Bootstrapping high-frequency jump tests*

Prosper Dovonon
Department of Economics, Concordia University

Sílvia Gonçalves
Department of Economics, University of Western Ontario

Ulrich Hounyo
CREATES, Department of Economics and Business Economics, Aarhus University

Nour Meddahi
Toulouse School of Economics, Toulouse University

December 19, 2016

Abstract

The main contribution of this paper is to propose a bootstrap test for jumps based on functions of realized volatility and bipower variation. Bootstrap intraday returns are randomly generated from a mean zero Gaussian distribution with a variance given by a local measure of integrated volatility (which we call $\{\hat{v}_i^n\}$). We first discuss a set of high level conditions on $\{\hat{v}_i^n\}$ such that any bootstrap test of this form has the correct asymptotic size and is alternative-consistent. Our results show that the choice of $\{\hat{v}_i^n\}$ is crucial for the power of the test. In particular, we should choose $\{\hat{v}_i^n\}$ in a way that is robust to jumps. We then focus on a thresholding-based estimator for $\{\hat{v}_i^n\}$ and provide a set of primitive conditions under which our bootstrap test is asymptotically valid. We also discuss the ability of the bootstrap to provide second-order asymptotic refinements under the null of no jumps. The cumulants expansions that we develop show that our proposed bootstrap test is unable to mimic the first-order cumulant of the test statistic. The main reason is that it does not replicate the bias of the bipower variation as a measure of integrated volatility. We propose a modification of the original bootstrap test which contains an appropriate bias correction term and for which second-order asymptotic refinements obtain.

1 Introduction

A well accepted fact in financial economics is the fact that asset prices do not always evolve continuously over a given time interval, being instead subject to the possible occurrence of jumps (or discontinuous movements in prices). The detection of such jumps is crucial for asset pricing and risk management because the presence of jumps has important consequences for the performance of asset pricing models and hedging strategies, often introducing parameters that are hard to estimate (see e.g. Bakshi et al. (1997), Bates (1996), and Johannes (2004)). For this reason, many tests for jumps have been

*We are grateful for comments from participants at the SoFie Annual Conference in Toronto, June 2014, and at the IAAE 2014 Annual Conference, Queen Mary, University of London, June 2014. We are also grateful to two anonymous referees and an associate editor for many valuable suggestions. Dovonon, Gonçalves and Meddahi acknowledge financial support from a ANR-FQRSC grant. In addition, Ulrich Hounyo acknowledges support from CREATES - Center for Research in Econometric Analysis of Time Series (DNRF78), funded by the Danish National Research Foundation, as well as support from the Oxford-Man Institute of Quantitative Finance. Finally, Nour Meddahi has also benefited from the financial support of the chair "Marché des risques et création de valeur" Fondation du risque/SCOR.

proposed in the literature over the years, most of the recent ones exploiting the rich information contained in high frequency data. These include tests based on bipower variation measures (such as in Barndorff-Nielsen and Shephard (2004, 2006), henceforth BN-S (2004, 2006), Huang and Tauchen (2005), Andersen et al. (2007), Jiang and Oomen (2008), and more recently Mykland, Shephard and Sheppard (2012)); tests based on power variation measures sampled at different frequencies (such as in Aït-Sahalia and Jacod (2009), Aït-Sahalia, Jacod and Li (2012)), and tests based on the maximum of a standardized version of intraday returns (such as in Lee and Mykland (2008, 2012)). In addition, tests based on thresholding or truncation-based estimators of volatility have also been proposed with the objective of disentangling big from small jumps, as in Aït-Sahalia and Jacod (2009) and Cont and Mancini (2011), based on Mancini (2001). See Aït-Sahalia and Jacod (2012, 2014) for a review of the literature on the econometrics of high frequency-based jump tests.

In this paper, we focus on the class of tests based on bipower variation originally proposed by Barndorff-Nielsen and Shephard (2004, 2006). Our main contribution is to propose a bootstrap implementation of these tests with better finite sample properties than the original tests based on the asymptotic normal distribution. In particular, our aim is to improve finite sample size while retaining good power. In order to do so, we generate the bootstrap observations under the null of no jumps, by drawing them randomly from a mean zero Gaussian distribution with a variance given by a local measure of integrated volatility (which we call $\{\hat{v}_i^n\}$).

Our first contribution is to give a set of high level conditions on $\{\hat{v}_i^n\}$ such that any bootstrap method of this form has the correct asymptotic size and is alternative-consistent. We then verify these conditions for a specific example of $\{\hat{v}_i^n\}$ based on a threshold-based volatility estimator constructed from blocks of intraday returns which are appropriately truncated to remove the effect of the jumps. In particular, we provide primitive assumptions on the continuous price process such that the bootstrap jump test based on the thresholding local volatility estimator is able to replicate the null distribution of the BN-S test (2004, 2006) under both the null and the alternative of jumps. Our assumptions are very general, allowing for leverage effects and general activity jumps both in prices and volatility. We show that although truncation is not needed for the bootstrap jump test to control the asymptotic size under the null of no jumps, it is important to ensure that the bootstrap jump test is consistent under the alternative of jumps. Other choices of $\{\hat{v}_i^n\}$ could be considered provided they are robust to jumps. For instance, we could rely on multipower variation volatility measures rather than truncation-based methods to compute $\{\hat{v}_i^n\}$ and use our high level conditions to show the first order validity of this bootstrap method. For brevity, we focus on the thresholding-based volatility estimator, which is one of the most popular methods of obtaining jump robust test statistics.

The second contribution of this paper is to prove that an appropriate version of the bootstrap jump test based on thresholding provides a second-order asymptotic refinement under the null of no jumps. To do so, we impose more restrictive assumptions on the data generating process that assume away the presence of drift and leverage effects. For this simplified model, we develop second-order asymptotic expansions of the first three cumulants of the BN-S test statistic and of its bootstrap version. Our results show that the first-order cumulant of the BN-S test depends on the bias of bipower variation under the null of no jumps. Even though this bias does not impact the validity of the test to first order because bipower variation is a consistent estimator of integrated volatility under the null, it has an impact on the first order cumulant of the statistic at the second order (i.e. at the order $O\left(n^{-1/2}\right)$). Our bootstrap test statistic is unable to capture this higher order bias and therefore does not provide a second-order refinement. We propose a modification of the bootstrap statistic that is able to do so. Specifically, the modified bootstrap test statistic contains a correction term that is based on an estimate of the contribution to the first order cumulant of the test statistic due to the bias of bipower variation. Our simulations show that although both bootstrap versions of the test outperform the asymptotic test, the modified bootstrap test statistic has lower size distortions than the original bootstrap statistic. In the empirical application, where we apply the bootstrap jump tests to 5-minutes returns on the SPY index over the period June 15, 2004 through June 13, 2014, this version of the bootstrap test detects about half of the number of jump days detected by the asymptotic theory-based tests.

The rest of the paper is organized as follows. In Section 2, we provide the framework and state our assumptions. In Section 3, we investigate the first-order asymptotic validity of the Gaussian wild bootstrap based on a given $\{\hat{v}_i^n\}$. Specifically, Section 3.1 contains a set of high level conditions on $\{\hat{v}_i^n\}$ such that any bootstrap method is asymptotically valid when testing for jumps. Section 3.2 provides a set of primitive assumptions under which the bootstrap based on a thresholding estimator $\{\hat{v}_i^n\}$ verifies these high level conditions and is therefore asymptotically valid to first order. Section 4 investigates the ability of the bootstrap to provide asymptotic refinements. In particular, Section 4.1 contains the second-order expansions of the cumulants of the original statistic whereas Section 4.2 contains their bootstrap versions. Section 5 gives the Monte Carlo simulations while Section 6 provides an empirical application. Section 7 concludes. Appendix A contains a law of large numbers for smooth functions of consecutive local truncated volatility estimates. This result is crucial for establishing the properties of the bootstrap jump test based on the thresholding approach. It is of independent interest as it extends some existing results in the literature, namely results by Jacod and Protter (2012), Jacod and Rosenbaum (2013) and Li, Todorov and Tauchen (2016)), who focused on smooth functions of a single local volatility estimate. In addition, an online supplementary appendix contains the proofs of all the results in the main text. Specifically, Appendix S1 contains the proofs of the bootstrap consistency results presented in Section 3 whereas Appendix S2 contains the proofs of the results in Section 4 (on the asymptotic refinements of the bootstrap). Finally, Appendix S3 contains formulas for the log version of our tests.

To end this section, a word on notation. As usual in the bootstrap literature, we let P^* describe the probability of bootstrap random variables, conditional on the observed data. Similarly, we write E^* and Var^* to denote the expected value and the variance with respect to P^*, respectively. For any bootstrap statistic $Z_n^* \equiv Z_n^*(\cdot,\omega)$ and any (measurable) set A, we write $P^*(Z_n^* \in A) = P^*(Z_n^*(\cdot,\omega) \in A) = \Pr(Z_n^*(\cdot,\omega) \in A|\mathcal{X}_n)$, where $\mathcal{X}_n$ denotes the observed sample. We say that $Z_n^* \to^{P^*} 0$ in prob-P (or $Z_n^* = o_{P^*}(1)$ in prob-P) if for any $\varepsilon, \delta > 0$, $P(P^*(|Z_n^*| > \varepsilon) > \delta) \to 0$ as $n \to \infty$. Similarly, we say that $Z_n^* = O_{P^*}(1)$ in prob-P if for any $\delta > 0$, there exists $0 < M < \infty$ such that $P(P^*(|Z_n^*| \geq M) > \delta) \to 0$ as $n \to \infty$. For a sequence of random variables (or vectors) Z_n^*, we also need the definition of convergence in distribution in prob-P. In particular, we write $Z_n^* \to^{d^*} Z$, in prob-P (a.s.-P), if $E^*(f(Z_n^*)) \to E(f(Z))$ in prob-P for every bounded and continuous function f (a.s.$-P$).

2 Assumptions and statistics of interest

We assume that the log-price process X_t is an Itô semimartingale defined on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration $(\mathcal{F}_t)_{t\geq 0}$ such that

$$X_t = Y_t + J_t,\ t \geq 0, \tag{1}$$

where Y_t is a continuous Brownian semimartingale process and J_t is a jump process. Specifically, Y_t is defined by the equation

$$Y_t = Y_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s,\ t \geq 0, \tag{2}$$

where a and σ are two real-valued random processes and W is a standard Brownian semimartingale process. The jump process is defined as

$$J_t = \int_0^t \int_{\mathbb{R}} \left(\delta(s,x) I_{\{|\delta(s,x)|\leq 1\}}\right)(\mu - \nu)(ds,dx) + \int_0^t \int_{\mathbb{R}} \left(\delta(s,x) I_{\{|\delta(s,x)|>1\}}\right)\mu(ds,dx), \tag{3}$$

where μ is a Poisson random measure on $\mathbb{R}_+ \times \mathbb{R}$ with intensity measure $\nu(ds, dx) = ds \otimes \lambda(dx)$, with λ the Lebesgue measure on $\mathbb{R}$, and δ a real function on $\Omega \times \mathbb{R}_+ \times \mathbb{R}$. The first term in the definition of J_t represents the "small" jumps of the process whereas the second term represents the "big" jumps.

We make the following assumptions on a, σ and J_t, where $r \in [0, 2]$.

Assumption H-r The process a is locally bounded, σ is càdlàg, and there exists a sequence of stopping times (τ_n) and a deterministic nonnegative function γ_n on $\mathbb{R}$ such that $\int \gamma_n(x)^r \lambda(dx) < \infty$ and $|\delta(\omega, s, x)| \wedge 1 \leq \gamma_n(x)$ for all (ω, s, x) satisfying $s \leq \tau_n(\omega)$.

