Bootstrapping pre-averaged realized volatility under market microstructure noise *

Ulrich Hounyo[†], Sílvia Gonçalves[‡] and Nour Meddahi[§]

Aarhus University, Western University and Toulouse School of Economics

October 3, 2015

Abstract

The main contribution of this paper is to propose a bootstrap method for inference on integrated volatility based on the pre-averaging approach, where the pre-averaging is done over all possible overlapping blocks of consecutive observations. The overlapping nature of the pre-averaged returns implies that the leading martingale part in the pre-averaged returns are k_n-dependent with k_n growing slowly with the sample size n. This motivates the application of a blockwise bootstrap method. We show that the "blocks of blocks" bootstrap method is not valid when volatility is time-varying. The failure of the blocks of blocks bootstrap is due to the heterogeneity of the squared pre-averaged returns when volatility is stochastic. To preserve both the dependence and the heterogeneity of squared pre-averaged returns, we propose a novel procedure that combines the wild bootstrap with the blocks of blocks bootstrap. We provide a proof of the first order asymptotic validity of this method for percentile and percentile-t intervals. Our Monte Carlo simulations show that the wild blocks of blocks bootstrap improves the finite sample properties of the existing first order asymptotic theory. We use empirical work to illustrate its use in practice.

Keywords: Block bootstrap, high frequency data, market microstructure noise, pre-averaging, realized volatility, wild bootstrap.

---

*We would like to thank Ilze Kalnina, Kevin Sheppard and Neil Shephard for many useful comments and discussions. This work was supported by grants FQRSC-ANR and SSHRC. 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 benefited from the financial support of the chair "Marché des risques et création de valeur" Fondation du risque/SCOR.

[†]Department of Economics and Business Economics, Aarhus University, 8210 Aarhus V., Denmark. Email: uhounyo@econ.au.dk.

[‡]Department of Economics, Faculty of Social Science, Western University, 1151 Richmond Street N., London, Ontario, Canada, N6A 5C2. Tel: (519) 661-2111. Ext: 85232. Email: sgoncal9@uwo.ca.

[§]Toulouse School of Economics, 21 allée de Brienne -Manufacture des Tabacs-31000, Toulouse, France. Email: nour.meddahi@tse-fr.eu.

1 Introduction

Estimation of integrated volatility is complicated by the existence of market microstructure noise. This noise represents the discrepancy between the true efficient price of an asset and its observed counterpart and is caused by a multitude of market microstructure effects (such as bid-ask bounds, the discreteness of price changes and the existence of rounding errors, the gradual response of prices to a block trade, the existence of data recording errors such as prices entered as zero, misplaced decimal points, etc).

In frictionless markets, and when the log-price process follows a continuous semimartingale, realized volatility computed as the sum of squared intraday returns converges to the integrated volatility as the sampling frequency goes to infinity (see e.g. Andersen, Bollerslev, Diebold, and Labys (2001), Barndorff-Nielsen and Shephard (2002)). See also related work discussed in Jacod and Protter (1998) and Barndorff-Nielsen and Shephard (2001). However, realized volatility is no longer consistent for integrated volatility under the presence of market microstructure noise. This has motivated the development of alternative estimators. One popular method is the pre-averaging approach first introduced by Podolskij and Vetter (2009) and further studied by Jacod et al. (2009). The basic underlying idea consists of first averaging out the noise by computing pre-averaged returns and then computing a realized volatility-like estimator using the pre-averaged returns. Although the pre-averaged realized volatility estimator is consistent for integrated volatility, its convergence rate is much slower than that of realized volatility (when there is no noise) and this can result in finite sample distortions that persist even at very large sample sizes. For this reason, the bootstrap is a useful alternative method of inference in this context.

In this paper, we propose a bootstrap method that can be used to estimate the distribution and the variance of the pre-averaged realized volatility estimator of Jacod et al. (2009). Our proposal is to resample the pre-averaged returns instead of resampling the original noisy returns. To be valid, the bootstrap needs to mimic the dependence and heterogeneity properties of the (squared) pre-averaged returns. When pre-averaging occurs over overlapping blocks of returns, as in Jacod et al. (2009), the leading martingale part in the squared pre-averaged returns are k_n-dependent, where k_n denotes the block length of the interval over which the pre-averaging is done and n denotes the sample size. Since k_n is proportional to $\sqrt{n}$, $k_n \to \infty$ as $n \to \infty$, which implies that the pre-averaged returns are strongly dependent. This suggests that a block bootstrap applied to the pre-averaged returns is appropriate and its application amounts to a "blocks of blocks" bootstrap, as proposed by Politis and Romano (1992) and further studied by Bühlmann and Künsch (1995) (see also Künsch (1989)). Nevertheless, as we show here, such a bootstrap scheme is not valid when volatility is time-varying. The reason is that squared pre-averaged returns are heterogenously distributed (in particular, their mean and variance are time-varying) and this creates a bias term in the blocks of blocks bootstrap variance estimator when volatility is stochastic. Thus, to handle both the dependence and heterogeneity of the squared pre-averaged returns, we propose a novel bootstrap approach that combines the wild bootstrap with the blocks of blocks bootstrap. We name this novel approach the wild blocks of blocks bootstrap. One of our main contributions is to show that this method consistently estimates the variance and the entire distribution of the pre-averaged estimator of Jacod et al. (2009). We provide a proof of the first order asymptotic validity of this method for constructing bootstrap unstudentized (percentile) as well as bootstrap studentized (percentile-t) intervals.

The pre-averaging approach can also be implemented with non-overlapping intervals, as in Podolskij and Vetter (2009). However, the overlapping methods is expected to provide more precise estimates of the integrated variance. We provide intuition of this in Section 2.2.

Gonçalves, Hounyo and Meddahi (2014) study the consistency of the wild bootstrap for the non-overlapping estimator of Podolskij and Vetter (2009). The wild bootstrap exploits the asymptotic independence of the pre-averaged returns when these are computed over non-overlapping intervals. This method is no longer valid when overlapping intervals are used to compute pre-averaged returns since these are strongly dependent. For this reason, a new bootstrap method is needed for the Jacod et al.'s (2009) approach. Although the wild blocks of blocks bootstrap that we propose here requires the choice of an additional tuning parameter (the block size), we suggest an empirical procedure to select the block size that performs well in our simulations.

Other estimators of integrated volatility that are consistent under market microstructure noise include the subsampling approach of Zhang et al. (2005) (see also the multiscale realized volatility estimator of Zhang (2006)) and the realized kernel estimator of Barndorff-Nielsen et al. (2008) (the maximum likelihood-based estimator of Xiu (2010) is also a recent addition to this literature). The bootstrap could also be useful for inference in the context of these estimators. Indeed, Zhang et al. (2011) showed that the asymptotic normal approximation is often inaccurate for the subsampling realized volatility estimator,[1] whose finite sample distribution is skewed and heavy tailed. They proposed Edgeworth corrections for this estimator as a way to improve upon the standard normal approximation. Unfortunately, Zhang et al. (2011) provided the Edgeworth corrections of the normalized statistic (where the denominator equals the variance of the estimator in population) rather than studentized statistic (where the denominator is a consistent estimator of the estimator's variance), while Gonçalves and Meddahi (2008) proved that Edgeworth corrections based on normalized statistic is worse than the asymptotic theory when there is no noise.

The main reason why we focus on the pre-averaging approach here is that it naturally lends

[1]Similarly, Bandi and Russell (2011) discussed the limitations of asymptotic approximations in the context of realized kernels and proposed an alternative solution.

itself to the bootstrap. In particular, we resample the pre-averaged returns instead of the individual returns and exploit the dependence and heterogeneity properties of the pre-averaged returns to prove the consistency of the bootstrap.

The rest of this paper is organized as follows. In the next section, we first introduce the setup, our assumptions and review the existing asymptotic theory of Jacod et al. (2009). Section 3 contains the bootstrap results. In Section 3.1 we show that the blocks of blocks bootstrap is consistent only when volatility is constant whereas Section 3.2 describes the wild blocks of blocks bootstrap and shows its consistency under stochastic volatility and i.i.d. noise. Section 4 presents the simulation results whereas Section 5 contains an empirical application. Section 6 concludes. Two appendices are provided. Appendix A contains the tables with simulation results whereas Appendix B is a mathematical appendix with the proofs.

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

2 Setup, assumptions and review of existing results

2.1 Setup and assumptions

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

$$X_t = X_0 + \int_0^t a_s ds + \int_0^t \sigma_s dW_s,\ t \geq 0, \tag{1}$$

where $a = (a_t)_{t\geq 0}$ is an adapted càdlàg drift process, $\sigma = (\sigma_t)_{t\geq 0}$ is an adapted càdlàg volatility process and $W = (W_t)_{t\geq 0}$ a standard Brownian motion.

The object of interest is the quadratic variation of X, i.e. the process

$$C_t = \int_0^t \sigma_s^2 ds,$$

also known as the integrated volatility. Without loss of generality, we let $t = 1$ and define $C_1 = \int_0^1 \sigma_s^2 ds$ as the integrated volatility of X over a given time interval $[0, 1]$, which we think of as a given day.

The presence of market frictions such as price discreteness, rounding errors, bid-ask spreads, gradual response of prices to block trades, etc, prevent us from observing the true efficient price process X. Instead, we observe a noisy price process Y, observed at time points $t = \frac{i}{n}$ for $i = 0, \ldots, n$, given by

$$Y_t = X_t + \epsilon_t, \tag{2}$$

where ϵ_t represents the noise term that collects all the market microstructure effects.

In order to make both X and Y measurable with respect to the filtration, we define a new probability space $\left(\Omega, (\mathcal{F}_t)_{t\geq 0}, P\right)$, which accommodates both processes. To this end, we follow Jacod et al. (2009) and assume one has a second space $\left(\Omega^1, (\mathcal{F}_t^1)_{t\geq 0}, P^1\right)$, where Ω^1 denotes $\mathbb{R}^{[0,1]}$ and $\mathcal{F}^1$ the product Borel-σ-field on Ω^1. Next, for any $t \in [0,1]$, we define $Q_t\left(\omega^{(0)}, dy\right)$ to be the probability measure on $\mathbb{R}$, which corresponds to the transition from $X_t\left(\omega^{(0)}\right)$ to the observed process Y_t. In the case of i.i.d. noise, this transition kernel is rather simple (see e.g. equation (2.7) of Vetter (2008)), but it becomes more pronounced in a general framework. $P^1\left(\omega^{(0)}, d\omega^{(1)}\right)$ denotes the product measure $\otimes_{t\in[0,1]} Q_t\left(\omega^{(0)}, \cdot\right)$. The filtered probability space $\left(\Omega, (\mathcal{F}_t)_{t\in[0,1]}, P\right)$ on which the process Y lives is then defined with $\Omega = \Omega^0 \times \Omega^1$, $\mathcal{F} = \mathcal{F}^0 \times \mathcal{F}^1$, $\mathcal{F}_t = \bigcap_{s>t} \mathcal{F}_s^0 \times \mathcal{F}_s^1$, and $P\left(d\omega^{(0)}, d\omega^{(1)}\right) = P^0\left(\omega^{(0)}\right) P^1\left(\omega^{(0)}, d\omega^{(1)}\right)$.

We assume that ϵ_t is centered and independent, conditionally on the efficient price process X. In addition, we assume that the conditional variance of ϵ_t is càdlàg. Assumption 1 below collects these assumptions.

Assumption 1.

(i) $E\left(\epsilon_t|X\right) = 0$ and ϵ_t and ϵ_s are independent for all $t \neq s$, conditionally on X.

(ii) $\alpha_t = E\left(\epsilon_t^2|X\right)$ is càdlàg and $E\left(\epsilon_t^8\right) < \infty$.

Assumption 1 amounts to Assumption (K) in Jacod et al. (2009). As they explain, this assumption is rather general, allowing for time varying variances of the noise and dependence between X and ϵ. See Jacod et al. (2009) for particular examples of market microstructure noise that satisfy Assumption 1. However, empirically the conditional independence assumption on ϵ may be unrealistic especially at the highest frequencies (see e.g. Hansen and Lunde (2006)). We will investigate the impact of autocorrelated noise on the bootstrap performance in Section 4.

