79), it holds true that

$$E(\|x\|^p) \leq K, \quad E(\|y\| \wedge 1) \leq K u_n^{1-\beta/2} \phi_n \quad \text{and} \quad E((\|y\| \wedge w_n)^p) \leq K u_n^{\omega(p-\beta) - \frac{(p-2)}{2}} \phi_n, \quad (93)$$

where $\phi_n \rightarrow 0$ as $n \rightarrow \infty$. Therefore,

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

which concludes the proof of (90) and, hence, Theorem 3.1. ■

Proof of Lemma 3.1. The linearity of the expectation operator implies that

$$\begin{aligned} E^*(\check{B}\check{V}(l, r)^{n*}) &= E^* \left[\check{B}\check{V}(l, r)^n - \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} \Delta \check{B}(l, r)_j^n u_j \right] \\ &= \check{B}\check{V}(l, r)^n - \frac{1}{\sqrt{b_n}} \sum_{j=1}^{J_n} \Delta \check{B}(l, r)_j^n E^*(u_j). \end{aligned}$$

Then, if $E^*(u_j) = 0$, it follows that $E^*(\check{B}\check{V}(l, r)^{n*}) = \check{B}\check{V}(l, r)^n$. The second part of the lemma follows from (27) and (33), as for $1 \leq i, j \leq 2$,

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

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

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

where

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

■

Proof of Corollary 3.1. Given (20), (22), and (46) the results follows from the properties of stable convergence. ■

Proof of Corollary 3.2. The result follows directly given (38) and the consistency result of $\check{\Sigma}^{l_1, r_1, l_2, r_2}$ in Theorem 3.1. ■

Proof of Theorem 3.2. We again prove the theorem in model (1) only, noting that this is enough, as it nests both (80) and (7). Write

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

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

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

where u_j are i.i.d. with $\text{var}^*(u_j) = 1/2$. 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 $\lambda \in D$, $\lambda' Z_n^* \xrightarrow{d^*} N(0, 1)$, in probability- P , as $n \rightarrow \infty$. We note that

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

It follows from Lemma 3.1 and Corollary 3.2 that $E^*(\lambda' Z_n^*) = 0$ and $\text{var}^*(\lambda' Z_n^*) = 1$ for all n . To conclude, it thus remains to prove that $\lambda' Z_n^*$ 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 $K > 0$, $\sup_{x \in \mathbb{R}} \left| P^* \left(\sum_{j=1}^{J_n} n^{1/4} \lambda' z_j^* \leq x \right) - \Phi(x) \right| \leq K \sum_{j=1}^{J_n} E^* \left| n^{1/4} \lambda' z_j^* \right|^{2+\varepsilon}$. Next, we show that $\sum_{j=1}^{J_n} E^* \left| n^{1/4} \lambda' z_j^* \right|^{2+\varepsilon} = o_p(1)$. 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:

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

Thus,

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

implying that

$$\begin{aligned}
\sum_{j=1}^{J_n} E^* |n^{1/4} \lambda' z_j^*|^{2+\varepsilon} &\leq K n^{\frac{2+\varepsilon}{4}} \left\| (\check{\Sigma}^n)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \|D_j\|^{2+\varepsilon} \\
&\leq K n^{\frac{2+\varepsilon}{4}} \left\| (\check{\Sigma}^n)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \left((\Delta \check{B}(l_1, r_1)_j^n)^{2+\varepsilon} + (\Delta \check{B}(l_2, r_2)_j^n)^{2+\varepsilon} \right) \\
&\leq K n^{\frac{2+\varepsilon}{4}} \left\| (\check{\Sigma}^n)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} (\check{B}(l_1, r_1)_{j+b_n}^n - \check{B}(l_1, r_1)_j^n)^{2+\varepsilon} \\
&\quad + K n^{\frac{2+\varepsilon}{4}} \left\| (\check{\Sigma}^n)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} (\check{B}(l_2, r_2)_{j+b_n}^n - \check{B}(l_2, r_2)_j^n)^{2+\varepsilon} \\
&\leq K n^{\frac{2+\varepsilon}{4}} \left\| (\check{\Sigma}^n)^{-1/2} \right\|^{2+\varepsilon} \sum_{j=1}^{J_n} \left((\check{B}(l_1, r_1)_j^n)^{2+\varepsilon} + (\check{B}(l_2, r_2)_j^n)^{2+\varepsilon} \right), \tag{94}
\end{aligned}$$