Assumption H-r is rather standard in this literature, implying that the r^{th} absolute power value of the jumps size is summable over any finite time interval, i.e. $\sum_{s \leq t} |\Delta X_s|^r < \infty$ for all $t > 0$. Since H-r for some r implies that H-r' holds for all $r' > r$, the weakest form of this assumption occurs for $r = 2$ (and essentially corresponds to the class of Itô semimartingales). As r decreases towards 0, fewer jumps of bigger size are allowed. In the limit, when $r = 0$, we get the case of finite activity jumps.

The quadratic variation process of X is given by $[X]_t = IV_t + JV_t$, where $IV_t \equiv \int_0^t \sigma_s^2 ds$ is the quadratic variation of Y_t, also known as the integrated volatility, and $JV_t \equiv \sum_{s \leq t} (\Delta J_s)^2$ is the jump quadratic variation, with $\Delta J_s = J_s - J_{s-}$ denoting the jumps in X. Without loss of generality, we let $t = 1$ and we omit the index t. For instance, we write $IV = IV_1$ and $JV = JV_1$.

We assume that prices are observed within the fixed time interval $[0, 1]$ (which we think of as a given day) and that the log-prices X_t are recorded at regular time points $t_i = i/n$, for $i = 0, \ldots, n$, from which we compute n intraday returns at frequency $1/n$,

$$r_i \equiv X_{i/n} - X_{(i-1)/n}, \ i = 1, \ldots, n,$$

where we omit the index n in r_i to simplify the notation.

Our focus is on testing for "no jumps" using the bootstrap. In particular, following Aït-Sahalia and Jacod (2009), we would like to decide on the basis of the observed intraday returns $\{r_i : i = 1, \ldots, n\}$ in which of the two following complementary sets the path we actually observed falls:

$$\begin{aligned} \Omega_0 &= \{\omega : t \longmapsto X_t(\omega) \text{ is continuous on } [0,1]\} \\ \Omega_1 &= \{\omega : t \longmapsto X_t(\omega) \text{ is discontinuous on } [0,1]\}, \end{aligned}$$

where $\Omega = \Omega_0 \cup \Omega_1$ and $\Omega_0 \cap \Omega_1 = \varnothing$. Formally, our null hypothesis can be defined as $H_0 : \omega \in \Omega_0$ whereas the alternative hypothesis is $H_1 : \omega \in \Omega_1$.

Let $RV_n = \sum_{i=1}^n r_i^2$ denote the realized volatility and let

$$BV_n = \frac{1}{k_1^2} \sum_{i=2}^n |r_{i-1}| |r_i|$$

be the bipower variation, where we let $k_1 \equiv E\left(\left|\chi_1^2\right|^{1/2}\right) = E(|Z|) = \sqrt{2}/\sqrt{\pi}$, where $Z \sim N(0,1)$. This is a special case of $k_q = E\left(\left|\chi_1^2\right|^{q/2}\right) = E(|Z|^q) = 2^{q/2}\frac{\Gamma\left(\frac{1+q}{2}\right)}{\Gamma\left(\frac{1}{2}\right)}$, $q > 0$.

The class of statistics we consider is based on the comparison between RV_n and BV_n. It is now well known that under certain regularity conditions including the assumption that X is continuous (see BN-S (2006) and Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006)) the following joint CLT holds:

$$\sqrt{n} \begin{pmatrix} RV_n - IV \\ BV_n - IV \end{pmatrix} \xrightarrow{st} N(0, \Sigma), \tag{4}$$

where $\xrightarrow{st}$ denotes stable convergence and

$$\Sigma = \begin{pmatrix} 2 & 2 \\ 2 & \theta \end{pmatrix} IQ, \tag{5}$$

with $IQ \equiv \int_0^1 \sigma_u^4 du$ and $\theta = \left(k_1^{-4} - 1\right) + 2\left(k_1^{-2} - 1\right) \simeq 2.6090$.

An implication of (4) is that under "no jumps", i.e. in restriction to Ω_0,

$$\frac{\sqrt{n}\left(RV_n - BV_n\right)}{\sqrt{V}} \xrightarrow{st} N\left(0,1\right),$$

where $V \equiv \tau \cdot IQ$ is the asymptotic variance of $\sqrt{n}\left(RV_n - BV_n\right)$ and $\tau = \theta - 2$. Hence, a linear version of the test is given by

$$T_n = \frac{\sqrt{n}\left(RV_n - BV_n\right)}{\sqrt{\hat{V}_n}}, \tag{6}$$

where

$$\hat{V}_n \equiv \tau \cdot \widehat{IQ}_n \quad \text{with} \quad \widehat{IQ}_n = \frac{n}{\left(k_{4/3}\right)^3} \sum_{i=3}^{n} \left|r_i\right|^{4/3} \left|r_{i-1}\right|^{4/3} \left|r_{i-2}\right|^{4/3}.$$

Choosing the tripower realized quarticity ensures that $\widehat{IQ}_n \xrightarrow{P} IQ$ on both Ω_0 and Ω_1. Thus, $T_n \xrightarrow{st} N\left(0,1\right)$, in restriction to Ω_0, and the test that rejects the null of "no jumps" at significance level α whenever $T_n > z_{1-\alpha}$, where $z_{1-\alpha}$ is the $100\left(1-\alpha\right)\%$ percentile of the $N\left(0,1\right)$ distribution has asymptotically correct strong size, i.e. the critical region $\mathcal{C}_n = \left\{T_n > z_{1-\alpha}\right\}$ is such that for any measurable set $S \subset \Omega_0$ such that $P(S) > 0$,

$$\lim_{n\to\infty} P\left(\omega \in \mathcal{C}_n | S\right) = \alpha.$$

Under the alternative hypothesis, we can show that T_n is alternative-consistent, i.e. the probability that we make the incorrect decision of "accepting the null" when this is false goes to zero:

$$\lim_{n\to\infty} P\left(\Omega_1 \cap \bar{\mathcal{C}}_n\right) = 0,$$

where $\bar{\mathcal{C}}_n$ is the complement of $\mathcal{C}_n$. Since the above condition implies that $P\left(\bar{\mathcal{C}}_n | \Omega_1\right) \to 0$, as $n \to \infty$, we have that $P\left(\mathcal{C}_n | \Omega_1\right) \to 1$ as $n \to \infty$, which we can interpret as saying that the test has asymptotic power equal to 1.

3 The bootstrap

When bootstrapping hypothesis tests, imposing the null hypothesis on the bootstrap data generating process is not only natural, but may be important to minimize the probability of a type I error. In particular, Davidson and MacKinnon (1999) (see also MacKinnon (2009)) show that in order to minimize the error in rejection probability under the null (type I error) of a bootstrap test, we should estimate the bootstrap DGP as efficiently as possible. This entails imposing the null hypothesis on the bootstrap DGP.

In this paper, we follow this rule and impose the null hypothesis of no jumps when generating the bootstrap intraday returns. Specifically, we let

$$r_i^* = \sqrt{\hat{v}_i^n} \cdot \eta_i, \qquad i = 1, \ldots, n, \tag{7}$$

for some variance measure $\hat{v}_i^n$ based on $\{r_i : i = 1, \ldots, n\}$, and where η_i is generated independently of the data as an i.i.d. $N\left(0,1\right)$ random variable. For simplicity, we again write r_i^* instead of $r_{i,n}^*$.

According to (7), bootstrap intraday returns are conditionally (on the original sample) Gaussian with mean zero and volatility $\hat{v}_i^n$, and therefore do not contain jumps. This bootstrap DGP is motivated by the simplified model $X_t = \int_0^t \sigma_s dW_s$, where σ is independent of W and there is no drift nor jumps[1]. Under these assumptions, conditionally on the path of volatility, $r_i \sim N(0, v_i^n)$, independently across i, where $v_i^n = \int_{(i-1)/n}^{i/n} \sigma_u^2 du$. Thus, we can think of $\hat{v}_i^n$ as the sample analogue of v_i^n. The main goal of Section 3.1 is to provide a set of general conditions on $\hat{v}_i^n$ under which the bootstrap is asymptotically valid. In practice, we need to choose $\hat{v}_i^n$ and our recommendation is to use a thresholding estimator that we define formally in Section 3.2.

The bootstrap analogues of RV_n and BV_n are

$$RV_n^* = \sum_{i=1}^n r_i^{*2} \quad \text{and} \quad BV_n^* = \frac{1}{k_1^2} \sum_{i=2}^n \left| r_{i-1}^* \right| \left| r_i^* \right|.$$

The first class of bootstrap statistics we consider is described as

$$T_n^* = \frac{\sqrt{n} \left(RV_n^* - BV_n^* - E^* \left(RV_n^* - BV_n^* \right) \right)}{\sqrt{\hat{V}_n^*}}, \tag{8}$$

where

$$E^* \left(RV_n^* - BV_n^* \right) = \sum_{i=1}^n \hat{v}_i^n - \sum_{i=2}^n \left(\hat{v}_{i-1}^n \right)^{1/2} \left(\hat{v}_i^n \right)^{1/2},$$

and

$$\hat{V}_n^* = \tau \cdot \widehat{IQ}_n^* \quad \text{with} \quad \widehat{IQ}_n^* = \frac{n}{\left(k_{4/3} \right)^3} \sum_{i=3}^n \left| r_i^* \right|^{4/3} \left| r_{i-1}^* \right|^{4/3} \left| r_{i-2}^* \right|^{4/3}.$$

Thus, T_n^* is exactly as T_n except for the recentering of $RV_n^* - BV_n^*$ around the bootstrap expectation $E^* \left(RV_n^* - BV_n^* \right)$. This ensures that the bootstrap distribution of T_n^* is centered at zero, as is the case for T_n under the null hypothesis of no jumps when n is large.

Nevertheless, and as we will study in Section 4, T_n has a higher order bias under the null which is not well mimicked by T_n^*, implying that this test does not yield asymptotic refinements. For this reason, we consider a second class of bootstrap statistics based on

$$\bar{T}_n^* = \frac{\sqrt{n} \left(RV_n^* - BV_n^* - E^* \left(RV_n^* - BV_n^* \right) \right)}{\sqrt{\hat{V}_n^*}} + \frac{1}{2} \frac{\sqrt{n} \left(\hat{v}_1^n + \hat{v}_n^n \right)}{\sqrt{\hat{V}_n^*}}, \tag{9}$$

where the second term accounts for the higher-order bias in T_n. This correction has an impact in finite samples, as our simulation results show. In particular, $\bar{T}_n^*$ has lower size distortions than T_n^* under the null , especially for the smaller sample sizes.

Next, we provide general conditions on $\hat{v}_i^n$ under which $T_n^* \xrightarrow{d^*} N(0,1)$, in prob-$P$ independently of whether $\omega \in \Omega_0$ or $\omega \in \Omega_1$. The consistency of the bootstrap then follows by verifying these high level conditions for a particular choice of $\hat{v}_i^n$. We verify them for a thresholding-based estimator, but other choices of $\hat{v}_i^n$ could be considered. For instance, we could rely on local multipower realized volatility estimators of v_i^n, following the approach of Mykland, Shephard and Sheppard (2012)) (see also Mykland and Zhang (2009)). Asymptotic refinements of the bootstrap based on $\bar{T}_n^*$ will be discussed in Section 4.

[1]Although (7) is motivated by this very simple model, as we will prove below, this does not prevent the bootstrap method to be valid more generally. In particular, its validity extends to the case where there is a leverage effect and the drift is non-zero.