2.2 The pre-averaged estimator and its asymptotic theory

We observe Y at regular time points $\frac{i}{n}$, for $i = 0, \ldots, n$, from which we compute n intraday returns at frequency $\frac{1}{n}$,

$$r_i \equiv Y_{\frac{i}{n}} - Y_{\frac{i-1}{n}}, \quad i = 1, \ldots, n.$$

Given that $Y = X + \epsilon$, we can write

$$r_i = \left(X_{\frac{i}{n}} - X_{\frac{i-1}{n}}\right) + \left(\epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}\right) \equiv r_i^e + \Delta\epsilon_i,$$

where $r_i^e = X_{\frac{i}{n}} - X_{\frac{i-1}{n}}$ denotes the $\frac{1}{n}$-frequency return on the efficient price process. Under Assumption 1, the order of magnitude of $\Delta\epsilon_i \equiv \epsilon_{\frac{i}{n}} - \epsilon_{\frac{i-1}{n}}$ is $O_P(1)$. In contrast, the ex-post variation of r_i^e is given by $\int_{(i-1)/n}^{i/n} \sigma_s^2 ds$. The order of magnitude of r_i^e is then $O_P\left(n^{-1/2}\right)$. This decomposition shows that the noise completely dominates the observed return process as $n \to \infty$, implying that the usual realized volatility estimator is biased and inconsistent. See Zhang et al. (2005) and Bandi and Russell (2008).

To describe the Jacod et al. (2009) pre-averaging approach, let k_n be a sequence of integers which will denote the window length over which the pre-averaging of returns is done. Similarly, let g be a weighting function on $[0,1]$ such that $g(0) = g(1) = 0$ and $\int_0^1 g(s)^2 ds > 0$, and assume g is continuous and piecewise continuously differentiable with a piecewise Lipschitz derivative g'. An example of a function that satisfies these restrictions is $g(x) = \min(x, 1-x)$.

We introduce the following additional notation. Let

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

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

For $i = 0, \ldots, n-k_n+1$, the pre-averaged returns $\bar{Y}_i$ are obtained by computing the weighted sum of all consecutive $\frac{1}{n}$-horizon returns over each block of size k_n,

$$\bar{Y}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right) r_{i+j}.$$

The effect of pre-averaging is to reduce the impact of the noise in the pre-averaged return. Specifically, as shown by Vetter (2008),

$$\bar{X}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(X_{\frac{i+j}{n}} - X_{\frac{i+j-1}{n}}\right) = O_P\left(\sqrt{\frac{k_n}{n}}\right),$$

and

$$\bar{\epsilon}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(\epsilon_{\frac{i+j}{n}} - \epsilon_{\frac{i+j-1}{n}}\right) = O_P\left(\frac{1}{\sqrt{k_n}}\right).$$

Thus, the impact of the noise is reduced the larger k_n is. To get the efficient $n^{-1/4}$ rate of convergence, Jacod et al. (2009) propose to choose a sequence of integers k_n such that the following assumption holds.

Assumption 2. For $\theta \in (0, \infty)$, we have that

$$\frac{k_n}{\sqrt{n}} = \theta + o\left(n^{-1/4}\right). \tag{3}$$

This choice implies that the orders of the two terms ($\bar{X}_i$ and $\bar{\epsilon}_i$) are balanced and equal to $O_P\left(n^{-1/4}\right)$. An example that satisfies (3) is $k_n = [\theta\sqrt{n}]$.

Based on the pre-averaged returns $\bar{Y}_i$, Jacod et al. (2009) propose the following estimator of integrated volatility,

$$PRV_n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^2 - \frac{\psi_1}{2n\theta^2\psi_2} \sum_{i=1}^{n} r_i^2, \tag{4}$$

where ψ_1 and ψ_2 are as defined above.

The first term in (4) is an average of realized volatility-like estimators based on pre-averaged returns of length k_n whereas the second term is a bias correction term. As discussed in Jacod et al. (2009), this bias term does not contribute to the asymptotic variance of PRV_n.

In order to give the central limit theorem for PRV_n, we introduce the following numbers that are associated with g,

$$\Phi_{ij} = \int_0^1 \phi_i(s)\,\phi_j(s)\,ds, \quad \text{and} \quad \Psi_{ij} = -\int_0^1 s\phi_i(s)\,\phi_j(s)\,ds.$$

For the simple function $g(x) = \min(x, 1-x)$, $\Phi_{11} = 1/6$, $\Phi_{12} = 1/96$ and $\Phi_{22} = 151/80640$. Under Assumption 1 and (k_n, θ) satisfying (3), Jacod et al. (2009) show that as $n \to \infty$,

$$\frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{V}} \to^{st} N(0,1), \tag{5}$$

where $\to^{st}$ denotes stable convergence, and

$$V = \frac{4}{\psi_2^2} \int_0^1 \left(\Phi_{22}\theta\sigma_s^4 + 2\Phi_{12}\frac{\sigma_s^2\alpha_s}{\theta} + \Phi_{11}\frac{\alpha_s^2}{\theta^3}\right) ds \tag{6}$$

is the conditional variance of PRV_n. To estimate V consistently, Jacod et al. (2009) propose

$$\begin{aligned}\hat{V}_n &= \frac{4\Phi_{22}}{3\theta\psi_2^4} \sum_{i=0}^{n-k_n+1} \overline{Y}_i^4 + \frac{4}{n\theta^3}\left(\frac{\Phi_{12}}{\psi_2^3} - \frac{\Phi_{22}\psi_1}{\psi_2^4}\right) \sum_{i=0}^{n-2k_n+1} \overline{Y}_i^2 \sum_{j=i+k_n}^{i+2k_n-1} r_j^2 \\ &+ \frac{1}{n\theta^3}\left(\frac{\Phi_{11}}{\psi_2^2} - 2\frac{\Phi_{12}\psi_1}{\psi_2^3} + \frac{\Phi_{22}\psi_1^2}{\psi_2^4}\right) \sum_{i=0}^{n-2k_n+1} r_i^2 r_{i+2}^2. \end{aligned} \tag{7}$$

Together with the CLT result (5), we have that

$$T_n \equiv \frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{\hat{V}_n}} \to^{st} N(0,1).$$

We can use this feasible asymptotic distribution result to build confidence intervals for integrated volatility. In particular, a two-sided feasible $100(1-\alpha)\%$ level interval for $\int_0^1 \sigma_s^2 ds$ is given by:

$$IC^{a}_{Feas,1-\alpha} = \left(PRV_n - z_{1-\alpha/2} n^{-1/4} \sqrt{\hat{V}_n}, PRV_n + z_{1-\alpha/2} n^{-1/4} \sqrt{\hat{V}_n} \right), \tag{8}$$

where $z_{1-\alpha/2}$ is such that $\Phi\left(z_{1-\alpha/2}\right) = 1-\alpha/2$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution. For instance, $z_{0.975} = 1.96$ when $\alpha = 0.05$.

Note that the pre-averaging approach can also be implemented with non-overlapping intervals, as in Vetter (2008) and Podolskij and Vetter (2009). By Theorem 3.7 of Vetter (2008), we can also build confidence intervals for integrated volatility based on the non-overlapping pre-averaged realized volatility estimator. In particular, a two-sided feasible $100(1-\alpha)\%$ level interval for $\int_0^1 \sigma_s^2 ds$ based on non-overlapping intervals, is given by:

$$IC^{PV}_{Feas,1-\alpha} = \left(PRV_n^{PV} - z_{1-\alpha/2} n^{-1/4} \sqrt{\hat{V}_n^{PV}}, PRV_n^{PV} + z_{1-\alpha/2} n^{-1/4} \sqrt{\hat{V}_n^{PV}} \right),$$

where PRV_n^{PV} and $\hat{V}_n^{PV}$ are the non-overlapping pre-averaged realized volatility estimator and a consistent estimator of the asymptotic variance, respectively. Following Corollary 3.3 and 3.6 of Vetter (2008), we have that

$$PRV_n^{PV} = \frac{1}{\psi_2} \sum_{i=0}^{\left\lfloor \frac{n}{k_n} - 1 \right\rfloor} \bar{Y}_{ik_n}^2 - \frac{\psi_1}{2n\theta^2\psi_2} \sum_{i=1}^{n} r_i^2, \text{ and } \hat{V}_n^{PV} = \frac{2\sqrt{n}}{3\psi_2^2} \sum_{i=0}^{\left\lfloor \frac{n}{k_n} - 1 \right\rfloor} \bar{Y}_{ik_n}^4.$$

It is well known that the pre-averaged estimator based on overlapping blocks is more efficient than the pre-averaged estimator based on non-overlapping returns. In particular, for the parametric model of zero drift, constant volatility $\sigma > 0$ and constant conditional variance of ϵ_t, $\alpha_t = \alpha > 0$, if we choose for example the weight function $g(x) = \min(x, 1-x)$, some simple derivations shows that the "optimal" choice of θ which minimizes the asymptotic variance V given in (6) is approximately equal to $4.777\sqrt{\alpha}/\sigma$. Under this special setting, using $\theta = 4.777\sqrt{\alpha}/\sigma$, the non-overlapping pre-averaged estimator of Podolskij and Vetter (2009) has an asymptotic variance that is approximately equal to $20.11\sigma^3\sqrt{\alpha}$, whereas the overlapping estimator of Jacod et al. (2009) has a variance equal to $8.545\sigma^3\sqrt{\alpha}$ (this is very close to the lower bound $8\sigma^3\sqrt{\alpha}$ derived by Gloter and Jacod (2001)). Therefore, based on efficiency, we should choose the overlapping approach of Jacod et al. (2009) rather than the non-overlapping one. However, as shown by our simulation results (see Appendix A), we find that the pre-averaging approach leads to important coverage probability distortions when returns are not sampled too frequently. This motivates Gonçalves, Hounyo and Meddahi (2014) to propose the wild bootstrap as alternative method of inference for the Podolskij and Vetter's (2009) estimator. Here, we focus on bootstrapping the more efficient pre-averaged realized volatility estimator of

Jacod et al. (2009).

3 The bootstrap

The goal of this section is to propose a bootstrap method that can be used to consistently estimate the distribution of $n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)$ as well as for the studentized statistic $n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)/\sqrt{\hat{V}_n}$. This justifies the construction of bootstrap percentile and percentile-t confidence intervals for integrated volatility, respectively.

Gonçalves and Meddahi (2009) proposed bootstrap methods for realized volatility in the absence of market microstructure noise. In their ideal setting, intraday returns r_i (conditionally on the path of the volatility σ and the drift a) are uncorrelated, but possibly heteroskedastic due to stochastic volatility, thus motivating the use of a wild bootstrap method.

When intraday returns are contaminated by market microstructure noise, they are no longer conditionally uncorrelated, as in Gonçalves and Meddahi (2009). This implies that the wild bootstrap is no longer valid when applied to r_i. Instead, a block bootstrap method applied to the intraday returns would seem appropriate.

One complication arises in this context: the statistic of interest is not symmetric in the observations and the block bootstrap generates blocks of observations that are conditionally independent. In particular, since the first term in PRV_n is an average of the squared pre-averaged returns $\bar{Y}_i^2$, it depends on all the products of intraday returns inside blocks of size k_n. If we generate block bootstrap intraday returns, these will be independent between blocks, implying that the bootstrap statistic may look at many pairs of intraday returns that are independent in the bootstrap world. This not only renders the analysis very complicated but can induce biases in the bootstrap estimator. To avoid this problem when dealing with statistics that are not symmetric in the underlying observations, Künsch (1989), Politis and Romano (1992) and Bühlmann and Künsch (1995) studied the "blocks of blocks" bootstrap, where one applies the block bootstrap to appropriately pre-specified blocks of observations. In our context, the blocks of blocks bootstrap consists of applying a traditional block bootstrap to the squared pre-averaged returns $\bar{Y}_i^2$. As we will see next, this approach is not valid when volatility is time-varying. The reason is that when volatility is stochastic, squared pre-averaged returns are not only dependent but also heterogeneous. The block bootstrap does not capture this heterogeneity unless volatility is constant[2]. In order to capture both the time dependence and the heterogeneity in $\bar{Y}_i^2$, we propose a novel bootstrap procedure that combines the wild bootstrap with the block bootstrap.

Although the consistent estimator of integrated volatility is PRV_n, only the first term in

[2]See Gonçalves and White (2002) for a discussion of the impact of mean heterogeneity on the validity of the block bootstrap for the sample mean.