where the second inequality is due to that, for any j , $\|D_j\|^2 = (\Delta \check{B}(l_1, r_1)_j^n)^2 + (\Delta \check{B}(l_2, r_2)_j^n)^2$, while the third is by expression of $\Delta \check{B}(l, r)_j^n$. Next, note that by definition of $\check{B}(l, r)_j^n$:

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

We can therefore write (94) as follows:

$$\begin{aligned}
\sum_{j=1}^{J_n} E^* (|n^{1/4} \lambda' z_j^*|^{2+\varepsilon}) &= O_p\left(n^{\frac{2+\varepsilon}{4}} n^{(\delta_1-1)(1+\varepsilon)} \right) \\
&= o_p(1),
\end{aligned}$$

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 (52). The last results then follow by application of the delta rule. ■

Proof of Theorem 3.3. First, we define:

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

where

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

It follows from Theorem 3.2 that $Z^{n*} \xrightarrow{d^*} N(0, I_2)$. Thus, the central limit theory for H^{n*} is established, if we can show that $(\check{\Sigma}^{n*})^{-1} \check{\Sigma}^n \xrightarrow{p^*} I_2$. To do this, we prove that

$$E^* \left[(\check{\Sigma}^n)^{-1} \check{\Sigma}^{n*} \right] \xrightarrow{p^*} I_2 \quad \text{and} \quad \text{var}^* \left[(\check{\Sigma}^n)^{-1} \check{\Sigma}^{n*} \right] \xrightarrow{p^*} 0. \quad (95)$$

The first equation in (95) holds by the definition of $\check{\Sigma}^n$ and $\check{\Sigma}^{n*}$. Next, again by definition:

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

As in the proof of Theorem 4.2 in Hounyo (2015):

$$\begin{aligned} E \left(\left\| \frac{n}{b_n^2} \sum_{j=1}^{J_n} (\check{\xi}_j \check{\xi}'_j) \oplus (\check{\xi}_j \check{\xi}'_j) \right\| \right) &\leq K \frac{n}{b_n^2} \sum_{j=1}^{J_n} \left(\sqrt{E(|\check{B}(l_1, r_1)_j^n|^4)} \sqrt{E(|\check{B}(l_2, r_2)_j^n|^4)} \right. \\ &\quad + \sqrt{E(|\check{B}(l_1, r_1)_{j+b_n}^n|^4)} \sqrt{E(|\check{B}(l_2, r_2)_j^n|^4)} \\ &\quad + \sqrt{E(|\check{B}(l_1, r_1)_j^n|^4)} \sqrt{E(|\check{B}(l_2, r_2)_{j+b_n}^n|^4)} \\ &\quad \left. + \sqrt{E(|\check{B}(l_1, r_1)_{j+b_n}^n|^4)} \sqrt{E(|\check{B}(l_2, r_2)_{j+b_n}^n|^4)} \right) \leq K \frac{b_n^2}{n^2} \rightarrow 0, \end{aligned}$$

where the last inequality follows, because $\frac{l+r-1}{2(l+r-\beta)} \leq \varpi < 1/2$ means that $\sqrt{E(|\check{B}(l, r)_j^n|^4)} \leq K \frac{b_n^2}{n^2}$. As $J_n = O(n)$ and $b_n = O(n^{\delta_1})$ such that $1/2 < \delta_1 < 2/3$ from (25), it follows that

$$\text{var}^* \left[(\check{\Sigma}^n)^{-1} \check{\Sigma}^{n*} \right] \xrightarrow{p^*} 0.$$