3.1 Bootstrap validity under general conditions on $\hat{v}_i^n$

We first provide a set of conditions under which a joint bootstrap CLT holds for $(RV_n^*, BV_n^*)'$. In particular, we would like to establish that

$$\Sigma_n^{*-1/2}\sqrt{n}\begin{pmatrix} RV_n^* - E^*(RV_n^*) \\ BV_n^* - E^*(BV_n^*) \end{pmatrix} \xrightarrow{d^*} N(0, I_2),$$

in prob-P, where

$$\Sigma_n^* \equiv Var^*\left(\sqrt{n}\begin{pmatrix} RV_n^* \\ BV_n^* \end{pmatrix}\right) = \begin{pmatrix} Var^*(\sqrt{n}RV_n^*) & Cov^*(\sqrt{n}RV_n^*, \sqrt{n}BV_n^*) \\ & Var^*(\sqrt{n}BV_n^*) \end{pmatrix},$$

is such that $\Sigma_n^* \xrightarrow{P} \Sigma$. The following result gives the first and second order bootstrap moments of $(RV_n^*, BV_n^*)'$. Note that since $r_i^* = \sqrt{\hat{v}_i^n} \cdot \eta_i$, we can write

$$RV_n^* = \sum_{i=1}^n \hat{v}_i^n \cdot u_i \quad \text{and} \quad BV_n^* = \frac{1}{k_1^2}\sum_{i=2}^n \left(\hat{v}_{i-1}^n\right)^{1/2}\left(\hat{v}_i^n\right)^{1/2} \cdot w_i$$

where $u_i \equiv \eta_i^2$ and $w_i \equiv |\eta_{i-1}|\,|\eta_i|$, with $\eta_i \sim$ i.i.d. $N(0,1)$. The bootstrap moments of $(RV_n^*, BV_n^*)'$ depend on the moments and dependence properties of (u_i, w_i). The proof is trivial and is omitted for brevity.

Lemma 3.1 *If $r_i^* = \sqrt{\hat{v}_i^n} \cdot \eta_i$, $\quad i = 1, \ldots, n$, where $\eta_i \sim$ i.i.d. $N(0,1)$, then*

(a1) $E^*(RV_n^*) = \sum_{i=1}^n \hat{v}_i^n$.

(a2) $E^*(BV_n^*) = \sum_{i=2}^n \left(\hat{v}_{i-1}^n\right)^{1/2}\left(\hat{v}_i^n\right)^{1/2}$.

(a3) $Var^*(\sqrt{n}RV_n^*) = 2n\sum_{i=1}^n \left(\hat{v}_i^n\right)^2$.

(a4) $Var^*(\sqrt{n}BV_n^*) = \left(k_1^{-4} - 1\right)n\sum_{i=2}^n \left(\hat{v}_i^n\right)\left(\hat{v}_{i-1}^n\right) + 2\left(k_1^{-2} - 1\right)n\sum_{i=3}^n \left(\hat{v}_i^n\right)^{1/2}\left(\hat{v}_{i-1}^n\right)\left(\hat{v}_{i-2}^n\right)^{1/2}$.

(a5) $Cov^*(\sqrt{n}RV_n^*, \sqrt{n}BV_n^*) = n\sum_{i=2}^n \left(\hat{v}_i^n\right)^{3/2}\left(\hat{v}_{i-1}^n\right)^{1/2} + n\sum_{i=2}^n \left(\hat{v}_i^n\right)^{1/2}\left(\hat{v}_{i-1}^n\right)^{3/2}$.

Lemma 3.1 shows that the bootstrap moments of RV_n^* and BV_n^* depend on multipower variation measures of $\{\hat{v}_i^n\}$. In particular, they depend on $n^{-1+q/2}\sum_{i=K}^n \prod_{k=1}^K \left(\hat{v}_{i-k+1}^n\right)^{q_k/2}$, where K is a positive integer and $q \equiv \sum_{k=1}^K q_k$, with $q_k \geq 0$.

The following assumption imposes a convergence condition on these measures as well as other additional high level conditions on $\hat{v}_i^n$ that are sufficient for a bootstrap CLT to hold. Note that this is a high level condition that does not depend on specifying whether we are on Ω_0 or on Ω_1.

Condition A Suppose that $\{\hat{v}_i^n\}$ satisfies the following conditions:

(i) For any $K \in \mathbb{N}$ and any sequence $\{q_k \in \mathbb{R}_+ : k = 1, \ldots, K\}$ of nonnegative numbers such that $0 \leq q \equiv \sum_{k=1}^{K} q_k \leq 8$,

$$n^{-1+q/2} \sum_{i=K}^{n} \prod_{k=1}^{K} \left(\hat{v}_{i-k+1}^{n}\right)^{q_k/2} \xrightarrow{P} \int_0^1 \sigma_u^q du > 0,$$

as $n \to \infty$.

(ii) There exists $\alpha \in [0, \frac{3}{7})$ such that $n \sum_{j=1}^{[n/(L_n+1)]} \left(\hat{v}_{j(L_n+1)}^{n}\right)^2 = o_P(1)$, where $L_n \propto n^{\alpha}$ and $[x]$ denotes the largest integer smaller or equal to x.

Condition A(i) requires the multipower variations of $\hat{v}_i^n$ to converge to $\int_0^1 \sigma_u^q du$ for any $q \leq 8$. Under this condition, the probability limit of Σ_n^*, the bootstrap covariance matrix of $\sqrt{n}\left(RV_n^*, BV_n^*\right)$, is equal to Σ (for this result, convergence of the multipower variations of $\hat{v}_i^n$ with $q = 4$ suffices). Together with Condition A(i), Condition A(ii) is used to show that a CLT holds for $\sqrt{n}\left(RV_n^* - E^*\left(RV_n^*\right), BV_n^* - E^*\left(BV_n^*\right)\right)'$ in the bootstrap world. In particular, since the vector $(u_i, w_i)'$ is lag-one dependent, we adopt a large-block-small-block argument, where the large blocks are made of L_n consecutive observations and the small block is made of a single element. Part (ii) ensures that the contribution of the small blocks is asymptotically negligible. The proof of Theorem 3.1 then follows by showing that $\hat{V}_n^* = \tau \cdot \widehat{IQ}_n^* \xrightarrow{P^*} V = \tau \cdot IQ$ under Condition A(i) (this follows from the convergence of the multipower variations of $\hat{v}_i^n$ of eighth order, explaining why we require $q \leq 8$).

Under this high level condition, we can prove the following result.

Theorem 3.1 *Under Condition A, if $n \to \infty$, $T_n^* \xrightarrow{d^*} N(0,1)$, in prob-$P$.*

Since $T_n \xrightarrow{st} N(0,1)$ on Ω_0, the fact that $T_n^* \xrightarrow{d^*} N(0,1)$, in prob-$P$, ensures that the test has correct size asymptotically. Under the alternative (i.e. on Ω_1) since T_n diverges at rate $\sqrt{n}$, but we still have that $T_n^* \xrightarrow{d^*} N(0,1)$, the test has power asymptotically. More formally, let the bootstrap critical region be defined as follows,

$$\mathcal{C}_n^* = \left\{\omega : T_n(\omega) > q_{n,1-\alpha}^*(\omega)\right\},$$

where $q_{n,1-\alpha}^*(\omega)$ is such that

$$P^*\left(T_n^*(\cdot, \omega) \leq q_{n,1-\alpha}^*(\omega)\right) = 1 - \alpha.$$

The bootstrap test rejects $H_0 : \omega \in \Omega_0$ against $H_1 : \omega \in \Omega_1$ whenever $\omega \in \mathcal{C}_n^*$. The following theorem follows from Theorem 3.1 and the asymptotic properties of T_n under H_0 and under H_1.

Theorem 3.2 *Suppose $T_n \xrightarrow{st} N(0,1)$, in restriction to Ω_0, and $T_n \xrightarrow{P} +\infty$ on Ω_1. If Condition A holds, then the bootstrap test based on T_n^* controls the asymptotic strong size and is alternative-consistent.*

3.2 Bootstrap validity when $\hat{v}_i^n$ is based on thresholding

The results of the previous subsection ensure the consistency of the bootstrap distribution of T_n^* for any choice of $\hat{v}_i^n$ that verifies Condition A. In this section, we verify this condition for the following choice of $\hat{v}_i^n$:

$$\hat{v}_{j+(i-1)k_n}^{n} = \frac{1}{k_n} \sum_{m=1}^{k_n} r_{(i-1)k_n+m}^2 1\left(\left|r_{(i-1)k_n+m}\right| \leq u_n\right),$$

where $i = 1, \ldots, \frac{n}{k_n}$ and $j = 1, \ldots, k_n$. Here, k_n is an arbitrary sequence of integers such that $k_n \to \infty$ and $k_n/n \to 0$ and u_n is a sequence of threshold values defined as

$$u_n = \alpha n^{-\varpi} \quad \text{for some constant } \alpha > 0 \quad \text{and} \quad 0 < \varpi < 1/2.$$

We will maintain these assumptions on k_n and u_n throughout. The estimator $\hat{v}_i^n$ is equal to n^{-1} times a "spot volatility" estimator that is popular in the high-frequency econometrics literature under jumps (see e.g. Mancini (2001) and Aït-Sahalia and Jacod (2009)). By excluding all returns containing jumps over a given threshold when computing $\hat{v}_i^n$, we guarantee that the bootstrap distribution of T_n^* converges to a $N(0,1)$ random variable, independently of whether there are jumps or not. This is crucial for the bootstrap test to control asymptotic size and at the same time have power.

The following lemma is auxiliary in verifying Condition A.

Lemma 3.2 *Assume that X satisfies (1), (2) and (3) such that Assumption H-2 holds. Let $q = \sum_{k=1}^{K} q_k$ with $q_k \geq 0$ and $K \in \mathbb{N}$. If either of the following conditions holds:*

(a) $q > 0$ and X is continuous;

(b) $q < 2$;

(c) $q \geq 2$, Assumption H-r holds for some $r \in [0, 2)$, and $\frac{q-1}{2q-r} \leq \varpi < \frac{1}{2}$; then

$$n^{-1+q/2} \sum_{i=K}^{n} \prod_{k=1}^{K} \left(\hat{v}_{i-k+1}^n\right)^{q_k/2} \xrightarrow{P} \int_0^1 \sigma_u^q du > 0.$$

Lemma 3.2 follows from Theorem A.1 in Appendix A, a result that is of independent interest and can be seen as an extension of Theorem 9.4.1 of Jacod and Protter (2012) (see also Jacod and Rosenbaum (2013)). In particular, Theorem A.1 provides a law of large numbers for smooth functions of consecutive truncated local realized volatility estimators defined on non-overlapping time intervals. Instead, Theorem 9.4.1 of Jacod and Protter (2012) only allows for functions that depend on a single local realized volatility estimate even though they are possibly based on overlapping intervals. Recently, Li, Todorov and Tauchen (2016) focus on single local realized volatility estimate based on non-overlapping intervals and extend the limit results of Theorem 9.4.1 of Jacod and Protter (2012) to a more general class of volatility functionals that do not have polynomial growth. Here we restrict our attention to functions that have at most polynomial growth, which is enough to accommodate the blocked multipower variations measures of Lemma 3.2.

Given this result, we can state the following theorem.

Theorem 3.3 *Assume that X satisfies (1), (2), (3) such that Assumption H-2 holds. If in addition, either of the two following conditions holds:*

(a) X is continuous; or

(b) Assumption H-r holds for some $r \in [0, 2)$ and $\frac{7}{16-r} \leq \varpi < \frac{1}{2}$;

then the conclusion of Theorem 3.1 holds for the thresholding-based bootstrap test T_n^*.