PRV_n drives the variance of the limiting distribution of PRV_n. In particular, as Jacod et al. (2009) have shown, the second term is a bias correction term which does not contribute to the asymptotic variance (it only ensures that the estimator is well centered at the integrated volatility). For this reason, our proposal is to bootstrap only the first contribution to PRV_n,

$$\widetilde{PRV}_n = \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^2.$$

This statistic depends only on the pre-averaged returns, to which we apply a particular bootstrap scheme. More specifically, let $\left\{\bar{Y}_i^* : i = 0, 1, \ldots, n - k_n + 1\right\}$ denote a bootstrap sample from $\left\{\bar{Y}_i : i = 0, 1, \ldots, n - k_n + 1\right\}$. The bootstrap analogue of PRV_n is

$$PRV_n^* = \widetilde{PRV}_n^* - \frac{\psi_1}{2n\theta^2\psi_2} \sum_{i=1}^{n} r_i^2,$$

where[3]

$$\widetilde{PRV}_n^* = \frac{1}{\psi_2 k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^{*2}.$$

Since the (conditional) expected value of $n^{1/4}\left(PRV_n^* - PRV_n\right)$ induced by the bootstrap resampling methods considered in this paper is not always zero, we center PRV_n^* around $E^*\left(PRV_n^*\right)$. Thus, we use the bootstrap distribution of

$$n^{1/4}\left(PRV_n^* - E^*\left(PRV_n^*\right)\right) = n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)$$

as an estimator of the distribution[4] of $n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)$.

Next, we consider the blocks of blocks bootstrap approach applied to $\widetilde{PRV}_n$ and show that it is asymptotically invalid when volatility is time-varying. This motivates a new bootstrap method that combines the wild bootstrap with the block bootstrap, which we study in the last subsection.

3.1 The blocks of blocks bootstrap

To describe this approach, let $N_n = n - k_n + 2$ denote the total number of pre-averaged returns and let b_n denote the block size. We suppose that $N_n = J_n \cdot b_n$, so that J_n denotes the number of blocks of size b_n one needs to draw to get $N_n = n - k_n + 2$ bootstrap observations. The blocks of

[3] This implies that our bootstrap statistic actually contains the bias term. Nevertheless, since this term is evaluated on the original sample rather than on the bootstrap data, our bootstrap method does not capture the added uncertainty caused by estimation of this term. Our simulations show that despite this, the bootstrap is very accurate, outperforming the asymptotic normal approximation.

[4] In particular, we can explicitly compute the bootstrap expectation of $\widetilde{PRV}_n^*$ (and we do so in (9) and (14)), for the blocks of blocks bootstrap and the wild blocks of blocks bootstrap, respectively. For instance, under the wild blocks of blocks bootstrap scheme, using an external random variable η with mean 1, it follows that $E^*\left(PRV_n^*\right) = PRV_n$.

blocks bootstrap generates a bootstrap resample $\left\{\bar{Y}^*_{i-1} : i = 1, \ldots, N_n\right\}$ by applying the moving blocks bootstrap of Künsch (1989) to the scaled pre-averaged returns $\left\{\bar{Y}_{i-1} : i = 1, \ldots, N_n\right\}$.

Letting $I_1, \ldots, I_{J_n}$ be i.i.d. random variables distributed uniformly on $\{0, 1, \ldots, N_n - b_n\}$, we set

$$\bar{Y}^*_{i-1+(j-1)b_n} = \bar{Y}_{i-1+I_j} \text{ for } 1 \leq j \leq J_n \text{ and } 1 \leq i \leq b_n.$$

The bootstrap analogue of $\widetilde{PRV}_n$ is

$$\widetilde{PRV}^*_n = \frac{1}{\psi_2 k_n} \sum_{i=1}^{N_n} \bar{Y}^{*2}_{i-1} = \frac{1}{J_n} \sum_{j=1}^{J_n} \left(\frac{1}{b_n} \sum_{i=1}^{b_n} \underbrace{\frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}^2_{I_j+i-1}}_{\equiv Z_{I_j+i}} \right),$$

where we let $Z_i \equiv \frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}^2_{i-1}$. Note that in our setup, $\bar{Y}_i = \bar{X}_i + \bar{\epsilon}_i = O_P\left(n^{-1/4}\right)$ given that k_n is such that $k_n/\sqrt{n} = \theta + o\left(n^{-1/4}\right)$. This implies that $\bar{Y}^2_{i-1} = O_P\left(n^{-1/2}\right)$ and therefore $Z_i = \frac{n-k_n+2}{k_n} \frac{1}{\psi_2} \bar{Y}^2_{i-1}$ is $O_P(1)$.

We can easily show that

$$E^*\left(\widetilde{PRV}^*_n\right) = \frac{1}{J_n} \sum_{j=1}^{J_n} E^*\left(\frac{1}{b_n} \sum_{i=1}^{b_n} Z_{I_j+i}\right) = \frac{1}{N_n - b_n + 1} \sum_{j=0}^{N_n - b_n} \left(\frac{1}{b_n} \sum_{i=1}^{b_n} Z_{j+i}\right). \tag{9}$$

Similarly,

$$\begin{aligned} V^*_n &\equiv Var^*\left(n^{1/4} \widetilde{PRV}^*_n\right) = \sqrt{n} E^*\left[\left(\frac{1}{J_n} \sum_{j=1}^{J_n} \frac{1}{b_n} \sum_{i=1}^{b_n} \left(Z_{I_j+i} - E^*\left(\widetilde{PRV}^*_n\right)\right)\right)^2\right] \\ &= \sqrt{n} \frac{1}{J_n} E^*\left(\frac{1}{b_n} \sum_{i=1}^{b_n} \left(Z_{I_1+i} - E^*\left(\widetilde{PRV}^*_n\right)\right)\right)^2 \\ &= \sqrt{n} \frac{b_n}{N_n} \frac{1}{N_n - b_n + 1} \sum_{j=0}^{N_n - b_n} \left(\frac{1}{b_n} \sum_{i=1}^{b_n} \left(Z_{j+i} - E^*\left(\widetilde{PRV}^*_n\right)\right)\right)^2. \end{aligned} \tag{10}$$

Our next result studies the convergence of V^*_n when $b_n = (p+1) k_n$, and $p \geq 1$ is either fixed as $n \to \infty$ or $p \to \infty$ after $n \to \infty$ (which we denote by writing $(n, p)_{seq} \to \infty$). To emphasize the dependence of V^*_n on p we write $V^*_{n,p}$.

Lemma 3.1 *Suppose Assumption 1 holds and $k_n \to \infty$ as $n \to \infty$ such that Assumption 2 holds. Let $V^*_{n,p} \equiv Var^*\left(n^{1/4} \widetilde{PRV}^*_n\right)$ denote the moving blocks bootstrap variance of $n^{1/4} \widetilde{PRV}^*_n$ based on a block length equal to $b_n = (p+1) k_n$, where $p \geq 1$. Then,*

a) *For any fixed $p \geq 1$, as $n \to \infty$,*

$$V^*_{n,p} \xrightarrow{P} V_p + B_p,$$

where

$$V_p = \int_0^1 \gamma^2 (p)_t \, dt$$

with

$$\gamma^2 (p)_t = \frac{4}{\psi_2^2} \left[\left(\Phi_{22} + \frac{1}{p+1} \Psi_{22} \right) \theta \sigma_t^4 + 2 \left(\Phi_{12} + \frac{1}{p+1} \Psi_{12} \right) \frac{\sigma_t^2 \alpha_t}{\theta} + \left(\Phi_{11} + \frac{1}{p+1} \Psi_{11} \right) \frac{\alpha_t^2}{\theta^3} \right],$$

and

$$B_p = \theta (p+1) \left[\int_0^1 \left(\sigma_t^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_t \right)^2 dt - \left(\int_0^1 \left(\sigma_t^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_t \right) dt \right)^2 \right].$$

b) When $\sigma_t = \sigma$ and $\alpha_t = \alpha$ are constants, $B_p = 0$ for any $p \geq 1$ and $V_p \xrightarrow{P} V \equiv \lim_{n\to\infty} Var\left(n^{1/4} PRV_n\right)$ as $p \to \infty$. In this case, $V_{n,p}^* \xrightarrow{P} V$ as $(n,p)_{seq} \to \infty$.

c) More generally, when σ_t and/or α_t are stochastic, $V_{n,p}^* \xrightarrow{P} \infty$ as $(n,p)_{seq} \to \infty$.

Part a) of Lemma 3.1 shows that when the bootstrap block size b_n is a fixed proportion of the pre-averaging block size k_n, the blocks of blocks bootstrap variance converges in probability to $V_p + B_p$, where B_p is a bias term due to the fact that volatility is time-varying. When both the volatility σ_t and α_t, the conditional variance of ϵ_t, are constants, B_p is equal to zero for any value of p. If $p \to \infty$ (i.e. if $b_n/k_n \to \infty$ as $n \to \infty$), then $V_p \xrightarrow{P} V$, the asymptotic variance of $n^{1/4} PRV_n$. Therefore, under these conditions, $V_{n,p}^* \xrightarrow{P} V$ as $(n,p) \to \infty$ sequentially. Although this result does not necessarily imply the consistency of $V_{n,p}^*$ towards V as $(n,p) \to \infty$ jointly (because sequential convergence does not by itself imply joint convergence), it is a first step in that direction (see in particular Lemma 6 of Phillips and Moon, 1999). We do not pursue the derivation of the joint limit of $V_{n,p}^*$ here because that would distract us from the main message of Lemma 3.1, which is the invalidity of the blocks of blocks bootstrap variance estimator when σ_t and/or α_t are time varying. In this more general and practically relevant case, part c) of Lemma 3.1 shows that $V_{n,p}^*$ diverges to ∞ in probability as $(n,p)_{seq} \to \infty$. The main reason for this inconsistency result is that $B_p \xrightarrow{P} \infty$ as $p \to \infty$. Notice that even though the limit derived in part c) is sequential, we can conclude that the same result holds as $(n,p) \to \infty$ jointly. The argument is as follows. Suppose it was the case that $V_{n,p}^* \xrightarrow{P} V \equiv \lim Var\left(n^{1/4} PRV_n\right)$, as $(n,p) \to \infty$ jointly. Then by Lemma 5 of Phillips and Moon (1999), we should have that $V_n^*(p) \xrightarrow{P} V$ sequentially as $(n,p)_{seq} \to \infty$, which is in contradiction with the result of part c). Hence, $V_{n,p}^*$ *cannot* converge in probability to V, as $(n,p) \to \infty$ jointly. More generally, we can show that if the joint limit of $V_{n,p}^*$ exists, then by the same argument, it must coincide with the sequential limit. Since we actually proved that $V_{n,p}^* \xrightarrow{P} \infty$ sequentially as $(n,p)_{seq} \to \infty$, this implies $V_{n,p}^*$ must diverge as $(n,p) \to \infty$ jointly.

Lemma 3.1 suggests that the blocks of blocks bootstrap is consistent for the variance of PRV_n only under constant volatility, constant conditional variance of noise and if we let the bootstrap block size b_n grow at a faster rate than the pre-averaging block size k_n. This result is related to a consistency result of the blocks of blocks bootstrap established in Bühlmann and Künsch (1995). As they showed, when the statistic of interest is an average of smooth functions of blocks of consecutive stationary strong mixing observations of size k_n, where k_n tends to infinity, the crucial condition for the block bootstrap to be valid is that the block size b_n grows at a faster rate than k_n. This is because the blocks over k_n observations (which in our case correspond to the pre-averaged returns) are strongly dependent for $|i-j| \leq k_n$, where $k_n \to \infty$, and b_n must be large enough to capture this dependence. Bühlmann and Künsch (1995) consider observations generated from a stationary strong mixing process and therefore they do not find any bias problem related to heterogeneity. Nevertheless, this becomes a problem in our context when volatility is stochastic. Therefore, a different bootstrap method is required to handle both the time dependence and the heterogeneity of pre-averaged returns.

Note that the inconsistency of the blocks of blocks bootstrap variance estimator for the asymptotic variance of PRV_n when the volatility is time-varying is not in contrast to the i.i.d. bootstrap results in Gonçalves and Meddahi (2009) for realized volatility (in the absence of noise). In particular, the i.i.d. bootstrap variance estimator of Gonçalves and Meddahi (2009) (cf. page 287) for the asymptotic variance of the realized volatility is given by

$$n \sum_{i=1}^{n} (r_i^e)^4 - \left(\sum_{i=1}^{n} (r_i^e)^2 \right)^2 \to^P \underbrace{3 \int_0^1 \sigma_t^4 dt - \left(\int_0^1 \sigma_t^2 dt \right)^2}_{\equiv V_{GM} < \infty},$$

which is equal to $2\int_0^1 \sigma_t^4 dt$ (i.e. the asymptotic conditional variance of the realized volatility) only when the volatility is constant.