This finishes the proof of the first part Theorem 3.3. The last result again follows by a direct application of the delta rule. ■

References

- Aït-Sahalia, Y., 1996, “Testing continuous-time models of the spot interest rate,” *Review of Financial Studies*, 9(2), 385–426.
- , 2004, “Disentangling diffusion from jumps,” *Journal of Financial Economics*, 74(3), 487–528.
- 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*, (Forthcoming).
- 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., D. Dobrev, and E. Schaumburg, 2012, “Jump-robust volatility estimation using nearest neighbour truncation,” *Journal of Econometrics*, 169(1), 75–93.
- 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.
- Black, F., and M. Scholes, 1973, “The pricing of options and corporate liabilities,” *Journal of Political Economy*, 81(3), 637–654.
- 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, 2010, “Realised quantile-based estimation of the integrated variance,” *Journal of Econometrics*, 159(1), 74–98.
- , 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*, (Forthcoming).
- Constantinides, G. M., 1992, “A theory of the nominal term structure of interest rates,” *Review of Financial Studies*, 5(4), 531–552.
- Corradi, V., and W. Distaso, 2006, “Semi-parametric comparison of stochastic volatility models using realized measures,” *Review of Economic Studies*, 73(3), 635–667.
- Corradi, V., and H. White, 1999, “Specification tests for the variance of a diffusion,” *Journal of Time Series Analysis*, 20(3), 253–270.
- 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.
- Cox, J. C., J. E. Ingersoll, and S. A. Ross, 1985, “A theory of the term structure of interest rates,” *Econometrica*, 53(2), 385–407.
- 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. Papanoditis, and D. N. Politis, 2014, “A generalized block bootstrap for seasonal time series,” *Journal of Time Series Analysis*, 35(2), 89–114.
- Duffie, D., and J. M. Harrison, 1993, “Arbitrage pricing of Russian options and perpetual Lookback options,” *Annals of Applied Probability*, 3(3), 641–651.
- Duffie, D., and R. Kan, 1996, “A yield-factor model of interest rates,” *Mathematical Finance*, 6(4), 379–406.
- 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.
- Hounyo, U., 2014, “The wild tapered block bootstrap,” Working paper, CREATES, Aarhus University.
- , 2015, “Bootstrapping integrated covariance matrix estimators in noisy jump-diffusion models with non-synchronous trading,” Working paper, CREATES, Aarhus University.
- Hounyo, U., S. Gonçalves, and N. Meddahi, 2015, “Bootstrapping pre-averaged realized volatility under market microstructure noise,” Working paper, CREATES, Aarhus University.
- 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.
- 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.
- Lee, S. S., and P. A. Mykland, 2012, “Jumps in equilibrium prices and market microstructure noise,” *Journal of Econometrics*, 168(2), 396–406.
- Li, J., V. Todorov, and G. Tauchen, 2013, “Volatility occupation times,” *Annals of Statistics*, 41(4), 1865–1891.
- , 2015, “Estimating the volatility occupation time via regularized Laplace inversion,” *Econometric Theory*, (Forthcoming).
- 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.
- Markowitz, H. M., 1952, “Portfolio selection,” *Journal of Finance*, 7(1), 77–91.
- Mykland, P. A., and L. Zhang, 2009, “Inference for continuous semimartingales observed at high frequency: A general approach,” *Econometrica*, 77(5), 1403–1445.
- , 2016, “Assessment of uncertainty in high frequency data: The observed asymptotic variance,” *Econometrica*, (Forthcoming).
- Oomen, R. C. A., 2006, “Properties of realized variance under alternative sampling schemes,” *Journal of Business and Economic Statistics*, 24(2), 219–237.
- 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.
- 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.
- Todorov, V., G. Tauchen, and I. Gryniv, 2014, “Volatility activity: Specification and estimation,” *Journal of Econometrics*, 178(1), 180–193.
- Vasicek, O., 1977, “An equilibrium characterization of the term structure,” *Journal of Financial Economics*, 5(2), 177–188.
- 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), 235–250.
- 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.