Theorem 3.3 shows that the thresholding-based statistic T_n^* is asymptotically distributed as a standard normal random variable independently of whether the null of no jumps is true or not. This guarantees that the bootstrap jump test has the correct asymptotic size and is consistent under the alternative of jumps. Note that under the null, when X is continuous, the result holds for any level of truncation, including the case where $u_n = \infty$, which corresponds to no truncation. Nevertheless, to ensure that T_n^* is also asymptotically normal under the alternative hypothesis of jumps some truncation is required. Part (b) of Theorem 3.3 shows that we should choose $u_n = \alpha n^{-\varpi}$ with $\frac{7}{16-r} \leq \varpi < \frac{1}{2}$, a condition that is more stringent than the usual condition on ϖ (which is $0 < \varpi < 1/2$). The lower bound on ϖ is an increasing function of r, a number that is related to the degree of activity of jumps as specified by Assumption H-r. For finite activity jumps where $r = 0$, ϖ should be larger or equal than 7/16 but strictly smaller than 1/2. As r increases towards 2 (allowing for an increasing number of small jumps), the range of values of ϖ becomes narrower, implying that we need to choose a smaller level of truncation in order to be able to separate the Brownian motion from the jumps contributions to returns.

The following result is a corollary to Theorem 3.3.

Corollary 3.1 *Assume that X satisfies (1), (2), (3) such that Assumption H-r holds for some $r \in [0, 2)$ and let $u_n = \alpha n^{-\varpi}$ with $\frac{7}{16-r} \leq \varpi < \frac{1}{2}$. Then, the conclusions of Theorem 3.2 are true for the thresholding-based bootstrap test T_n^*.*

This result shows that the thresholding-based bootstrap jump test has the correct asymptotic size and is consistent under the alternative of jumps provided we choose a truncation level $u_n = \alpha n^{-\varpi}$ with $\frac{7}{16-r} \leq \varpi < \frac{1}{2}$, where r is $[0, 2)$. In particular, if we choose ϖ in the vicinity of 1/2, as commonly done in applications, the bootstrap test is consistent under the alternative of jumps for a wide spectrum of jump activities including finite activity. For example, if $\varpi = 0.45$, the set of r such that $\varpi \geq 7/(16 - r)$ is $[0, 0.444]$ and when $\varpi = 0.48$, this set becomes $[0, 1.417]$.

4 Second-order accuracy of the bootstrap

In this section, we investigate the ability of the bootstrap test based on the thresholding local realized volatility estimator to provide asymptotic higher-order refinements under the null hypothesis of no jumps. Our analysis is based on the following simplified model for X_t,

$$X_t = \int_0^t \sigma_s dW_s, \tag{10}$$

where σ is càdlàg locally bounded away from 0 and $\int_0^t \sigma_s^2 ds < \infty$ for all $t \in [0, 1]$. In addition, we assume that σ is independent of W. Thus, we not only impose the null hypothesis of no jumps under which $J_t = 0$, but we also assume that there is no drift nor leverage effects. Under these assumptions, conditionally on the path of volatility, $r_i \sim N(0, v_i^n)$ independently across i, a result that we will use throughout this section. Allowing for the presence of drift and leverage effects would complicate substantially our analysis. In particular, we would not be able to condition on the volatility path σ when deriving our expansions if we relaxed the assumption of independence between σ and W. Allowing for the presence of a drift would require a different bootstrap method, the main reason being that the effect of the drift on the test statistic is of order $O\left(n^{-1/2}\right)$ and our bootstrap returns have mean zero by construction (see Gonçalves and Meddahi, 2009). We leave these important extensions for future research.

To study the second-order accuracy of the bootstrap, we rely on second-order Edgeworth expansions of the distribution of our test statistics T_n and T_n^*. As is well known, the coefficients of the polynomials entering a second-order Edgeworth expansion are a function of the first three cumulants of the test statistics (cf. Hall, 1992). In order to derive these higher-order cumulants, we make the following additional assumption. We rely on it to obtain the limit of the first order cumulant of T_n (cf. $\kappa_{1,1}$ below).

Assumption V The volatility process σ_u^2 is pathwise continuous, bounded away from zero and Holder-continuous in $L^2(P)$ on $[0,1]$ of order $\delta > 1/2$, i.e., $E\left(\left(\sigma_u^2 - \sigma_s^2\right)^2\right) = O(|u-s|^{2\delta})$.

Thus, we not only impose that the volatility path is continuous, but we also rule out stochastic volatility models driven by a Brownian motion. Examples of processes that satisfy Assumption V include fractional Brownian motion with Hurst parameter $H > 1/2$.

4.1 Second-order expansions of the cumulants of T_n

Next we provide asymptotic expansions for the cumulants of T_n. For any positive integer i, let $\kappa_i(T_n)$ denote the i^{th} cumulant of T_n. In particular, recall that $\kappa_1(T_n) = E(T_n)$, $\kappa_2(T_n) = Var(T_n)$ and $\kappa_3(T_n) = E(T_n - E(T_n))^3$. In addition, for any $q > 0$, we let $\overline{\sigma^q} = \int_0^1 \sigma_u^q du$.

Theorem 4.1 *Assume that X satisfies (10) and Assumption V holds, where σ is independent of W. Then, conditionally on σ, we have that*

$$\begin{aligned}
\kappa_1(T_n) &= \underbrace{\frac{1}{\sqrt{n}}\left(\frac{1}{2\sqrt{\tau}}\frac{\sigma_0^2+\sigma_1^2}{\sqrt{\overline{\sigma^4}}} - \frac{a_1}{2}\frac{\overline{\sigma^6}}{\left(\overline{\sigma^4}\right)^{3/2}}\right)}_{\equiv\kappa_1=\kappa_{1,1}+\kappa_{1,2}} + O\left(\frac{1}{n}\right); \\
\kappa_2(T_n) &= 1 + O\left(\frac{1}{n}\right); \text{ and} \\
\kappa_3(T_n) &= \underbrace{\frac{1}{\sqrt{n}}\left(a_2 + \frac{3}{2}(a_1 - a_3)\right)\frac{\overline{\sigma^6}}{\left(\overline{\sigma^4}\right)^{3/2}}}_{\equiv\kappa_3} + O\left(\frac{1}{n}\right),
\end{aligned}$$

where $\tau = \theta - 2 = \left(k_1^{-4} - 1\right) + 2\left(k_1^{-2} - 1\right) - 2$ and the constants a_1, a_2 and a_3 also depend on $k_q = E|Z|^q$, $Z \sim N(0,1)$, for certain values of $q > 0$; their specific values are given in Lemma S2.5 in the Appendix.

Theorem 4.1 shows that the first and third order cumulants of T_n are subject to a higher order bias of order $O\left(n^{-1/2}\right)$, given by the constants κ_1 and κ_3. Since the asymptotic normal approximation assumes that the values of these cumulants are zero, this induces an error of order $O\left(n^{-1/2}\right)$ for the asymptotic normal approximation when approximating the null distribution of T_n.

The bootstrap is asymptotically second-order accurate if the bootstrap first and third order cumulants mimic κ_1 and κ_3. As it turns out, this is not true for the bootstrap test based on T_n^*. The main reason is that it fails to capture $\kappa_{1,1}$, a bias term that is due to the fact that bipower variation is a biased (but consistent) estimator of IV. To understand how this bias impacts the first order cumulant of T_n, note that we can write

$$T_n = \frac{\sqrt{n}\,(RV_n - BV_n)}{\sqrt{\hat{V}_n}} = (S_n + A_n)\left(1 + \frac{1}{\sqrt{n}}(U_n + B_n)\right)^{-1/2}, \tag{11}$$

where

$$\begin{aligned} S_n &= \frac{\sqrt{n}\left(RV_n - BV_n - E\left(RV_n - BV_n\right)\right)}{\sqrt{V_n}}; \\ A_n &= \frac{\sqrt{n}E\left(RV_n - BV_n\right)}{\sqrt{V_n}};\ U_n = \frac{\sqrt{n}\left(\hat{V}_n - E\hat{V}_n\right)}{\sqrt{V_n}}; \\ B_n &= \frac{\sqrt{n}\left(E\hat{V}_n - V_n\right)}{\sqrt{V_n}},\ \text{and } V_n = Var\left(\sqrt{n}\left(RV_n - BV_n\right)\right). \end{aligned}$$

By construction, conditionally on σ, $E\left(S_n\right) = 0$ and $Var\left(S_n\right) = 1$; the variable S_n drives the usual asymptotic normal approximation. The term A_n is deterministic (conditionally on σ) and reflects the fact that $E\left(RV_n - BV_n\right) \neq 0$ under the null of no jumps. In particular, we can easily see that $E\left(RV_n - BV_n\right) = IV - E\left(BV_n\right)$. Thus, A_n reflects the bias of BV_n as an estimator of IV. We can show that $A_n = O\left(n^{-1/2}\right)$, implying that to order $O\left(n^{-1}\right)$, the first-order cumulant of T_n is

$$\kappa_1\left(T_n\right) = \frac{1}{\sqrt{n}}\underbrace{\left(\sqrt{n}A_n - \frac{1}{2}E\left(S_n U_n\right)\right)}_{\longrightarrow \kappa_{1,1}+\kappa_{1,2}\equiv\kappa_1} + O\left(\frac{1}{n}\right).$$

The limit of $\sqrt{n}A_n$ is $\kappa_{1,1}$. This follows by writing

$$\sqrt{n}A_n = \frac{nE\left(IV - BV_n\right)}{\sqrt{V_n}} = \frac{n}{\sqrt{V_n}}\left(\sum_{i=1}^{n} v_i^n - \sum_{i=2}^{n}\left|v_{i-1}^n\right|^{1/2}\left|v_i^n\right|^{1/2}\right),$$

where $IV = \sum_{i=1}^{n} v_i^n$, and noting that by Lemma S2.3 (in Appendix S2), under Assumption V,

$$n\left(\sum_{i=1}^{n} v_i^n - \sum_{i=2}^{n}\left|v_{i-1}^n\right|^{1/2}\left|v_i^n\right|^{1/2}\right) \xrightarrow{P} \frac{1}{2}\left(\sigma_0^2 + \sigma_1^2\right), \tag{12}$$

and $V_n \xrightarrow{P} \tau\overline{\sigma^4}$.

Next we show that the bootstrap test based on T_n^* does not replicate $\kappa_{1,1}$ and therefore is not second-order correct. We then propose a correction of this test and show that it matches κ_1 and κ_3.

4.2 Second-order expansions of the bootstrap cumulants

Let κ_{1n}^* and κ_{3n}^* denote the leading terms of $\kappa_1^*\left(T_n^*\right)$ and $\kappa_3^*\left(T_n^*\right)$, the first and third order cumulants of T_n^*, respectively. In particular,

$$\kappa_1^*\left(T_n^*\right) = \frac{1}{\sqrt{n}}\kappa_{1n}^* + o_P\left(\frac{1}{\sqrt{n}}\right) \text{ and } \kappa_3^*\left(T_n^*\right) = \frac{1}{\sqrt{n}}\kappa_{3n}^* + o_P\left(\frac{1}{\sqrt{n}}\right),$$

where κ_{1n}^* and κ_{3n}^* depend on n since they are a function of the original sample. Their probability limits are denoted by κ_1^* and κ_3^* and the following theorem derives their values.

Theorem 4.2 *Assume that X satisfies (10) and Assumption V holds, where σ is independent of W. Suppose that $k_n \to \infty$ such that $k_n/n \to 0$, $\sqrt{n}/k_n$ is bounded and u_n is a sequence of threshold values defined as $u_n = \alpha n^{-\varpi}$ for some constant $\alpha > 0$ and $0 < \varpi < 1/2$. Then, conditionally on σ, we have*

$$\kappa_1^* = \kappa_{1,2} \neq \kappa_1 \text{ and } \kappa_3^* = \kappa_3$$

where $\kappa_{1,2}$, κ_1 and κ_3 are defined as in Theorem 4.2.

Theorem 4.2 shows that the bootstrap test based on T_n^* only captures the first order cumulant κ_1 partially and therefore fails to provide a second order asymptotic refinement. The main reason is that by construction the bootstrap analogue of A_n (which we denote by A_n^*) is zero for T_n^*. Because the original test has $A_n \neq 0$, the bootstrap fails to capture this source of uncertainty. Note that the conditions on u_n used by Theorem 4.2 specify that $\varpi \in (0, 1/2)$, but the result actually follows under no restrictions on u_n since we assume that X is continuous (this explains also why we do not require strengthening the restrictions on ϖ as we did when proving Theorem 3.3).