This means that even in the absence of noise, when the volatility is time-varying we would not use the i.i.d. bootstrap method of Gonçalves and Meddahi (2009)) to compute standard errors of statistics based on functional of realized volatility. However, note that although the i.i.d. bootstrap method in Gonçalves and Meddahi (2009) does not consistently estimate the asymptotic variance of realized volatility, their bootstrap method is still asymptotically valid for studentized (percentile-t) bootstrap intervals. This is not necessary the case for the blocks of blocks bootstrap method applied to PRV_n. The main reason is that when the volatility is time-varying, and the bootstrap block size b_n grow faster than k_n (i.e., the more realistic case of choice of b_n), $V_{n,p}^* \xrightarrow{P} \infty$ as $(n,p) \to \infty$ jointly.

4 The wild blocks of blocks bootstrap

In this section, we propose and study the consistency of a novel bootstrap method for pre-averaged returns based on overlapping blocks of k_n intraday returns. It combines the blocks of blocks bootstrap with the wild bootstrap and in this manner gets rid of the bias term B_p associated with the blocks of blocks bootstrap variance V_n^* in (10).

Here, let b_n a sequence of integers such that

$$b_n \propto n^\delta, \tag{11}$$

where $\delta \in (0,1)$, and assume that J_n is such that $J_n \cdot b_n = N_n$. Let $\upsilon_1, \ldots, \upsilon_{J_n}$ be i.i.d. random variables whose distribution is independent of the original sample. Denote by $\mu_q^* = E^*\left(\upsilon_j^q\right)$ its q-th order moments. For $j = 1, \ldots, J_n$, let

$$\bar{B}_j = \frac{1}{b_n}\sum_{i=1}^{b_n} \bar{Y}^2_{i-1+(j-1)b_n}$$

denote the block average of the squared pre-averaged returns $\bar{Y}^2_{i-1+(j-1)b_n}$ for block j, we also let $\eta_j = \upsilon_j^2$. We then generate the bootstrap pre-averaged squared returns as follows,

$$\bar{Y}^{*2}_{i-1+(j-1)b_n} = \bar{B}_{j+1} + \left(\bar{Y}^2_{i-1+(j-1)b_n} - \bar{B}_{j+1}\right)\eta_j, \text{ for } 1 \le j \le J_n - 1 \text{ and for } 1 \le i \le b_n. \tag{12}$$

For the last block $j = J_n$, $\bar{B}_{j+1}$ is not available and therefore we let

$$\bar{Y}^{*2}_{i-1+(j-1)b_n} = \bar{B}_j + \left(\bar{Y}^2_{i-1+(j-1)b_n} - \bar{B}_j\right)\eta_j, \text{ for } 1 \le i \le b_n. \tag{13}$$

Our method is related to the wild bootstrap approach of Wu (1986) and Liu (1988). More specifically, in Wu (1986) and Liu (1988), the statistic of interest is $\bar{X}_n$, where X_i is independently but heterogeneously distributed with mean μ_i and variance σ_i^2. Their wild bootstrap generates X_i^* as

$$X_i^* = \bar{X}_n + \left(X_i - \bar{X}_n\right)\eta_i, \text{ for } 1 \le i \le n,$$

where η_i is i.i.d. $(0,1)$. Liu (1988) shows that the bootstrap distribution of $\sqrt{n}\left(\bar{X}_n^* - \bar{X}_n\right)$ is consistent for the distribution of $\sqrt{n}\left(\bar{X}_n - \bar{\mu}_n\right)$, where $\bar{\mu}_n = n^{-1}\sum_{i=1}^n \mu_i$, provided $\frac{1}{n}\sum_{i=1}^n \left(\mu_i - \bar{\mu}_n\right)^2 \to 0$ (and some other regularity conditions).

Our bootstrap method can be seen as a generalization of the wild bootstrap of Wu (1986) and Liu (1988) to the k_n-dependent case. In particular, here the statistic of interest is an average of blocks of observations of size k_n,

$$\widetilde{PRV}_n = \frac{1}{N_n}\sum_{i=1}^{N_n} Z_i,$$

where $Z_i \equiv \frac{N_n}{k_n}\frac{1}{\psi_2}\bar{Y}^2_{i-1}$ has time-varying moments and is k_n-dependent (conditionally on X), i.e. Z_i is independent of Z_j for all $|i-j| > k_n$.

To preserve the serial dependence, we divide the data into J_n non-overlapping blocks of size b_n and generate the bootstrap observations within a given block j using the same external random variable η_j. This preserves the dependence within each block. When there is no dependence, we can take $b_n = 1$, in which case our bootstrap method amounts to Liu's wild bootstrap with one difference: instead of centering each bootstrap observation Z_i^* around the overall mean $\widetilde{PRV}_n$, we center Z_i^* around Z_{i+1}. The reason for the new centering is that μ_i in our context does not satisfy Liu's condition $\frac{1}{n}\sum_{i=1}^n (\mu_i - \bar{\mu}_n)^2 \to 0$ (unless volatility is constant). Hence centering around $\widetilde{PRV}_n$ does not work here. Instead, we show that centering around Z_{i+1} yields an asymptotically valid bootstrap method for $\widetilde{PRV}_n$ even when volatility is stochastic.

The bootstrap data generating process (12) and (13) yields a bootstrap sample $\left\{\bar{Y}_0^{*2}, \ldots, \bar{Y}_{N_n-1}^{*2}\right\}$ which we use to compute

$$\widetilde{PRV}_n^* = \frac{1}{\psi_2 k_n} \sum_{i=1}^{N_n} \bar{Y}_{i-1}^{*2},$$

the wild blocks of blocks bootstrap analogue of $\widetilde{PRV}_n$. Let

$$\bar{B}_j^* = \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}_{i-1+(j-1)b_n}^{*2}$$

be the bootstrap analogue of $\bar{B}_j$. Given (12), we have that for $j = 1, \ldots, J_n - 1$,

$$\bar{B}_j^* = \bar{B}_{j+1} + \left(\bar{B}_j - \bar{B}_{j+1}\right)\eta_j,$$

whereas from (13), $\bar{B}_j^* = \bar{B}_j$ for $j = J_n$. This implies that we can write

$$\begin{aligned}
\widetilde{PRV}_n^* &= \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n} \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}_{i-1+(j-1)b_n}^{*2} = \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n-1} \bar{B}_j^* + \frac{b_n}{\psi_2 k_n} \bar{B}_{J_n}^* \\
&= \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n-1} \left[\bar{B}_{j+1} + \left(\bar{B}_j - \bar{B}_{j+1}\right)\eta_j\right] + \frac{b_n}{\psi_2 k_n} \bar{B}_{J_n}.
\end{aligned}$$

We can now easily obtain the bootstrap mean and variance of PRV_n^*. In particular,

$$E^*\left(\widetilde{PRV}_n^*\right) = \frac{b_n}{\psi_2 k_n}\left(\sum_{j=1}^{J_n-1} \bar{B}_{j+1} + \bar{B}_{J_n}\right) + \frac{b_n}{\psi_2 k_n} \sum_{j=1}^{J_n-1} \left(\bar{B}_j - \bar{B}_{j+1}\right) E^*\left(\eta_j\right), \tag{14}$$

and

$$V_n^* \equiv Var^*\left(n^{1/4}\widetilde{PRV}_n^*\right) = \frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \left(\bar{B}_j - \bar{B}_{j+1}\right)^2 Var^*\left(\eta_j\right).$$

Our next result studies the convergence of V_n^* when b_n satisfies (11) such that $1/2 < \delta < 2/3$. To prove the consistency of V_n^* for V we impose the following additional condition.

Assumption 3. σ_t is locally bounded away from zero and is a continuous semimartingale.

This assumption rule out jumps in σ_t and is common in the realized volatility literature (e.g., equation (3) of Barndorff-Nielsen et al. (2008) or equation (3) of Gonçalves and Meddahi (2009)). We can prove the following results.

Lemma 4.1 *Suppose Assumptions 1, 2 and 3 hold and the block size b_n satisfies (11) such that $1/2 < \delta < 2/3$. Let $V_n^* \equiv Var^*\left(n^{1/4}\widetilde{PRV}_n^*\right)$ denote the wild blocks of blocks bootstrap variance of $n^{1/4}\widetilde{PRV}_n^*$ based on a block length equal to b_n and external random variables $\eta_j \sim$ i.i.d. with mean $E^*(\eta_j)$ and variance $Var^*(\eta_j) = 1/2$. Then,*

$$p\lim_{n\to\infty} V_n^* = V \equiv \lim_{n\to\infty} Var\left(n^{1/4}PRV_n\right),$$

This result shows that if we let $\delta > 1/2$, i.e., b_n grow faster than k_n (i.e., $b_n/k_n \to \infty$) but such that $b_n/n \to 0$ and $Var^*(\eta_j) = 1/2$, the wild blocks bootstrap variance estimator is consistent for the asymptotic variance of PRV_n under Assumptions 1, 2 and 3. Given the consistency of the bootstrap variance estimator, and the fact that it is possible to obtain an exact and explicit formula of V_n^*, one may simply use V_n^* in place of $\hat{V}_n$ given by (7) as alternative consistent estimator of V. Together with the CLT result (5), we have that

$$\frac{n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right)}{\sqrt{V_n^*}} \to^{st} N(0,1).$$

As alternative method of inference (which does not require any resampling of one's data), we can use this feasible asymptotic distribution result to build confidence intervals for integrated volatility. In particular, a two-sided feasible $100(1-\alpha)\%$ level interval for $\int_0^1 \sigma_s^2 ds$ is given by:

$$IC^b_{Feas,1-\alpha} = \left(PRV_n - z_{1-\alpha/2} n^{-1/4}\sqrt{V_n^*}, PRV_n + z_{1-\alpha/2} n^{-1/4}\sqrt{V_n^*}\right), \tag{15}$$

where

$$V_n^* = \frac{n^{1/2} b_n^2}{2\psi_2^2 k_n^2} \sum_{j=1}^{J_n-1} \left(\bar{B}_j - \bar{B}_{j+1}\right)^2, \tag{16}$$

$z_{1-\alpha/2}$ is such that $\Phi\left(z_{1-\alpha/2}\right) = 1 - \alpha/2$, and $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution.

The structure of the wild blocks of blocks bootstrap method somehow seems to be related to the ideas in the recent paper of Mykland and Zhang (2014). To see this, it may be helpful to rewrite V_n^* given by (16) as follows

$$V_n^* = n^{1/2} \cdot \left[\frac{1}{2} \sum_{j=1}^{J_n-1} \left(\hat{B}_{j+1} - \hat{B}_j\right)^2\right], \tag{17}$$

where

$$\hat{B}_j = \frac{1}{\psi_2 k_n} \sum_{i=1}^{b_n} \bar{Y}^2_{i-1+(j-1)b_n}$$

denote the analogue of $\widetilde{PRV}_n$ computed for the block j. Hence, one can show that the formula for V_n^* (given by (17)) is related to the general nonparametric method proposed by Mykland and Zhang (2014). In particular, given results in equations (7) and (11) in Mykland and Zhang (2014), it is easy to see that under some regularity conditions the asymptotic variance (AVAR) of many estimators, say $\Theta = \int_0^1 \tilde{\theta}_t dt$, in the high-frequency literature can be estimated based on

$$\widehat{\text{AVAR}}\left(\hat{\Theta} - \Theta\right) = \frac{1}{2}\sum_{j=1}^{J_n-1}\left(\hat{\Theta}_{j+1} - \hat{\Theta}_j\right)^2, \tag{18}$$

where $\hat{\Theta}_j$ is the estimator Θ calculated on the j-th block and such that $\Theta_j = \int_{(j-1)b_n/n}^{jb_n/n} \tilde{\theta}_t dt$. More precisely, under some regularity conditions (including negligible edge effect and continuous spot process $\widetilde{\theta}_t$) equations (7) and (11) in Mykland and Zhang (2014) amount to,

$$\sum_{j=1}^{J_n-1}\left(\hat{\Theta}_{j+1} - \hat{\Theta}_j\right)^2 = \left(2\sum_{j=1}^{J_n-1}\text{AVAR}\left(\hat{\Theta}_j - \Theta_j\right) + \frac{2}{3}\left(\frac{b_n}{n}\right)^2\left[\widetilde{\theta}, \widetilde{\theta}\right]_1\right)(1 + o_p(1)), \tag{19}$$

and

$$\text{AVAR}\left(\hat{\Theta} - \Theta\right) = \left(\sum_{j=1}^{J_n-1}\text{AVAR}\left(\hat{\Theta}_j - \Theta_j\right)\right)(1 + o_p(1)), \tag{20}$$

respectively, where $\left[\widetilde{\theta}, \widetilde{\theta}\right]_1$ is the total quadratic variation of the spot process $\widetilde{\theta}_t$ over the whole interval from 0 to 1. Given (19) and (20), it follows that

$$\text{AVAR}\left(\hat{\Theta} - \Theta\right) = \frac{1}{2}\sum_{j=1}^{J_n-1}\left(\hat{\Theta}_{j+1} - \hat{\Theta}_j\right)^2 - \frac{4}{3}\left(\frac{b_n}{n}\right)^2\left[\widetilde{\theta}, \widetilde{\theta}\right]_1(1 + o_p(1)). \tag{21}$$

Thus, if $\frac{b_n}{n}$ can be taken to be small enough, then one can simply use (18), i.e., a one scale estimator by ignoring the $\left[\widetilde{\theta}, \widetilde{\theta}\right]_1$ term. Note that given the normalization of AVAR in Mykland and Zhang (2014) (cf. footnote 1), we have AVAR=AVAR$_n = n^{-2\alpha}V$, where $\hat{\Theta}$ is such that $n^{\alpha}\left(\hat{\Theta} - \Theta\right) \to^{st} N(0, V)$, for some $\alpha > 0$. Thus, in our context, the one scale estimator formula applied to $\hat{\Theta} = \widetilde{PRV}_n$ with $\alpha = 1/4$, gives

$$\widehat{\text{AVAR}}\left(\hat{\Theta} - \Theta\right) = \frac{1}{2}\sum_{j=1}^{J_n-1}\left(\hat{B}_{j+1} - \hat{B}_j\right)^2 = n^{-1/2}V_n^*.$$

We emphasize that the paper by Mykland and Zhang (2014) goes much further in developing the asymptotic variance estimator, including estimators with hard edge effect and allowing non continuous spot process. In particular, Mykland and Zhang (2014) show that by subsampling and averaging one can still use a result akin to (19) when $\widetilde{\theta}_t$ is a general semimartingale. In addition, they argue that subsampling and averaging can at the same time help to deal with hard edge effect (which can lead the additivity in (20) to fail. Thus, the final estimator is more complicated and is based on two- or multi-scale construction. Their approach aims to avoid using the information on the asymptotic variance of $\hat{\Theta}$ (for instance, no need to know the closed form of the AVAR). However, this is not without consequences for their method. For example they have to introduce an additional layer of blocks to implement the bias-correction term. The asymptotic variance estimator of $\hat{\Theta}$ can also go negative in finite samples, which is not the case of the bootstrap. This relationship with Mykland and Zhang (2014), in particular the way both method managed blocks of adjacent summands suggests that our wild blocks of blocks bootstrap approach may be applied very generally in the field of nonparametric estimation with infill asymptotic. The exploration of this is beyond the scope of this paper.

Our next result proves the consistency of the bootstrap distribution of $n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)$.

Theorem 4.1 *Suppose Assumptions 1, 2 and 3 hold such that for any $\varepsilon \geq 2$, $E\left(\epsilon_t^{4(2+\varepsilon)}\right) < \infty$, and the block size b_n satisfies (11) such that $1/2 < \delta < 2/3$. Let $\widetilde{PRV}_n^*$ be the pre-averaged realized volatility estimator based on a block length equal to b_n and an external random variable $\eta_j \sim$ i.i.d. $\left(E^*\left(\eta_j\right), Var^*\left(\eta_j\right)\right)$ such that $Var^*\left(\eta_j\right) = \frac{1}{2}$, and for some $\varepsilon \geq 2$ $E^*\left|\eta_j\right|^{2+\varepsilon} \leq \Delta < \infty$. Then*

$$\sup_{x\in\mathbb{R}}\left|P^*\left(n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right) \leq x\right) - P\left(n^{1/4}\left(PRV_n - \int_0^1 \sigma_s^2 ds\right) \leq x\right)\right| \to^P 0 \text{ as } n \to \infty.$$

Theorem 4.1 justifies using the wild blocks of blocks bootstrap to construct bootstrap percentile intervals for integrated volatility. Specifically, a $100(1-\alpha)\%$ symmetric bootstrap percentile interval for integrated volatility based on the bootstrap is given by

$$IC_{perc,1-\alpha}^* = \left(PRV_n - n^{-1/4}p_{1-\alpha}^*, PRV_n + n^{-1/4}p_{1-\alpha}^*\right), \tag{22}$$

where $p_{1-\alpha}^*$ is the $1-\alpha$ quantile of the bootstrap distribution of $\left|n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)\right|$. Next, we propose a consistent bootstrap variance estimator that allows us to form bootstrap percentile-t intervals. More specifically, we can show that the following bootstrap variance estimator consistently estimates V_n^* for any choice of the external random variable η_j:

$$\hat{V}_n^* = \frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\sum_{j=1}^{J_n-1}\left(\bar{B}_j^* - \bar{B}_{j+1}\right)^2.$$

Our proposal is to use this estimator to construct a bootstrap studentized statistic,

$$T_n^* \equiv \frac{n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)}{\sqrt{\hat{V}_n^*}},$$

the bootstrap analogue of T_n.

Theorem 4.2 *Suppose Assumptions 1, 2 and 3 hold such that for any $\varepsilon \geq 2$, $E\left(\epsilon_t^{4(2+\varepsilon)}\right) < \infty$, and the block size b_n satisfies (11) such that $1/2 < \delta < 2/3$. Let $\widetilde{PRV}_n^*$ be the pre-averaged realized volatility estimator based on a block length equal to b_n and an external random variable $\eta_j \sim$ i.i.d. $\left(E^*\left(\eta_j\right), Var^*\left(\eta_j\right)\right)$ such that for some $\varepsilon \geq 2$ $E^*\left|\eta_j\right|^{2+\varepsilon} \leq \Delta < \infty$. Then*

$$\sup_{x\in\mathbb{R}} \left|P^*\left(T_n^* \leq x\right) - P\left(T_n \leq x\right)\right| \to^P \quad as\ n \to \infty.$$

Theorem 4.2 justifies constructing bootstrap percentile-t intervals. In particular, a $100\left(1-\alpha\right)\%$ symmetric bootstrap percentile-t interval for integrated volatility is given by

$$IC^{*a}_{perc-t,1-\alpha} = \left(PRV_n - q^*_{1-\alpha} n^{-1/4}\sqrt{\hat{V}_n}, PRV_n + q^*_{1-\alpha} n^{-1/4}\sqrt{\hat{V}_n}\right), \tag{23}$$

or alternatively we can use

$$IC^{*b}_{perc-t,1-\alpha} = \left(PRV_n - q^*_{1-\alpha} n^{-1/4}\sqrt{V_n^*}, PRV_n + q^*_{1-\alpha} n^{-1/4}\sqrt{V_n^*}\right), \tag{24}$$

where $q^*_{1-\alpha}$ is the $(1-\alpha)$-quantile of the bootstrap distribution of $|T_n^*|$.

5 Monte Carlo results

In this section, we compare the finite sample performance of the bootstrap with the feasible asymptotic theory for confidence intervals of integrated volatility in the case of i.i.d. and autocorrelated market microstructure noise.

We consider two data generating processes in our simulations. First, following Zhang et al. (2005), we use the one-factor stochastic volatility (SV1F) model of Heston (1993) as our data-generating process, i.e.

$$dX_t = \left(\mu - \nu_t/2\right) dt + \sigma_t dB_t,$$

and

$$d\nu_t = \kappa\left(\widetilde{\alpha} - \nu_t\right) dt + \gamma\left(\nu_t\right)^{1/2} dW_t,$$

where $\nu_t = \sigma_t^2$, and we assume $Corr(B, W) = \rho$. The parameter values are all annualized. In particular, we let $\mu = 0.05/252$, $\kappa = 5/252$, $\widetilde{\alpha} = 0.04/252$, $\gamma = 0.05/252$, $\rho = -0.5$. For $i = 1, \ldots, n$, we let the market microstructure noise be defined as $\epsilon_{\frac{i}{n}} \sim$ i.i.d.$N\left(0, \alpha\right)$. The size of the noise is an important parameter. We follow Barndorff-Nielsen et al. (2008) and model the noise magnitude as $\xi^2 = \alpha/\sqrt{\int_0^1 \sigma_s^4 ds}$. We fix ξ^2 equal to 0.0001, 0.001 and 0.01 and let $\alpha = \xi^2\sqrt{\int_0^1 \sigma_s^4 ds}$. These values are motivated by the empirical study of Hansen and Lunde (2006), who investigate 30 stocks of the Dow Jones Industrial Average.

We also consider a more realistic two-factor stochastic volatility (SV2F) model analyzed by

Barndorff-Nielsen et al. (2008), where[5]

$$
\begin{aligned}
dX_t &= adt + \sigma_t dW_t, \\
\sigma_t &= s\text{-}exp\left(\beta_0 + \beta_1 \tau_{1t} + \beta_2 \tau_{2t}\right), \\
d\tau_{1t} &= \widetilde{\alpha}_1 \tau_{1t} dt + dB_{1t}, \\
d\tau_{2t} &= \widetilde{\alpha}_2 \tau_{2t} dt + \left(1 + \phi \tau_{2t}\right) dB_{2t}, \\
corr\left(dW_t, dB_{1t}\right) &= \varphi_1,\ corr\left(dW_t, dB_{2t}\right) = \varphi_2.
\end{aligned}
$$

We follow Huang and Tauchen (2005) and set $a = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\widetilde{\alpha}_1 = -0.00137$, $\widetilde{\alpha}_2 = -1.386$, $\phi = 0.25$, $\varphi_1 = \varphi_2 = -0.3$. We initialize the two factors at the start of each interval by drawing the persistent factor from its unconditional distribution, $\tau_{10} \sim N\left(0, \frac{-1}{2\widetilde{\alpha}_1}\right)$, and by starting the strongly mean-reverting factor at zero.

We simulate data for the unit interval $[0,1]$ and normalize one second to be $1/23400$, so that $[0,1]$ is thought to span 6.5 hours. The observed Y process is generated using an Euler scheme. We then construct the $\frac{1}{n}$-horizon returns $r_i \equiv Y_{i/n} - Y_{(i-1)/n}$ based on samples of size n.

We use two different values of θ: $\theta = 1/3$, as in Jacod et al. (2009), and $\theta = 1$, as in Christensen, Kinnebrock and Podolskij (2010). The latter value corresponds to a conservative choice of k_n. We also follow the literature and use the weight function $g(x) = \min(x, 1-x)$ to compute the pre-averaged returns.