Our solution is to add a bias correction term to T_n^* that relies on the explicit form of the limit of $\sqrt{n}A_n$. In particular, our adjusted bootstrap statistic is given by

$$\bar{T}_n^* = \frac{\sqrt{n}\left(RV_n^* - BV_n^* - E^*\left(RV_n^* - BV_n^*\right)\right)}{\sqrt{\hat{V}_n^*}} + \frac{1}{2}\frac{\sqrt{n}\left(\hat{v}_1^n + \hat{v}_n^n\right)}{\sqrt{\hat{V}_n^*}} = T_n^* + \bar{R}_n^*,$$

where $\bar{R}_n^*$ can be written as $\bar{R}_n^* = \sqrt{\frac{V_n^*}{\hat{V}_n^*}}A_n^*$ with

$$A_n^* = \frac{1}{2}\frac{\sqrt{n}\left(\hat{v}_1^n + \hat{v}_n^n\right)}{\sqrt{V_n^*}}.$$

Since $n\hat{v}_i^n$ is equal to a spot volatility estimator, it follows that

$$\sqrt{n}A_n^* = \frac{1}{2}\frac{n\left(\hat{v}_1^n + \hat{v}_n^n\right)}{\sqrt{V_n^*}} \xrightarrow{P} \frac{1}{2}\frac{\left(\sigma_0^2 + \sigma_1^2\right)}{\sqrt{\tau\overline{\sigma^4}}} \equiv \kappa_{1,1}$$

under our assumptions. Hence, $\bar{T}_n^*$ is able to replicate the first and third order cumulants through order $O\left(n^{-1/2}\right)$ and therefore provides a second-order refinement. The following theorem provides the formal derivation of the cumulants of $\bar{T}_n^*$. We let $\bar{\kappa}_1^*$ and $\bar{\kappa}_3^*$ denote the probability limits of $\bar{\kappa}_{1n}^*$ and $\bar{\kappa}_{3n}^*$, the leading terms of the first-order and third-order bootstrap cumulants of $\bar{T}_n^*$.

Theorem 4.3 *Under the same assumptions as Theorem 4.2, conditionally on σ, we have*

$$\bar{\kappa}_1^* = \kappa_1 \text{ and } \bar{\kappa}_3^* = \kappa_3$$

where κ_1 and κ_3 are defined as in Theorem 4.2.

5 Monte Carlo simulations

In this section, we assess by Monte Carlo simulations the performance of our bootstrap tests. Along with the asymptotic test of BN-S (2006), we report bootstrap results using $\hat{v}_i^n$ based on the thresholding estimator (cf. Section 3.2). We follow Jacod and Rosenbaum (2013) and set $k_n = [\sqrt{n}]$, the integer part of $\sqrt{n}$. As their results show (see also Jacod and Protter (2012)), this yields the optimal rate of convergence for the spot volatility estimator $n\hat{v}_i^n$. We also follow Podolskij and Ziggel (2010) and choose $\varpi = 0.4$ and $\alpha = 2.3\sqrt{BV_n}$ for the truncation parameters.

We present results for the SV2F model given by[2]

$$\begin{aligned}
dX_t &= adt + \sigma_{u,t}\sigma_{sv,t}dW_t + dJ_t,\\
\sigma_{u,t} &= C + A \cdot exp\left(-a_1 t\right) + B \cdot exp\left(-a_2\left(1-t\right)\right),\\
\sigma_{sv,t} &= s\text{-}exp\left(\beta_0 + \beta_1\tau_{1,t} + \beta_2\tau_{2,t}\right),\\
d\tau_{1,t} &= \alpha_1\tau_{1,t}dt + dB_{1,t},\\
d\tau_{2,t} &= \alpha_2\tau_{2,t}dt + \left(1 + \phi\tau_{2,t}\right)dB_{2,t},\\
corr\left(dW_t, dB_{1,t}\right) &= \rho_1,\ corr\left(dW_t, dB_{2,t}\right) = \rho_2.
\end{aligned}$$

[2]The function s-exp is the usual exponential function with a linear growth function splined in at high values of its argument: s-exp$(x) = \exp(x)$ if $x \leq x_0$ and s-exp$(x) = \frac{\exp(x_0)}{\sqrt{x_0 - x_0^2 + x^2}}$ if $x > x_0$, with $x_0 = \log(1.5)$.

The processes $\sigma_{u,t}$ and $\sigma_{sv,t}$ represent the components of the time-varying volatility in prices. In particular, $\sigma_{sv,t}$ denotes the two factors stochastic volatility model commonly used in this literature. We follow Huang and Tauchen (2005) and set $a = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$, $\rho_1 = \rho_2 = -0.3$. At the start of each interval, we initialize the persistent factor τ_1 by $\tau_{1,0} \sim N\left(0, \frac{-1}{2\alpha_1}\right)$, its unconditional distribution. The strongly mean-reverting factor τ_2 is started at $\tau_{2,0} = 0$. The process $\sigma_{u,t}$ models the diurnal U-shaped pattern in intraday volatility. In particular, we follow Hasbrouck (1999) and Andersen et al. (2012) and set the constants $A = 0.75$, $B = 0.25$, $C = 0.88929198$, and $a_1 = a_2 = 10$. These parameters are calibrated so as to produce a strong asymmetric U-shaped pattern, with variance at the open (close) more than 3 (1.5) times that at the middle of the day. Setting $C = 1$ and $A = B = 0$ yields $\sigma_{u,t} = 1$ for $t \in [0, 1]$ and rules out diurnal effects from the observed process X. In our experiment, we first consider a setup without diurnal effects followed by one with diurnal effects in X. Finally, J_t is a finite activity jump process modeled as a compound Poisson process with constant jump intensity λ and random jump size distributed as $N(0, \sigma^2_{jmp})$. We let $\sigma^2_{jmp} = 0$ under the null hypothesis of no jumps in the return process. Under the alternative, we let $\lambda = 0.058$, and $\sigma^2_{jmp} = 1.7241$. These parameters are motivated by empirical studies by Huang and Tauchen (2005) and Barndorff-Nielsen, Shephard, and Winkel (2006), which suggest that the jump component accounts for 10% of the variation of the price process.

We simulate data for the unit interval $[0, 1]$ and normalize one second to be $1/23,400$, so that $[0, 1]$ is meant to span 6.5 hours. The observed process X is generated using an Euler scheme. We then construct the $1/n$-horizon returns $r_i = X_{i/n} - X_{(i-1)/n}$ based on samples of size n. Results are presented for four different samples sizes: $n = 48$, 96, 288, and 576, corresponding approximately to "8-minute", "4-minute", "1,35-minute", and "40-second" frequencies.

Table 1 gives the rejection rates. We report results without jumps and with finite activity jumps. Test results from both the linear test statistic and its log version[3] are reported using asymptotic-theory based critical value as well as bootstrap critical values. All tests are carried out at 5% nominal level. The rejection rates reported in the left part of Table 1 (under no jumps) are obtained from 10,000 Monte Carlo replications with 999 bootstrap samples for each simulated sample for the bootstrap tests. For finite activity jumps, since J_t is a compound Poisson process, even under the alternative, it is possible that no jump occurs in some sample over the interval [0,1] considered. Thus, to compute the rejection rates under the alternative of jumps we rely on the number n_0 of replications, out of 10,000, for which at least one jump has occurred. For our parameter configuration, $n_0 = 570$.

Starting with size, the results show that the linear version of the test based on the asymptotic theory of BN-S (2006) (labeled "AT" in Table 1) is substantially distorted for the smaller sample sizes. In particular, for the SV2F model without diurnal effects, the rejection rate is three times larger than the nominal level of the test (at 15.69%) for $n = 48$. Although this rate drops as n increases, it remains significantly larger than the nominal level even when $n = 576$, with a value equal to 8.27%. As expected, the log version of the test statistic has smaller size distortions: the rejection rates are now 13.04% and 7.68% for $n = 48$ and $n = 576$, respectively. The rejection rates of the bootstrap tests are always smaller than those of the asymptotic tests and therefore the bootstrap outperforms the latter under the null. This is true for both bootstrap jump tests based on (8) and (9) (denoted "Boot1" and "Boot2", respectively) and for both the linear and the log versions of the test. Note that our bias correction adjustment of the bootstrap test is specific to the linear version of the statistic (as it depends on its cumulants). Since we have not developed cumulant expansions for the log version of the statistic, we do not report the analogue of "Boot2" for this test.

When X has diurnality patterns in volatility, we apply the tests to both raw returns and to transformed returns with volatility corrected for diurnal patterns. We use the nonparametric jump robust estimation of intraday periodicity in volatility suggested by Boudt, Croux and Laurent (2011)

[3] See the online Appendix S3 for details on the log-transform of the test statistic T_n and the bootstrap-related formulas.

for diurnal patterns correction. In the process, standardized returns are obtained using an estimate $\hat{\sigma}_{u,i}$ of intraday volatility pattern from 2,000 simulated days. The results for the tests based on the raw returns (without diurnality correction) appear in the middle panel of Table 1 whereas the bottom panel contains results for tests based on the transformed returns. We can see that the test based on the asymptotic theory of BN-S (2006) has large distortions driven by the difference in volatility across blocks, even if the sample size is large. For $n = 576$, the null rejection rate is 13.29% for the linear version of the test and 11.74% for the log version. These are more than twice as large as the desired nominal level of 5%. The overrejection is magnified for smaller sample sizes. For instance, for $n = 48$ they are equal to 32.61% and 28.91%, respectively. As expected, (in the bottom part of Table 1) corrections for diurnal effects help reduce the distortions. For $n = 48$, the rates are now equal to 14.82% and 12.10%, whereas for $n = 576$ they are 8.47% and 7.91%. The bootstrap null rejection rates are always smaller than those of the asymptotic theory-based tests, implying that the bootstrap outperforms the latter. This is true even for the bootstrap test applied to the non-transformed intraday returns, which yields rejection rates that are closer to the nominal level than those obtained with the asymptotic tests based on the correction of the diurnal effects (compare "Boot2" in the middle panel with "AT" in the bottom panel). This is a very interesting finding since it implies that our bootstrap method is more robust to the presence of diurnal effects than the asymptotic theory-based tests. Of course, even better results can be obtained for the bootstrap tests by resampling the transformed intraday returns and this is confirmed by Table 1, which shows that the results for bootstrap tests (especially Boot2) with diurnal effects correction are systematically closer to 5% than those with no correction of diurnal effects.

These results also reveal that Boot2 outperforms Boot1, in particular for small sample sizes. This shows that taking into account the asymptotically negligible bias in T_n, only relevant at the second-order, is very useful for smaller values of n.

Overall, the left panel of Table 1 shows that the bootstrap reduces dramatically the size distortions that we can see from asymptotic tests and this across sample sizes whether the linear or log version of the test is used.

Turning now to the power analysis, results in Table 1 (right panel) show that the main feature of notice is that the bootstrap tests have lower power than their asymptotic counterparts, especially in presence of diurnal effects. This is expected given that the asymptotic tests have much larger rejections under the null than the bootstrap tests. In particular, this explains the large discrepancy between the bootstrap and the asymptotic test when both are applied to the non-transformed data. As n increases, we see that this difference decreases. The results also show that power is largest for tests (both asymptotic and bootstrap-based) applied to the transformed returns. For these tests, the difference in power between the bootstrap and the asymptotic tests is very small. Given that the bootstrap essentially eliminates the size distortions of the asymptotic test, these two findings strongly favor the bootstrap over the asymptotic tests.