In order to reduce finite sample biases associated with Riemann integrals, we follow Jacod et al. (2009) and Hautsch and Podolskij (2013) and use the finite sample adjustments version of the pre-averaged realized volatility estimator,

$$
PRV_n^a = \left(1 - \frac{\psi_1^{k_n}}{2n\theta^2 \psi_2^{k_n}}\right)^{-1} \left(\frac{n}{n - k_n + 2} \frac{1}{\psi_2^{k_n} k_n} \sum_{i=0}^{n-k_n+1} \bar{Y}_i^2 - \frac{\psi_1^{k_n}}{2n\theta^2 \psi_2^{k_n}} \sum_{i=1}^{n} r_i^2\right),
$$

where $\psi_1^{k_n} = k_n \sum_{i=1}^{k_n} \left(g\left(\frac{i}{k_n}\right) - g\left(\frac{i-1}{k_n}\right)\right)^2$ and $\psi_2^{k_n} = \frac{1}{k_n} \sum_{i=1}^{k_n} g^2\left(\frac{i}{k_n}\right)$. Similarly, $\hat{V}_n$ as defined in (7) replaces Φ_{11}, Φ_{12} and Φ_{22} by their Riemann approximations,

$$
\Phi_{11}^{k_n} = k_n \left(\sum_{i=1}^{k_n} \left(\phi_1^{k_n}(j)\right)^2 - \frac{1}{2}\left(\phi_1^{k_n}(0)\right)^2\right), \quad \Phi_{12}^{k_n} = \frac{1}{k_n}\left(\sum_{i=1}^{k_n} \phi_1^{k_n}(j)\, \phi_2^{k_n}(j) - \frac{1}{2}\phi_1^{k_n}(0)\, \phi_2^{k_n}(0)\right), \text{ and}
$$

$$
\Phi_{22}^{k_n} = \frac{1}{k_n^3}\left(\sum_{i=1}^{k_n} \left(\phi_2^{k_n}(j)\right)^2 - \frac{1}{2}\left(\phi_2^{k_n}(0)\right)^2\right),
$$

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

where

$$\phi_1^{k_n}(j) = k_n \sum_{i=j+1}^{k_n-1} \left(g\left(\frac{i-1}{k_n}\right) - g\left(\frac{i}{k_n}\right) \right) \left(g\left(\frac{i-j-1}{k_n}\right) - g\left(\frac{i-j}{k_n}\right) \right), \text{ and}$$

$$\phi_2^{k_n}(j) = \sum_{i=j+1}^{k_n-1} g\left(\frac{i}{k_n}\right) g\left(\frac{i-j}{k_n}\right).$$

Tables 1, 2, 3 and 4 give the actual rates of 95% confidence intervals of integrated volatility as well as the average lengths of the confidence intervals for the SV1F and the SV2F models, respectively, computed over 10,000 replications. Results are presented for eight different samples sizes: n = 23400, 11700, 7800, 4680, 1560, 780, 390 and 195, corresponding to "1-second", "2-second", "3-second", "5-second", "15-second", "30-second", "1-minute" and "2-minute" frequencies. In our simulations, bootstrap intervals use 999 bootstrap replications for each of the 10,000 Monte Carlo replications. We consider the bootstrap percentile method computed at the 95% level. To generate the bootstrap data we use a two point distribution $\eta_j = v_j^2$ with $v_j \sim$ i.i.d. such that:

$$v_j = \begin{cases} \left(\frac{1}{2}\right)^{1/4} \frac{-1+\sqrt{5}}{2}, & \text{with prob } p = \frac{\sqrt{5}-1}{2\sqrt{5}} \\ \left(\frac{1}{2}\right)^{1/4} \frac{-1-\sqrt{5}}{2}, & \text{with prob } 1-p = \frac{\sqrt{5}+1}{2\sqrt{5}} \end{cases},$$

for which $\mu_2^* = \sqrt{2}$ and $\mu_4^* = 5/2$, implying that $Var^*(\eta_j) = 1/2$. This choice of η_j is asymptotically valid when used to construct bootstrap percentile as well as percentile-t intervals. The choice of the bootstrap block size is critical. We follow Politis, Romano and Wolf (1999) and use the minimum volatility method to choose the bootstrap block. Details of the algorithm are given in Appendix A.

5.1 i.i.d. noise

In this subsection, we simulate results for the case of i.i.d. market microstructure noise. For the CLT-based intervals and the wild blocks of blocks bootstrap-based intervals, Tables 1 and 2 show that for the two models, all intervals tend to undercover. The degree of undercoverage is especially large for smaller values of n, when sampling is not too frequent. The SV2F model exhibits overall larger coverage distortions than the SV1F model, for all sample sizes. Results are sensitive to the value of the tuning parameter θ. When $\theta = 1/3$, larger market microstructure effects induce larger coverage distortions. In particular, the coverage distortions are very important when $\xi^2 = 0.01$ in comparison to the case where market microstructure effects are moderate or negligible ($\xi^2 = 0.001$ and $\xi^2 = 0.0001$). This reflects the fact that for this value of θ, k_n is not sufficiently large to allow pre-averaging to remove the market microstructure bias. The pre-averaged estimator is biased in finite samples and this explains the finite sample distortions. In contrast, for the conservative choice of k_n, results are not very sensitive to the noise magnitude. The reason is that the larger is the block size over which the pre-averaging is done, the smaller is the impact of the noise.

In all cases, the wild blocks of blocks bootstrap outperforms the existing first order asymptotic theory. As expected, the average chosen block size is larger for larger sample sizes, but our results show that it is not sensitive to the noise magnitude. This is because the noise magnitude is almost irrelevant for the intensity of the autocorrelation of the square pre-averaged returns (as confirmed by simulations not reported here).

5.2 Autocorrelated noise

In a second set of experiments, we look at the case where the market microstructure noise is autocorrelated. Empirically the conditional independence noise assumption is somewhat unrealistic for ultra-high frequency data (see, among others, Hansen and Lunde (2006)). This is in fact one of the motivations behind the approach of Hautsch and Podolskij (2013). Their results relax the conditional independence assumption on ϵ to allow for q-dependent noise, at the cost of not allowing for time varying variances of the noise process and dependence between X and ϵ. Indeed, the main consistency result for Jacod et al. (2009) pre-averaged estimators (cf. their Theorem 3.1) still holds. The key difference is that the limit (of the required bias-correction term) now depends on the higher order autocorrelations of the noise process instead of depending on $\alpha_t = E\left(\epsilon_t^2|X\right)$ (in particular, α_t is replaced by the long run variance $\rho^2 = \rho\left(0\right) + 2\sum_{k=1}^{q} \rho\left(k\right)$, where $\rho\left(k\right) = Cov\left(\epsilon_1, \epsilon_{1+k}\right)$, and q is the order of dependence of the noise process $\left(\epsilon_i\right)_{i\geq 0}$). The main implication is that the bias correction for pre-averaged realized volatility must depend on an estimator of ρ^2. Hautsch and Podolskij (2013) discuss an estimator of ρ^2 given by

$$\rho_n^2 = \rho_n\left(0\right) + 2\sum_{k=1}^{q} \rho_n\left(k\right),$$

where $\rho_n\left(0\right), \ldots, \rho_n\left(q\right)$ are obtained by a simple recursion,

$$\begin{aligned}
\rho_n\left(q\right) &= -\gamma_n\left(q+1\right), \\
\rho_n\left(q-1\right) &= -\gamma_n\left(q\right) + 2\rho_n\left(q\right), \\
\rho_n\left(q-2\right) &= -\gamma_n\left(q-1\right) + 2\rho_n\left(q-1\right) - \rho_n\left(q\right), \\
\text{where } \gamma_n\left(k\right) &= \frac{1}{n}\sum_{i=1}^{n} r_i r_{i+k}, \qquad k = 0, \ldots, q+1.
\end{aligned}$$

This implies the following consistent estimator of integrated volatility under a q-dependent autocorrelated noise process:

$$PRV_n^d = \underbrace{\frac{1}{\psi_2 k_n} \sum_{i=1}^{n-k_n+1} \bar{Y}_i^2}_{RV\text{-like estimator}} - \underbrace{\frac{\psi_1}{\theta^2 \psi_2} \rho_n^2}_{\text{new bias correction term}} . \tag{25}$$

To obtain a feasible asymptotic procedure, Hautsch and Podolskij (2013) also propose the following consistent estimator of $V^d \equiv \lim_{n\to\infty} Var\left(n^{1/4} PRV_n^d\right)$:

$$\begin{aligned} \hat{V}_n^d &= \frac{4\Phi_{22}}{3\theta\psi_2^4} \sum_{i=0}^{n-k_n+1} \overline{Y}_i^4 + \frac{8\rho_n^2}{\theta^2\sqrt{n}} \left(\frac{\Phi_{12}}{\psi_2^3} - \frac{\Phi_{22}\psi_1}{\psi_2^4} \right) \sum_{i=0}^{n-2k_n+1} \overline{Y}_i^2 \\ &+ \frac{4\rho_n^4}{\theta^3} \left(\frac{\Phi_{11}}{\psi_2^2} - 2\frac{\Phi_{12}\psi_1}{\psi_2^3} + \frac{\Phi_{22}\psi_1^2}{\psi_2^4} \right). \end{aligned} \tag{26}$$

We conjecture that the wild blocks of blocks bootstrap remains valid when we relax the conditional independence assumption on ϵ_i provided we use it to approximate the distribution of PRV_n^d. Indeed, the conditional independence noise assumption used in our proof in Appendix B is not essential to guarantee the consistency of the wild blocks of blocks bootstrap variance since we do not use any prior knowledge on ϵ_i apart from the k_n-dependence of $\bar{\epsilon}_i$. If ϵ_i is a q-dependent sequence, then $\bar{\epsilon}_i$ becomes $(k_n + q)$-dependent, and the result of Lemma 4.1 still holds, although higher order autocorrelations of ϵ appear in the limit. So long as $E\left(\epsilon_t | X\right) = 0$, $\bar{\epsilon}_i$ admits asymptotic normality at the usual rate $k_n^{-1/2}$, (see e.g. the proof of Lemma 1 of Hautsch and Podolskij (2013)), and if we let the block size b_n grow faster than $k_n + q$ and set $Var^*\left(\eta_j\right) = 1/2$, then the wild blocks of blocks bootstrap variance estimator will remain consistent for V^d. Moreover, by using the wild blocks of blocks bootstrap, a stationarity condition on ϵ is not required, since by construction it is robust to the heterogeneity of square pre-averaged returns. These facts lead us to conjecture that the wild blocks of blocks bootstrap is valid when applied to the new bias adjusted pre-averaged volatility estimator under autocorrelated noise. Although we do not provide a detailed proof of this result, in this section we explore the finite sample properties of the wild bootstrap under autocorrelation in ϵ_i.

In particular, we follow Kalnina (2011) and let the market microstructure noise be generated as an $MA(1)$ process (for a given frequency of the observations):

$$\epsilon_{\frac{i}{n}} = u_{\frac{i}{n}} + \lambda u_{\frac{i-1}{n}}, \quad u_{\frac{i}{n}} \sim \text{i.i.d.} N\left(0, \frac{\alpha}{1+\lambda^2}\right), \tag{27}$$

so that $Var\left(\epsilon\right) = \alpha$. Three different values of λ are considered, $\lambda = -0.3$, -0.5, and $\lambda = -0.9$. We chose α as in the i.i.d. case discussed above, i.e. we let $\alpha = \xi^2 \sqrt{\int_0^1 \sigma_s^4 ds}$. We let $\theta = 1$ (conservative choice of k_n).

Our aim here is to evaluate by Monte Carlo simulation the performance of the wild blocks of blocks bootstrap when applied to the statistic that relies on the new bias correction of Hautsh and Podoskij (2013), which is robust to noise autocorrelation. We consider five types of intervals (two types of intervals based on the asymptotic normal distribution under the label CLT1 and CLT2 and three types of intervals based on the wild blocks of blocks bootstrap under the label Boot1, Boot2 and Boot3), computed at the 95% level. More specifically, for the asymptotic theory-based approach we consider the following intervals,

$$\left(PRV_n^d - 1.96n^{-1/4}\sqrt{\hat{V}_n^d}, PRV_n^d + 1.96n^{-1/4}\sqrt{\hat{V}_n^d}\right), \text{ and} \tag{28}$$

$$\left(PRV_n^d - 1.96n^{-1/4}\sqrt{V_n^*}, PRV_n^d + 1.96n^{-1/4}\sqrt{V_n^*}\right). \tag{29}$$

For the bootstrap, we consider

$$\left(PRV_n^d - n^{-1/4}p^*_{0.95}, PRV_n^d + n^{-1/4}p^*_{0.95}\right), \tag{30}$$

$$\left(PRV_n^d - q^*_{1-\alpha}n^{-1/4}\sqrt{\hat{V}_n}, PRV_n^d + q^*_{1-\alpha}n^{-1/4}\sqrt{\hat{V}_n}\right), \text{ and} \tag{31}$$

$$\left(PRV_n^d - q^*_{1-\alpha}n^{-1/4}\sqrt{V_n^*}, PRV_n^d + q^*_{1-\alpha}n^{-1/4}\sqrt{V_n^*}\right). \tag{32}$$

Whereas (30) corresponds to bootstrap percentile intervals,(31) and (32) correspond to bootstrap percentile-t intervals. Note that for the bootstrap based-intervals, the bootstrap quantile $p^*_{0.95}$ and $q^*_{0.95}$ are computed exactly as in the i.i.d. noise case (it is based on the absolute value of $n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)$ and $n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)/\sqrt{\hat{V}_n^*}$, respectively, whose form is unaffected by the new bias adjustment used in PRV_n^d).