Overall, Table 1 shows that Boot2 is the best choice. This is especially true when using smaller values of n. Therefore, our recommendation is to choose Boot2.

6 Empirical results

This empirical application uses trade data on the SPDR S&P 500 ETF (SPY), which is an exchange traded fund (ETF) that tracks the S&P 500 index. Data on SPY have been used by Mykland, Shephard and Sheppard (2012) (see also Bollerslev, Law and Tauchen (2008)). Our primary sample comprises 10 years of trade data on SPY starting from June 15, 2004 through June 13, 2014 as available in the New York Stock Exchange Trade and Quote (TAQ) database. This tick-by-tick dataset has been cleaned according to the procedure outlined by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009). We

Table 1: Rejection rates of of asymptotic and bootstrap tests, nominal level $\alpha = 0.05$.

	Size					Power				
	Linear test			Log test		Linear test			Log test	
n	AT	Boot1	Boot2	AT	Boot1	AT	Boot1	Boot2	AT	Boot1
	SV2F model without diurnal effects, no jumps					SV2F model without diurnal effects, jumps				
48	15.69	7.20	5.99	13.04	7.19	80.26	73.50	72.18	78.20	73.31
96	12.81	6.87	5.99	11.20	7.07	83.27	79.14	78.76	82.33	78.95
288	9.81	6.43	5.87	8.90	6.32	88.16	87.03	86.65	87.78	87.03
576	8.27	5.91	5.56	7.68	5.83	88.53	88.16	88.16	88.53	88.16
	SV2F model with diurnal effects, no correction, no jumps					SV2F model with diurnal effects, no correction, jumps				
48	32.61	16.31	14.28	28.91	14.68	86.09	78.20	77.26	85.53	77.82
96	25.28	13.98	12.26	22.23	14.32	86.65	82.71	81.02	85.34	82.52
288	16.42	10.01	8.95	14.39	9.53	88.91	86.65	86.28	88.35	86.47
576	13.29	8.64	7.99	11.74	8.41	88.72	87.59	87.22	88.16	87.59
	SV2F model with diurnal effects, correction, no jumps					SV2F model with diurnal effects, correction, jumps				
48	14.82	6.69	5.48	12.10	6.73	91.03	89.14	88.79	90.52	89.14
96	12.47	6.83	6.01	10.97	6.86	92.41	91.38	91.03	92.41	91.21
288	9.94	6.31	5.59	8.93	6.25	94.48	93.79	93.62	94.31	93.79
576	8.47	6.04	5.59	7.91	5.96	94.14	93.97	93.97	94.14	93.97

Notes: 'AT' is based on (6), i.e., the asymptotic theory of BN-S (2006); 'Boot1' and 'Boot2' are based on bootstrap test statistics T_n^* (cf. (8)) and $\bar{T}_n^*$ (cf. (9)), respectively. 'Boot2' takes into account the asymptotically negligible bias in T_n which may be relevant at the second-order, and under certain conditions provides the refinement for the bootstrap method. We use 10,000 Monte Carlo trials with 999 bootstrap replications each.

also removed short trading days leaving us with 2497 days of trade data.

Figure 1 shows the series of daily returns on SPY over the 2497 trading days considered. The 2008 financial crisis is noticeable with large returns appearing in the third quarter of 2008 and the first quarter of 2009. We can actually distinguish three subperiods for SPY: '*Before crisis*', from the beginning of the sample (June 15, 2004) through August 29 2008 (1053 trading days); '*Crisis*', from September 2, 2008 through May 29, 2009 (185 trading days), and '*After crisis*', from June 1, 2009 through June 13, 2014 (1259 trading days).

Table 2 (left panel) gives some summary statistics on daily returns and 5-min-return-based realized volatility (RV) and realized bipower variation (BV) over the mentioned periods. The average daily returns before and after the crisis are positive (1.53 and 6.42 basis points, respectively) whereas the average return over the crisis is -12.9 basis points. Daily averages of RV and BV are also quite high during the crisis period with both culminating to 6 times their respective levels across the whole sample. The average contribution of jumps to realized volatility as measured by $RJ = (RV - BV)/RV$ also deepens during the crisis period to 5%, whereas the 7% found for the full sample and in pre- and post-crisis periods seems in line with the findings of Huang and Tauchen (2005) for S&P 500 future index.

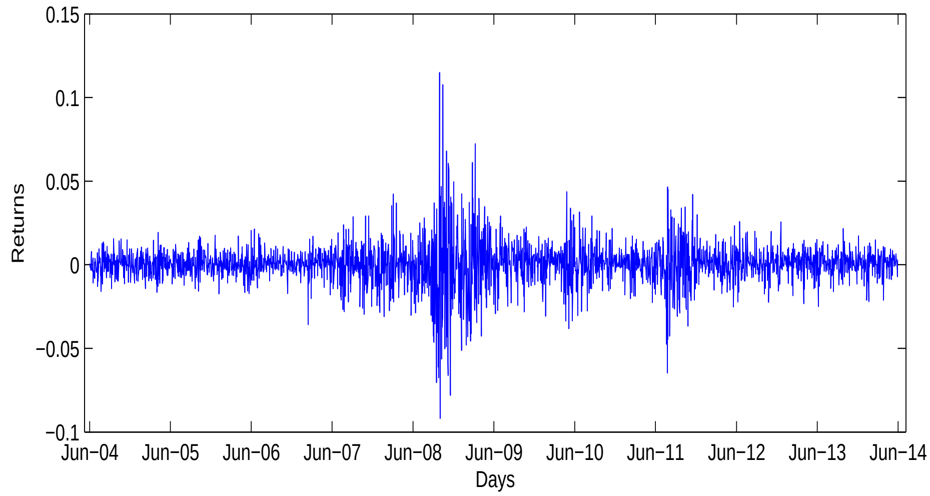

Figure 1: Daily returns on SPY from June 15, 2004 through June 13, 2014.

Table 2: This table gives the average daily return, realized volatility (RV), realized bipower variation (BV) and the contribution of jumps to realized volatility (RJ) of SPY over each period along with their standard deviations (SD). RV and BV are based on 5-min intra-day returns. These statistics are also reported over days identified with and without jumps by the asymptotic approximation of the log-test-statistic and the bootstrap approximation of the linear test statistic. ($\alpha = 0.05$). In the table, 'Ret.' stands for return.

	Ret. $\times 10^4$	$RV \times 10^4$	$BV \times 10^4$	RJ	Ret. $\times 10^4$	$RV \times 10^4$	$BV \times 10^4$	RJ
	Full sample: June 15, 2004 through June 13, 2014 (2497 days)				*Days identified with jumps by the asymptotic log test (581 days)*			
Mean	2.93	0.99	0.95	0.07	10.69	0.82	0.64	0.22
SD	126.00	2.60	2.52	0.11	129.63	1.96	1.52	0.07
	Before crisis: June 15, 2004 through August 29, 2008 (1053 days)				*Days identified without jumps by the asymptotic log test (1916 days)*			
Mean	1.53	0.55	0.51	0.07	0.58	1.05	1.04	0.02
SD	86.91	0.66	0.64	0.11	124.82	2.76	2.75	0.08
	During crisis: September 2, 2008 through May 29, 2009 (185 days)				*Days identified with jumps by the bootstrap linear test (342 days)*			
Mean	-12.90	6.06	5.82	0.05	14.06	0.81	0.60	0.25
SD	313.03	7.30	7.03	0.11	140.96	1.95	1.44	0.07
	After crisis: June 1, 2009 through June 13, 2014 (1259 days)				*Days identified without jumps by the bootstrap linear test (2155 days)*			
Mean	6.42	0.63	0.60	0.07	1.16	1.02	1.00	0.04
SD	103.94	1.07	1.14	0.12	123.41	2.68	2.65	0.09

Table 3 shows the percentage of days identified with a jump ("jump days") by the asymptotic and bootstrap tests. We consider the asymptotic version of the linear and the log test statistics as well as their bootstrap versions. For the linear bootstrap test, we rely on "Boot2", the adjusted bootstrap statistic that promises second-order refinements (and which does best in finite samples according to our simulations). For the log version of the bootstrap test, we rely on "Boot1". These tests are applied to data with and without correction for diurnal effects and are based on 5-min returns throughout. This yields 78 daily observations over the 6.5 hours of the trading session.

In line with the simulation findings, the asymptotic tests tend to substantially over detect jumps compared to the bootstrap tests, which throughout detect about half of the number of jump days detected by the asymptotic tests. More precisely, with no account for diurnal effects, the asymptotic (linear and log) tests detect 26.31% and 23.27% jump days, respectively out of the 2497 days in our sample, while the bootstrap tests detect 13.7% and 16.9% jump days. These percentages are about the same as what is obtained before and after crisis. During the crisis though, less jump days (in proportion) are detected, with the asymptotic tests detecting around 20%, while the bootstrap linear and log tests detect about 10.8% and 13.3% jump days, respectively.

Given the results of the jump tests (both asymptotic and bootstrap-based), we can compute the summary statistics for days with and without jumps. The results are contained in the right panel of Table 2. Besides the fact that the bootstrap finds less jump days, average returns are higher by around 4 basis points on bootstrap-jump-days with higher standard deviation. The average contribution of jumps to realized volatility is substantially higher on jump days than on no-jump-days by a ratio of about 10-to-1 for the asymptotic (log) test and 5-to-1 for the bootstrap test.

Table 3: Percentage of days identified as jumps day by daily statistics (nominal level $\alpha = 0.05$) using intra-day 5-min returns.

No correction for diurnal effects				With correction for diurnal effects			
AT-lin	AT-log	Boot2-lin	Boot1-log	AT-lin	AT-log	Boot2-lin	Boot1-log
Full sample: June 15, 2004 through June 13, 2014 (2497 days)							
26.31	23.27	13.70	16.90	24.23	20.54	12.66	14.46
Before crisis: June 15, 2004 through August 29, 2008 (1053 days)							
25.55	22.41	13.11	16.43	22.41	18.99	12.73	14.06
During crisis: September 2, 2008 through May 29, 2009 (185 days)							
21.62	19.46	10.81	13.51	24.32	21.62	11.35	12.97
After crisis: June 1, 2009 through June 13, 2014 (1259 days)							
27.64	24.54	14.61	17.79	25.73	21.68	12.79	15.01

Notes: 'AT-lin' and 'Boot2-lin' ('AT-log' and 'Boot1-log') stand for asymptotic and bootstrap tests using the linear (log) version of the test statistic. 'Boot2-lin' test uses the second-order corrected bootstrap test statistic for asymptotic refinement.