Tables 3 and 4 contains the results. We only report results for the SV2F model, since it is more empirically relevant and indeed it exhibits overall larger coverage distortions than the SV1F model. Two sets of results are presented. First, we present results for intervals based on PRV_n, the non-robust pre-averaged estimator discussed for the uncorrelated noise case (Table 3). Then, we present results for intervals based on PRV_n^d, the robust estimator based on the new bias correction of Hautsch and Podolskij (2013) (Table 4). The results show that intervals based on PRV_n are more distorted when market microstructure effects are moderate or high ($\xi^2 = 0.001$ and $\xi^2 = 0.01$) and there is autocorrelation in ϵ_i than otherwise. The main reason for the distortions is the fact that PRV_n is not correctly centered and standardized under autocorrelation. For instance, when $\lambda = -0.3$, n =195, and $\xi^2 = 0.01$ the CLT1-based interval has a coverage probability (from Table 3) equal to 72.98% under autocorrelated noise whereas its coverage rate is equal to 83.32% under uncorrelated noise. Although the difference is not very large for the smaller $|\lambda|$ (intensity of autocorrelation), it gets much bigger for larger values of $|\lambda|$. For $\lambda = -0.5$ and -0.9, and (n =195, and $\xi^2 = 0.01$) these rates equal 67.75% and 63.04%, respectively. Thus, the distortions increase with $|\lambda|$. Also for high effects of noise

($\xi^2 = 0.01$), the degree of undercoverage becomes especially large for larger values of n, when sampling is frequent. For instance when $\lambda = -0.5$, and ($n = 195$ and $n = 23400$) they are equal to 67.75% and 28.97%, respectively. This confirm the invalidity of intervals based on PRV_n under correlated noise. A similar pattern is observed for the CLT2-based intervals. We also see that these are close to the (percentile) Boot1-based intervals.

However, if we rely on PRV_n^d as a point estimator of integrated volatility, the corresponding intervals (both asymptotic and bootstrap) are better centered and standardized and the distortions are smaller and closer to their values under the uncorrelated noise case. For instance, for n =195, and $\xi^2 = 0.01$ the CLT1-based intervals now have coverage rates equal to 84.33% and 84.56% when $\lambda = -0.3$ and $\lambda = -0.5$, respectively. A similar pattern is observed for larger sample sizes, although the rates are overall larger. For instance, for $n = 23400$ they are equal to 93.59% and 93.77%, respectively.

When the wild blocks of blocks bootstrap method is used to compute critical values for the t-test based on PRV_n and the error is MA(1), for high effects of noise ($\xi^2 = 0.01$), coverage rates are usually smaller than those obtained when the noise is uncorrelated (and therefore distortions are larger). As for the CLT-based intervals, the larger differences occur for the larger values of $|\lambda|$. For the smaller values of $|\lambda|$, the difference in coverage probability between the two types of errors is almost negligible. As for the CLT-based intervals, using the wild blocks of blocks bootstrap to compute critical values for the t-statistic based on PRV_n^d essentially eliminates the difference in coverage probabilities observed between the uncorrelated and the MA(1) errors.

In summary, the results in Tables 3 and 4 show that under autocorrelated noise the statistic based on the bias correction of Hautsch and Podolskij (2013) works well and that the coverage rates of 95% nominal level intervals based on either the asymptotic mixed Gaussian distribution or the wild blocks of blocks bootstrap proposed in this paper are similar to those obtained under uncorrelated noise. In particular, the bootstrap (percentile-t) outperforms the asymptotic theory. Whereas, the results based on CLT2 and the (percentile) Boot1 intervals are close, but slightly different.

6 Empirical results

In this section, we implement the wild blocks of blocks bootstrap on high frequency data and compare it to the existing feasible asymptotic procedure of Jacod et al. (2009). The data consists of transaction log prices of General Electric (GE) shares carried out on the New York Stock Exchange (NYSE) in October 2011. We also consider transaction log prices of Microsoft (MSFT) in December 2010, taken from Thomson Reuter's Tick History. GE represents highly liquid stocks with approximately 27 trade arrivals per minute. Conversely, MSFT is significantly less liquid with approximately 6 trade arrivals per minute. Our procedure for cleaning the data is exactly identical to that used by Barndorff-Nielsen et al. (2008) (for further details see Barndorff-Nielsen et al. (2009)). For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. until 4 p.m.

We implement the pre-averaged realized volatility estimator of Jacod et al. (2009) on returns recorded every S transactions, where S is selected each day so that for GE and MSFT there are approximately 1493 and 82 observations a day, respectively. This means that on average, for GE and MSFT, these returns are recorded roughly every 15 seconds and 5 minutes, respectively. Table 5 in the Appendix provides the number of transactions per day, the sample size for the pre-averaged returns, and the dependent-noise robust version of the pre-averaged realized volatility estimator using (25) (for $q = 0, 1$ and 2). We also report the optimal value of q (the number of non-vanishing covariances) using the decision rule proposed by Hautsch and Podolskij (2013). To implement the pre-averaged realized volatility estimator, we select the tunning parameter θ by following the conservative rule ($\theta = 1$, implying that $k_n = \sqrt{n}$). To choose the block size b_n, we follow Politis, Romano and Wolf (1999) and use the minimum volatility method (see Appendix A for details). As illustrated below, these stocks represent different empirical features and thus allow to gain valuable insights into the empirical performance of the wild blocks of blocks bootstrap method.

For GE, Figure 1 in Appendix A shows daily 95% confidence intervals (CIs) for integrated volatility using both methods, the wild blocks of blocks bootstrap and the existing feasible asymptotic procedure of Jacod et al. (2009). In the latter case CIs are computed using (28) whereas for the bootstrap we use (32).

The confidence intervals based on the bootstrap method are usually wider than the confidence intervals using the feasible asymptotic theory.[6] This is especially true in periods with large volatility. To gain further insight on the behavior of our intervals for these periods, we implemented the test for jumps of Barndorff-Nielsen and Shephard (2006) using a moderate sample size (2-minute sampling intervals). It turns out that these days often correspond to days on which there is evidence for jumps (in particular for the 13, 17, 20 and 26 of October 2011). Since neither of the two types of intervals are valid in the presence of jumps, further analysis should be pursued for these particular days. In particular, we should rely on estimation methods that are robust to jumps such as the pre-averaged multipower variation method proposed by Podolskij and Vetter (2009) or the quantile estimation method of Christensen, Oomen, and Podolskij (2010).

Similarly for MSFT (the less liquid stock) Figure 2 in Appendix A shows daily 95% confidence intervals for integrated volatility. The same patterns also emerges as for GE. The confi-

[6]Nevertheless, as our Monte Carlo simulations showed, the latter typically have undercoverage problems whereas the bootstrap intervals have coverage rates closer to the desired level. Therefore if the goal is to control the coverage probability, shorter intervals are not necessarily better. The figures also show a lot of variability in the daily estimate of integrated volatility.

dence intervals based on the wild blocks of blocks bootstrap method are usually wider than the confidence intervals using the feasible asymptotic theory. In contrast to GE, for MSFT we have found no evidence of jumps at 5% significance level for days with large volatility. Importantly, the bootstrap based confidence sets of these days are larger than those based on the asymptotic theory, as suggested by the simulation study, which highlights the importance of using the bootstrap in these volatile days.

7 Conclusion

In this paper, we propose the bootstrap as a method of inference for integrated volatility in the context of the pre-averaged realized volatility estimator proposed by Jacod et al. (2009). We show that the "blocks of blocks" bootstrap method suggested by Politis and Romano (1992) is not valid when volatility is time-varying. This is due to the heterogeneity of the squared pre-averaged returns when volatility is stochastic.

To simultaneously handle the dependence and heterogeneity of the squared pre-averaged returns, we propose a novel bootstrap procedure that combines the wild and the blocks of blocks bootstrap. We provide a set of conditions under which this method is asymptotically valid to first order. Our Monte Carlo simulations show that the wild blocks of blocks bootstrap improves the finite sample properties of the existing first order asymptotic theory. In future work, we plan to generalize the wild blocks of blocks bootstrap for inference on multivariate integrated volatility as considered by Christensen, Kinnebrock and Podolskij (2010). Bootstrap variance-covariances matrices are naturally positive semi-definite, which is very important for empirical applications. Finally, taking into account the possible presence of jumps is an important extension that should be studied.

Appendix A: Simulation and empirical results

Here we describe the Minimum Volatility Method algorithm of Politis, Romano and Wolf (1999, Chapter 9) for choosing the block size b_n for a two-sided confidence interval.

Algorithm: Choice of the bootstrap block size by minimizing confidence interval volatility

(i) For $b = b_{small}$ to $b = b_{big}$ compute a bootstrap interval for IV at the desired confidence level, this resulting in endpoints $IC_{b,low}$ and $IC_{b,up}$.

(ii) For each b compute the volatility index VI_b as the standard deviation of the interval endpoints in a neighborhood of b. More specifically, for a smaller integer d, let VI_b equal to the standard deviation of the endpoints $\{IC_{b-d,low},\ldots,IC_{b+d,low}\}$ plus the standard deviation of the endpoints $\{IC_{b-d,up},\ldots,IC_{b+d,up}\}$, i.e.

$$VI_b \equiv \sqrt{\frac{1}{2d+1}\sum_{i=-d}^{d}\left(IC_{b+i,low}-\bar{IC}_{low}\right)^2}+\sqrt{\frac{1}{2d+1}\sum_{i=-d}^{d}\left(IC_{b+i,up}-\bar{IC}_{up}\right)^2},$$

where $\bar{IC}_{low}=\frac{1}{2d+1}\sum_{i=-d}^{d}IC_{b+i,low}$ and $\bar{IC}_{up}=\frac{1}{2d+1}\sum_{i=-d}^{d}IC_{b+i,up}$.

(iii) Pick the value b^* corresponding to the smallest volatility index and report $\{IC_{b^*,low},IC_{b^*,up}\}$ as the final confidence interval.

To make the algorithm more computationally efficient, we have skipped a number of b values in regular fashion between b_{small} and b_{big}. We have considered only the values of b such that $b=pk_n$ where p is a fixed integer. We employ $b_{small}=2k_n$, $b_{big}=\min(\theta\frac{N_n}{4},12k_n)$ and $d=2$.

Tables 1, 2, 3 and 4 report the actual coverage rates for the feasible asymptotic theory approach and for our bootstrap methods using the optimal block size by minimizing confidence interval volatility. In Table 5 we provide some statistics of GE and MSFT shares in October 2011 and December 2010, respectively.

Table 1. Coverage rates of nominal 95% intervals under i.i.d. noise with $\theta = 1/3$