We also report test results applied to returns corrected for diurnal effects. This is of particular relevance in the current application since, as shown by Figure 2, diurnal patterns seem to be in display in our samples. Figure 2 displays graphs of average absolute 5-min returns over the days in the specified sample. (See Andersen and Bollerslev (1997).) The U-shape of these graphs highlights the fact that

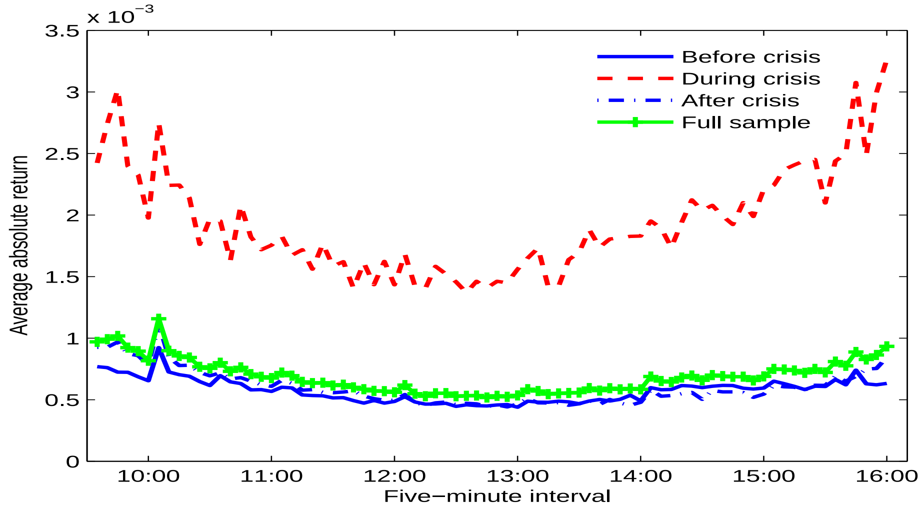

Figure 2: Diurnal pattern of SPY. The graph displays the average (over the specified samples) of absolute 5-min intraday returns of each trading day. 'Before crisis' refers to the sample from June 15, 2004 through August 29, 2008 (before the 2008 financial crash). 'During crisis' refers to the period from September 2, 2008 through May 29, 2009 and 'After crisis' refers to the period from June 1, 2009 through June 13, 2014.

the market seems to be more volatile early and late in daily trading sessions compared to the mid-day volatility. We can also see that the gap between early/late and mid-day volatilities is magnified in the crisis period. After correction for diurnal effects, less jumps days are detected by all the tests before and after crisis. For instance, before the crisis, the asymptotic tests move from 25.55% and 22.41% to 22.41% and 18.99% whereas the bootstrap tests move from 13.11% to 12.73% for the linear test and from 16.43% to 14.06% for the log test. However, in the crisis period, while the bootstrap still detects about the same number of jump days, the asymptotic tests detect substantially more jumps after diurnal effects correction. It is also worthwhile to point out that the gap between the bootstrap linear and log tests narrows as diurnal effects are accounted for. Overall, the bootstrap tests seem more robust to diurnality than the asymptotic tests.

7 Conclusion

The main contribution of this paper is to propose bootstrap methods for testing the null hypothesis of "no jumps". The methods generate bootstrap intraday returns from a Gaussian distribution with variance given by a local realized measure of integrated volatility $\{\hat{v}_i^n\}$. We first provide a set of high level conditions on $\{\hat{v}_i^n\}$ such that any bootstrap method of this form is asymptotically valid when testing for jumps using the test statistic proposed by Barndorff-Nielsen and Shephard (2006). This means in particular that the bootstrap is able to control size and is consistent under the alternative of jumps. Our results show that the choice of $\{\hat{v}_i^n\}$ is crucial for ensuring that the bootstrap test is asymptotically normally distributed, independently of whether the null or the alternative hypothesis is true. In particular, to ensure that this holds under the alternative, we should choose $\{\hat{v}_i^n\}$ in a manner that is robust to jumps. A popular estimator is the thresholding estimator and we provide a detailed analysis of the bootstrap test based on this choice of $\{\hat{v}_i^n\}$.

A second contribution of this paper is to discuss the ability of the bootstrap to provide second-order asymptotic refinements over the usual asymptotic mixed Gaussian distribution under the null of no jumps. We develop second-order asymptotic expansions of the cumulants of the test statistic of Barndorff-Nielsen and Shephard (2006) and show that the bias inherent in bipower variation as an estimator of integrated volatility has an impact on the first-order cumulant of this test, up to order $O\left(n^{-1/2}\right)$. More importantly, our bootstrap test is not able to match this cumulant effect and is therefore not second-order accurate. We then propose a modification of the original bootstrap test for which an asymptotic refinement exists. The modification consists of adding a bias correction term that estimates the contribution of the bipower variation bias to the first-order cumulant of the original test. Our simulations show that this adjustment is important in finite samples, especially for the smaller sample sizes when sampling is more sparse.

An interesting finding from the Monte Carlo simulations is that the bootstrap is more robust to the presence of diurnality effects in volatility than the usual asymptotic approximations. In particular, the adjusted bootstrap test statistic applied to raw returns generated from a two-factor stochastic volatility model with diurnal patterns in volatility has size properties that are analogous to those of the asymptotic test based on standardized returns constructed from on a nonparametric estimate of the diurnality pattern. Applying the bootstrap to the standardized returns yields even better size control. We also illustrate the usefulness of our bootstrap jumps test by applying it to 5-min returns on the SPY index over the period from June 15, 2004 through June 13, 2014. Overall, the main finding is that the bootstrap detects about half of the number of jump days detected by the asymptotic-theory based tests.

Appendix A: A law of large numbers for functions of non-overlapping local volatility estimates

In this section, we state and prove Theorem A.1, a result that is auxiliary in proving Lemma 3.2. As noted in the main test, Theorem A.1 has merit on its own right as it extends Theorem 9.4.1 of Jacod and Protter (2012) to the case of smooth functions of consecutive local realized volatility estimates rather than a single local estimate. Contrary to Jacod and Protter (2012) who allow for the possibility that the local estimates entering the sum are based on overlapping intervals, here we focus our attention in the non-overlapping case. This is enough to cover the blocked multipower variation measures of Lemma 3.2.

Let

$$\hat{c}_{j,n} = \frac{n}{k_n} \sum_{m=1}^{k_n} r^2_{(j-1)k_n+m} 1_{\{|r_{(j-1)k_n+m}| \leq u_n\}},$$

$j = 1, \ldots, \frac{n}{k_n}$, with $r_i \equiv X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$, $i = 1, \ldots, n$; $\quad u_n = \alpha n^{-\varpi}$, $\quad \varpi \in (0, \frac{1}{2})$ and k_n is a sequence of integers satisfying $k_n \to \infty$ and $\frac{k_n}{n} \to 0$ as $n \to \infty$.

Theorem A.1 *Assume that X satisfies Assumption H-2, and let g be a continuous function on $\mathbb{R}^\ell_+$ satisfying for some $p \geq 0$:*

$$|g(x_1, \ldots, x_\ell)| \leq K\left(1 + |x_1|^p + \cdots + |x_\ell|^p\right).$$

If either one of the following conditions holds:

(a) X is continuous;

(b) $p < 1$;

(c) Assumption H-r holds for some $r \in [0,2)$ and $p \geq 1$, $\varpi \geq \frac{2p-1}{4p-r}$; then

$$G_n \equiv \frac{k_n}{n} \sum_{j=\ell}^{n/k_n} g\left(\hat{c}_{j,n}, \hat{c}_{j-1,n}, \ldots, \hat{c}_{j-\ell+1,n}\right) \xrightarrow{P} \int_0^1 g(\sigma_s^2, \ldots, \sigma_s^2) ds.$$

In the proof, we will follow the standard localization argument of Jacod and Protter (2012) and assume without loss of generality that the following stronger version of Assumption H-r holds:

Assumption SH-r Assumption H-r holds, and in addition the processes a and σ are bounded, and $|\delta(\omega, t, x)| \wedge 1 \leq \gamma(x)$ with $\int |\gamma(x)|^r dx < \infty$.

Proof. The proof follows the lines of the proof of Theorem 9.4.1 of Jacod and Protter (2012). By localization, we assume without loss of generality that SH-r holds throughout the proof with $r = 2$ for $p < 1$. Also, upon proving separately the convergence for g^+ and g^-, we assume that $g \geq 0$. Throughout the proof, K is a generic constant.

Step 1: We first assume that g is bounded. For all $s \in [0,1]$ and $l = 1, \ldots, \ell$, let $\hat{c}_n^{(l)}(s) = \hat{c}_{j+l,n}$ when $(j-1)\frac{k_n}{n} \leq s < j\frac{k_n}{n}$, where we set $\hat{c}_{j,n}$ to 0 if $j > n/k_n$. We have

$$G_n = \frac{k_n}{n} g\left(\hat{c}_{\ell,n}, \hat{c}_{\ell-1,n}, \ldots, \hat{c}_{1,n}\right) + \int_0^{1-\ell\frac{k_n}{n}} g\left(\hat{c}_n^{(\ell)}(s), \ldots, \hat{c}_n^{(1)}(s)\right) ds.$$

Thus

$$E\left(\left|G_n - \int_0^1 g\left(\sigma_s^2, \ldots, \sigma_s^2\right) ds\right|\right) \leq K\frac{k_n}{n} + \int_0^{1-\ell\frac{k_n}{n}} a_n(s) ds, \tag{A.1}$$

with $a_n(s) = E\left|g\left(\hat{c}_n^{(\ell)}(s), \ldots, \hat{c}_n^{(1)}(s)\right) - g\left(\sigma_s^2, \ldots, \sigma_s^2\right)\right|$. By the right continuity assumption for σ_s^2, the proof of Theorem 9.3.2(a) of Jacod and Protter (2012) readily applies to the functions $\hat{c}_n^{(l)}(s)$: $l = 1, \ldots, \ell$ and we can claim that $\hat{c}_n^{(l)}(s) \xrightarrow{P} \sigma_s^2$ for all $s \in [0,1)$. Hence, $g\left(\hat{c}_n^{(\ell)}(s), \ldots, \hat{c}_n^{(1)}(s)\right) - g\left(\sigma_s^2, \ldots, \sigma_s^2\right)$ converges in probability to 0. Since g is bounded, the bounded convergence theorem guarantees that $a_n(s)$ tends to 0 as $n \to \infty$ and stays bounded from the boundedness of g. We can therefore claim that the right hand side of (A.1) tends to 0 by the dominated convergence theorem and the conclusion of the theorem follows.

Step 2: Let ψ be a C^∞ function : $\mathbb{R}_+ \to [0,1]$ with $1_{[0,\infty)}(x) \leq \psi(x) \leq 1_{[\frac{1}{2},\infty)}(x)$, and $\psi_\varepsilon(x) = \psi(|x|/\varepsilon)$ and $\psi'_\varepsilon = 1 - \psi_\varepsilon$. For $m \geq 2$, let

$$g'_m(x_1, \ldots, x_\ell) = g(x_1, \ldots, x_\ell) \prod_{l=1}^{\ell} \psi'_m(x_l)$$

and $g_m = g - g'_m$. The function g'_m is continuous and bounded and hence Step 1 allows us to claim that for m fixed,

$$\frac{k_n}{n} \sum_{j=\ell}^{n/k_n} g'_m\left(\hat{c}_{j,n}, \hat{c}_{j-1,n}, \ldots, \hat{c}_{j-\ell+1,n}\right) \xrightarrow{P} \int_0^1 g'_m(\sigma_s^2, \ldots, \sigma_s^2) ds.$$

Note also that $\int_0^1 g'_m(\sigma_s^2, \ldots, \sigma_s^2) ds = \int_0^1 g(\sigma_s^2, \ldots, \sigma_s^2) ds$ for m large enough since σ_s^2 is bounded under SH-r and the fact that $\psi'_m(x) = 1$ for $|x| \leq m/2$.

It remains to show that $\frac{k_n}{n}\sum_{j=\ell}^{n/k_n} g_m\left(\hat{c}_{j,n},\hat{c}_{j-1,n},\ldots,\hat{c}_{j-\ell+1,n}\right)$ is negligible for large n and m. By assumption, we have

$$g_m(x_1,\ldots,x_\ell)\le K\left(1+\sum_{l=1}^{\ell}|x_l|^p\right)\left(1-\prod_{l=1}^{\ell}\psi'_m(x_l)\right)$$

but

$$1-\prod_{l=1}^{\ell}\psi'_m(x_l)\le\sum_{l=1}^{\ell}1_{\{|x_l|\ge\frac{m}{2}\}},$$

since if $|x_l|\le m/2$, $\psi'_m(x_l)=1$ and both sides of the inequality are nil if $|x_l|\le m/2$ for all l. If $|x_l|>m/2$ for some l, then $\sum_{l=1}^{\ell}1_{\{|x_l|\ge\frac{m}{2}\}}\ge 1\ge 1-\prod_{l=1}^{\ell}\psi'_m(x_l)$. Also, if $|x_l|>m/2$ for some l, we have that $1+\sum_{l=1}^{\ell}|x_l|^p\le 2\sum_{l=1}^{\ell}|x_l|^p$. Thus,

$$g_m(x_1,\ldots,x_\ell)\le 2K\sum_{l,l'=1}^{\ell}|x_l|^p 1_{\{|x_{l'}|\ge\frac{m}{2}\}}.$$

Therefore, to complete the proof, it suffices to show that

$$\lim_{m\to\infty}\limsup_{n\to\infty}E\left(\frac{k_n}{n}\sum_{j=\ell}^{n/k_n}\hat{c}^p_{j-l+1,n}1_{\{\hat{c}_{j-l'+1,n}>m\}}\right)=0, \tag{A.2}$$

for all $l,l'=1,\ldots,\ell$.

Letting $\kappa=0$ when X is continuous and $\kappa=1$ otherwise, for $q\ge 2$, following the argument in the proof of Jacod and Protter (2012), we have, for $i=1,\ldots,n$,

$$E(|r_i|^q)\le K_q\left(\frac{1}{n^{q/2}}+\kappa\frac{1}{n^{(q/2)\wedge 1}}\right)$$

and, from the c_r-inequality, we deduce that

$$E\left(\hat{c}^p_{j,n}\right)\le K_q\left(1+\kappa\frac{1}{n^{q\wedge 1-q}}\right).$$

Now, by successive application of the Hölder and Markov inequalities, we have for any $q>p$:

$$\begin{aligned}E\left(\hat{c}^p_{i,n}1_{\{\hat{c}_{j,n}>m\}}\right) &\le \left(E\left(\hat{c}^q_{i,n}\right)\right)^{\frac{p}{q}}\left(P\left(\hat{c}_{j,n}\ge m\right)\right)^{1-\frac{p}{q}}\\ &\le \left(E\left(\hat{c}^q_{i,n}\right)\right)^{\frac{p}{q}}\left(\frac{1}{m^q}E\left(\hat{c}^q_{j,n}\right)\right)^{1-\frac{p}{q}}\le\frac{K_q}{m^{q-p}}\left(1+\kappa\frac{1}{n^{q\wedge 1-q}}\right).\end{aligned}$$

Take $q=2p$ if X is continuous and $q=1>p$ otherwise and conclude (A.2).

We consider now $p\ge 1$. With the same alternative decomposition of X as that in Jacod and Protter (2012, Eq. (9.2.7)), we write $r_i=r_{1i}+r_{2i}$, with r_{1i} and r_{2i} the increments of the process X' and X'', respectively and as defined in the reference. We have that

$$\begin{aligned}r_i^2 1_{\{|r_i|\le u_n\}} &= (r_{1i}+r_{2i})^2 1_{\{|r_{1i}+r_{2i}|\le u_n\}}\le 2\left(r_{1i}^2+(r_{2i}^2\wedge u_n^2)\right)=2\left(r_{1i}^2+u_n^2\left(\frac{r_{2i}^2}{u_n^2}\wedge 1\right)\right)\\ &\le K\left(r_{1i}^2+u_n^2\left(n^{\varpi}|r_{2i}|\wedge 1\right)^2\right).\end{aligned}$$

(Where we use for the last inequality the fact that $(a/b)\wedge 1 \le \max(1,1/b)[a\wedge 1]$, with $a,b>0$.) Thus $\hat{c}_{j,n} \le \zeta'_{j,n} + \zeta''_{j,n}$ with

$$\zeta'_{j,n} = K\frac{1}{k_n}\sum_{m=1}^{k_n}(\sqrt{n}r_{1,(j-1)k_n+m})^2, \qquad \zeta''_{j,n} = K\frac{v_n^2}{k_n}\sum_{m=1}^{k_n}\left(n^{\varpi}|r_{2,(j-1)k_n+m}|\wedge 1\right)^2,$$

with $v_n = \sqrt{n}u_n$.

Note that

$$\hat{c}_{i,n}1_{\{\hat{c}_{j,n}\ge m\}} \le \frac{\hat{c}_{i,n}\hat{c}_{j,n}}{m} \le \frac{1}{2m}\left(\hat{c}_{i,n}^2+\hat{c}_{j,n}^2\right) \le \frac{1}{m}\left(\zeta'^2_{i,n}+\zeta''^2_{i,n}+\zeta'^2_{j,n}+\zeta''^2_{j,n}\right)$$

and by the c_r-inequality,

$$E\left(\hat{c}_{i,n}^p 1_{\{\hat{c}_{j,n}\ge m\}}\right) \le \frac{4^{p-1}}{m^p}\left(E(\zeta'^{2p}_{i,n})+E(\zeta''^{2p}_{i,n})+E(\zeta'^{2p}_{j,n})+E(\zeta''^{2p}_{j,n})\right).$$

On the other hand, Eqs. (9.2.12) and (9.2.13) of Jacod and Protter (2012) ensure that

$$E\left((\sqrt{n}|r_{1i}|)^q|\mathcal{F}_{\frac{i-1}{n}}\right) \le K_q \quad \text{and} \quad E\left((n^{\varpi}|r_{2i}|)^2\wedge 1|\mathcal{F}_{\frac{i-1}{n}}\right) \le Kn^{-1+r\varpi}\phi_n,$$

with $\phi_n \to 0$ as $n\to\infty$. Thus, a further application of the c_r-inequality gives $E(\zeta'^{2p}_{j,n}) < K$ and

$$\begin{aligned} E(\zeta''^{2p}_{j,n}) &\le K\frac{v_n^{4p}}{k_n}\sum_{m=1}^{k_n}E\left(\left(n^{\varpi}|r_{2,(j-1)k_n+m}|\wedge 1\right)^{4p}\right) \le K\frac{v_n^{4p}}{k_n}\sum_{m=1}^{k_n}E\left(\left(n^{\varpi}|r_{2,(j-1)k_n+m}|\wedge 1\right)^2\right) \\ &\le Kn^{4p(-\varpi+\frac{1}{2})}n^{-1+r\varpi}\phi_n = Kn^{-w}\phi_n, \end{aligned}$$

with $w = 1-2p+\varpi(4p-r)$. Thus,

$$E\left(\hat{c}_{i,n}^p 1_{\{\hat{c}_{j,n}\ge m\}}\right) \le \frac{K}{m^p}\left(1+n^{-w}\phi_n\right).$$

Since $w\ge 0$ under the maintained assumptions, (A.2) follows. ■

References

[1] Aït-Sahalia, Y., and J. Jacod, 2009. "Testing for jumps in a discretely observed process," *Annals of Statistics* 37 (1): 184–222.

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

[3] Aït-Sahalia, Y. and J. Jacod, 2014. High Frequency Financial Econometrics, Princeton University Press.

[4] Aït-Sahalia, Y., J. Jacod and J. Li, 2012. "Testing for jumps in noisy high frequency data," *Journal of Econometrics*, 168, 207-222.

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

[6] Andersen, T.G., T. Bollerslev, and F. X. Diebold, 2007. "Roughing it up: including jump components in the measurement, modeling, and forecasting of return volatility," *Review of Economics and Statistics* 89 (4), 701–20.

[7] Andersen, T.G., D. Dobrev and E. Schaumburg, 2012. "Jump-robust volatility estimation using nearest neighbor truncation," *Journal of Econometrics*, 169, 75-93.

[8] Bakshi, G., Cao, C., Chen, Z., 1997. "Empirical performance of alternative option pricing models," *Journal of Finance* 52, 2003–2049.

[9] Barndorff-Nielsen, O., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard, 2006. "A central limit theorem for realised power and bipower variations of continuous semimartingales." In Y. Kabanov, R. Lipster, and J. Stoyanov (Eds.), From Stochastic Analysis to Mathematical Finance, Festschrift for Albert Shiryaev, 33 - 68. Springer.

[10] Barndorff-Nielsen, O.E., Hansen, P. R., Lunde, A. and N. Shephard, 2009. "Realized kernels in practice: trades and quotes," *Econometrics Journal* 12, 1–32.

[11] Barndorff-Nielsen, O.E., and N. Shephard, 2004. "Power and bipower variation with stochastic volatility and jumps," *Journal of Financial Econometrics* 2 (1), 1–37.

[12] Barndorff-Nielsen, O.E., and N. Shephard, 2006. "Econometrics of testing for jumps in financial economics using bipower variation," *Journal of Financial Econometrics* 4, 1–30.

[13] Barndorff-Nielsen, O. E., Shephard, N. and M. Winkel, 2006. "Limit theorems for multipower variation in the presence of jumps," *Stochastic Processes and Applications*, 116, 796-806.

[14] Bates, D.S., 1996. "Jumps and stochastic volatility: exchange rate processes implicit in deutsch mark options," *Review of Financial Studies* 9, 69–107.

[15] Bollerslev, T., Law, T. H. and G. Tauchen, 2008. "Risk, jumps, and diversification," *Journal of Econometrics* 144, 234–256.

[16] Boudt, K., C. Croux and S. Laurent, 2011. "Robust estimation of intraweek periodicity in volatility and jump detection," *Journal of Empirical Finance*, 18, 353-367.

[17] Cont, R., and C. Mancini, 2011. "Nonparametric tests for pathwise properties of semimartingales," *Bernoulli* 17 (2): 781–813.

[18] Davidson, R. and J.G. MacKinnon, 1999. "The size distortion of bootstrap tests," *Econometric Theory*, 15, 1999, 361-376.

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

[20] Hall, P., 1992. "The bootstrap and Edgeworth expansion,"Springer-Verlag, New York.

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

[22] Huang, X., and G. Tauchen, 2005. "The relative contribution of jumps to total price variance," *Journal of Financial Econometrics* 3 (4), 456–99.

[23] Jiang, G.J., and R. C. A. Oomen, 2008. "Testing for jumps when asset prices are observed with noise - a 'swap variance' approach," *Journal of Econometrics* 144 (2), 352–70.

[24] Jacod, J., and P. Protter, 2012. "Discretization of processes," Springer-Verlag, Berlin Heidelberg.

[25] Jacod, J., and M. Rosenbaum, 2013. "Quarticity and other functionals of volatility efficient estimation," *The Annals of Statistics*, 41, 1462-1484.

[26] Li, J., V. Todorov and G. Tauchen, 2016. "Adaptive estimation of continuous-time regression models using high-frequency data," *Working paper, Duke University*

[27] Johannes, M., 2004. "The statistical and economic role of jumps in interest rates," *Journal of Finance* 59, 227–260.

[28] Lee, S., and P.A. Mykland, 2008. "Jumps in financial markets: a new nonparametric test and jump dynamics," *Review of Financial Studies* 21 (6): 2535–63.

[29] Lee, S., and P.A. Mykland, 2012. "Jumps in equilibrium prices and market microstructure noise," *Journal of Econometrics*, 168, 396-406.

[30] MacKinnon, J.G., 2009. "Bootstrap hypothesis testing," in Handbook of Computational Econometrics, ed. David A. Belsley and John Kontoghiorghes, Chichester, Wiley, 183–213.

[31] Mancini, C., 2001. "Disentangling the jumps of the diffusion in a geometric jumping Brownian Motion." *Giornale dell'Istituto Italiano Attuari* 64, 19–47.

[32] Mykland, P.A., Zhang, L., 2009. "Inference for continuous semimartingales observed at high frequency," *Econometrica* 77, 1403–1445.

[33] Mykland, P.A., Shephard, N. and K. Sheppard, 2012. "Econometric analysis of financial jumps using efficient bi- and multipower variation.", mimeo, Oxford University.

[34] Podolskij, M., and D. Ziggel 2010. "New tests for jumps in semimartingale models," *Statistical Inference for Stochastic Processes* 13(1), 15-41.