	SV1F									SV2F								
	Coverage rate 95%					Avg. block size				Coverage rate 95%					Avg. block size			
n	CLT1	CLT2	Boot1	Boot2	Boot3	CLT2	Boot1	Boot2	Boot3	CLT1	CLT2	Boot1	Boot2	Boot3	CLT2	Boot1	Boot2	Boot3
$\xi^2 = 0.0001$																		
195	91.54	94.80	94.62	94.05	96.63	11.5	11.5	11.5	11.4	88.58	93.47	93.42	92.39	96.16	11.8	11.8	11.6	11.6
390	92.44	93.98	93.88	94.68	95.61	20.1	20.1	20.0	19.6	89.90	92.90	92.90	93.49	95.26	21.8	21.9	20.5	20.5
780	93.38	93.04	93.02	95.31	94.74	30.7	31.1	32.1	30.3	91.75	92.78	92.58	94.41	94.83	34.8	34.7	32.5	32.2
1560	93.83	93.20	93.34	95.25	94.82	60.8	61.6	59.3	55.9	92.78	94.59	94.38	95.52	96.03	74.5	71.4	64.0	64.3
4680	94.39	93.91	93.85	95.34	94.80	137.9	138.1	132.3	126.9	93.87	95.21	94.99	95.99	96.16	171.7	159.3	142.8	143.6
7800	94.60	93.66	93.71	95.37	94.35	165.5	166.3	172.9	158.0	94.24	95.18	94.86	96.04	96.09	201.1	189.1	171.3	172.6
11700	94.62	93.97	93.78	95.27	94.29	218.2	217.1	219.4	207.9	94.23	95.25	94.65	95.89	95.69	257.3	240.7	213.8	217.3
23400	94.67	94.10	94.31	95.04	94.60	323.4	325.2	330.5	304.6	94.30	95.08	94.74	95.82	95.72	382.0	346.6	319.5	318.8
$\xi^2 = 0.001$																		
195	90.91	95.00	95.04	93.77	96.61	11.5	11.5	11.5	11.3	88.27	93.90	93.77	91.93	96.44	11.7	11.7	11.5	11.5
390	91.79	94.08	94.04	94.11	95.78	20.1	20.2	19.8	19.5	89.66	93.12	93.02	92.85	95.54	21.7	21.7	20.3	20.4
780	93.17	93.35	93.39	94.78	94.97	30.3	30.9	32.2	29.8	91.25	92.98	92.88	94.04	95.18	34.6	34.3	32.3	32.1
1560	93.45	93.58	93.35	94.71	94.72	59.9	60.5	57.3	55.0	92.31	94.49	94.23	94.77	95.87	73.1	70.6	62.9	62.9
4680	94.32	94.04	94.10	95.24	94.95	135.1	134.6	129.2	122.6	93.39	95.10	94.93	95.73	95.90	167.3	156.9	139.5	141.0
7800	94.38	93.91	94.08	95.06	94.68	161.0	162.8	170.2	154.1	93.96	94.94	94.81	95.64	95.84	197.9	187.2	167.7	169.1
11700	94.09	93.85	93.71	94.76	94.27	213.9	214.5	217.4	203.1	93.81	94.58	94.37	95.44	95.27	251.7	236.2	210.7	214.0
23400	94.45	94.29	94.16	94.85	94.68	318.2	321.5	328.5	299.5	93.98	94.93	94.55	95.28	95.28	372.9	343.2	316.6	316.6
$\xi^2 = 0.01$																		
195	84.80	93.94	93.99	88.19	96.17	11.5	11.5	11.4	11.3	84.13	95.12	94.96	88.30	96.90	11.5	11.6	11.4	11.3
390	82.88	92.13	91.97	86.25	94.09	19.8	20.0	19.9	19.3	81.78	92.72	92.55	85.81	94.85	20.2	20.4	19.8	19.6
780	86.94	91.69	91.56	89.24	93.40	30.2	30.4	31.9	29.2	85.52	91.36	91.28	88.57	93.38	31.4	31.5	31.8	30.3
1560	87.72	91.28	91.23	89.86	92.70	56.7	57.4	55.0	50.7	87.72	92.10	92.06	90.17	93.59	60.8	59.9	57.1	54.4
4680	90.61	92.28	92.27	91.95	93.48	128.8	129.7	122.4	113.3	90.94	92.99	92.86	92.40	93.88	135.8	133.6	126.5	121.2
7800	91.69	92.59	92.52	92.59	93.20	151.9	156.5	164.3	144.6	91.77	92.86	92.87	92.79	93.63	165.2	164.6	165.2	152.0
11700	92.22	92.65	92.79	92.93	93.48	210.2	212.6	216.6	202.2	92.03	92.85	92.73	92.88	93.63	219.6	217.5	213.1	203.3
23400	92.97	93.61	93.38	93.59	93.77	303.6	307.1	318.4	281.4	92.92	93.71	93.69	93.62	94.06	319.8	315.7	315.0	288.0

Notes: CLT1 and CLT2-intervals based on the Normal using equations (8) and (15), respectively; Boot1, Boot2 and Boot3-intervals based on the wild blocks of blocks bootstrap. Boot1 is for bootstrap percentile interval using equation (22), whereas Boot2 and Boot3 are for percentile-t intervals given in equations (23) and (24), respectively. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 2. Coverage rates of nominal 95% intervals under i.i.d. noise with $\theta = 1$

	SV1F									SV2F								
	Coverage rate 95%					Avg. block size				Coverage rate 95%					Avg. block size			
n	CLT1	CLT2	Boot1	Boot2	Boot3	CLT2	Boot1	Boot2	Boot3	CLT1	CLT2	Boot1	Boot2	Boot3	CLT2	Boot1	Boot2	Boot3
$\xi^2 = 0.0001$																		
195	89.41	87.17	86.37	94.87	93.42	36.0	35.6	33.3	33.4	83.45	84.74	83.80	91.05	91.87	36.4	36.1	34.8	34.8
390	91.01	90.39	89.87	94.96	94.50	66.9	65.4	55.4	60.7	86.07	88.30	87.81	91.89	93.37	68.7	66.9	61.3	63.1
780	91.96	88.66	88.16	95.76	94.08	138.5	131.4	102.8	119.0	87.99	89.27	88.67	93.75	94.53	141.1	135.1	121.6	127.7
1560	92.97	91.07	90.62	96.52	95.04	293.1	275.1	233.5	254.2	89.57	92.48	91.91	95.09	96.32	296.8	286.8	277.6	276.6
4680	93.95	92.94	92.42	96.46	95.44	617.4	567.7	444.8	501.4	91.68	94.88	94.08	95.55	96.57	625.1	585.5	549.1	537.4
7800	94.38	93.10	92.96	96.35	95.19	806.4	751.9	577.5	689.7	92.42	95.15	94.21	95.76	96.43	820.0	745.0	665.5	673.4
11700	94.52	93.46	93.22	96.15	95.10	976.5	897.4	677.1	823.1	92.93	95.18	94.32	95.94	96.26	993.4	870.7	782.4	799.2
23400	94.64	93.86	93.68	96.10	95.21	1518.9	1336.4	987.1	1220.7	93.62	95.70	94.89	96.24	96.30	1530.0	1277.3	1137.8	1161.4
$\xi^2 = 0.001$																		
195	89.35	87.26	86.45	94.92	93.40	36.0	35.6	33.3	33.4	83.32	84.73	83.66	91.02	92.10	36.5	36.1	34.7	34.8
390	90.96	90.55	90.02	95.04	94.52	67.0	65.4	55.2	60.8	86.18	88.42	87.90	91.72	93.18	68.6	67.0	61.3	63.0
780	91.92	88.58	88.27	95.61	93.87	138.3	131.1	102.6	118.7	88.03	89.26	88.61	93.69	94.47	140.9	134.8	121.4	127.2
1560	93.07	91.12	90.62	96.65	95.05	292.8	274.4	231.7	254.2	89.53	92.55	92.07	95.00	96.22	297.3	286.4	276.5	275.4
4680	93.97	92.98	92.58	96.29	95.35	615.8	565.1	443.7	500.5	91.67	94.70	93.96	95.64	96.53	625.9	586.4	545.6	535.5
7800	94.20	93.17	92.97	96.28	95.13	807.0	753.5	575.0	687.6	92.46	94.93	94.40	95.80	96.40	820.3	747.2	666.7	674.8
11700	94.35	93.23	93.46	96.05	95.30	979.3	894.2	672.5	821.0	92.63	95.08	94.16	95.87	96.37	992.5	873.4	784.0	795.7
23400	94.74	93.87	93.68	96.12	95.07	1519.1	1335.5	980.4	1220.1	93.55	95.56	94.62	96.16	96.31	1533.8	1279.0	1143.2	1161.2
$\xi^2 = 0.01$																		
195	88.92	87.45	86.79	94.29	93.83	35.9	35.6	33.2	33.2	83.32	85.17	84.28	90.51	91.97	36.3	35.9	34.5	34.4
390	90.23	90.67	89.98	94.60	94.97	66.8	65.1	55.2	60.4	85.96	88.66	87.97	91.50	93.19	68.4	66.5	60.1	62.1
780	92.11	89.04	88.76	95.65	94.20	138.1	130.5	101.1	117.7	88.06	89.01	88.73	93.28	94.03	140.3	133.9	117.8	125.3
1560	92.96	91.66	91.30	96.45	95.63	291.6	274.3	226.9	251.4	89.69	92.68	92.13	94.62	95.99	295.8	283.9	268.2	267.7
4680	93.62	93.13	92.86	96.30	95.22	615.6	565.7	433.7	491.4	91.62	94.72	94.07	95.38	96.12	622.6	580.7	534.5	522.0
7800	93.90	93.44	93.29	95.97	95.19	805.1	747.7	570.5	686.4	92.40	94.93	94.23	95.69	96.30	818.6	745.0	650.0	666.3
11700	94.26	93.93	93.45	95.92	95.22	980.2	891.6	669.4	816.2	93.01	94.94	94.48	95.85	96.11	995.7	877.7	770.5	794.7
23400	94.63	93.83	93.60	96.10	95.32	1523.4	1343.9	987.9	1220.5	93.75	95.77	94.91	95.97	96.16	1533.6	1286.4	1124.1	1158.1

Notes: CLT1 and CLT2-intervals based on the Normal using equations (8) and (15), respectively; Boot1, Boot2 and Boot3-intervals based on the wild blocks of blocks bootstrap. Boot1 is for bootstrap percentile interval using equation (22), whereas Boot2 and Boot3 are for percentile-t intervals given in equations (23) and (24), respectively. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 3. Coverage rates of nominal 95% intervals under correlated noise based on $\hat{Bias} = \frac{\psi_1}{2n\theta^2\psi_2}\sum_{i=1}^{n} r_i^2$

n	$\lambda=-0.3$ CLT1	CLT2	Boot1	Boot2	Boot3	$\lambda=-0.5$ CLT1	CLT2	Boot1	Boot2	Boot3	$\lambda=-0.9$ CLT1	CLT2	Boot1	Boot2	Boot3
$\xi^2=0.0001$															
195	83.37	84.76	83.72	90.96	91.91	83.29	84.64	83.75	90.99	91.86	83.34	84.64	83.79	90.92	91.98
390	85.88	88.17	87.73	91.79	93.34	85.88	88.13	87.66	91.66	93.36	85.79	88.07	87.50	91.60	93.39
780	87.87	89.15	88.52	93.66	94.46	87.83	89.08	88.48	93.65	94.41	87.83	89.10	88.42	93.56	94.36
1560	89.51	92.49	91.72	95.00	96.30	89.46	92.45	91.51	94.92	96.26	89.43	92.42	91.53	94.91	96.25
4680	91.51	94.79	94.14	95.47	96.49	91.44	94.77	94.11	95.41	96.46	91.39	94.77	94.06	95.40	96.42
7800	92.24	94.97	94.06	95.67	96.33	92.18	94.94	94.08	95.71	96.31	92.14	94.80	94.10	95.61	96.31
11700	92.75	95.14	94.37	95.78	96.18	92.69	95.07	94.24	95.73	96.13	92.62	95.03	94.39	95.67	96.12
23400	93.44	95.50	94.80	96.09	96.27	93.40	95.49	94.70	96.04	96.16	93.35	95.45	94.57	95.91	96.19
$\xi^2=0.001$															
195	82.24	83.99	83.10	90.36	91.59	81.89	83.67	82.91	89.93	91.30	81.71	83.55	82.74	89.73	91.41
390	84.90	87.72	87.03	91.09	92.81	84.34	87.17	86.50	90.64	92.52	83.92	86.86	86.12	90.33	92.24
780	86.71	88.40	87.87	92.81	93.78	86.13	87.97	87.39	92.28	93.57	85.64	87.87	87.17	92.00	93.49
1560	88.18	91.83	90.99	94.22	95.85	87.45	91.27	90.63	93.67	95.52	86.89	90.78	90.12	93.43	95.30
4680	90.17	93.85	92.94	94.42	95.70	89.30	93.31	92.42	93.86	95.40	88.59	93.02	92.16	93.56	95.36
7800	90.67	93.85	93.01	94.70	95.66	89.93	93.30	92.40	93.94	95.11	89.35	92.63	91.96	93.33	94.87
11700	90.97	93.96	93.23	94.44	95.09	90.13	93.46	92.72	93.83	94.67	89.44	93.09	92.29	93.29	94.15
23400	91.79	94.36	93.27	94.73	95.14	90.45	93.77	92.62	93.96	94.56	89.42	93.10	91.97	93.20	93.94
$\xi^2=0.01$															
195	72.98	77.21	76.32	83.33	87.04	67.75	73.53	72.52	79.85	84.59	63.04	70.37	69.19	76.69	82.82
390	73.11	79.89	78.90	82.29	87.39	65.89	74.97	73.71	76.94	83.57	60.59	71.29	69.73	72.02	80.75
780	72.86	78.47	77.64	83.02	87.29	64.07	72.33	71.13	76.49	83.10	57.45	67.25	65.98	71.45	79.97
1560	72.18	81.18	80.03	82.87	89.46	61.43	74.85	73.28	74.85	84.58	52.59	68.41	66.44	68.29	80.55
4680	67.48	79.40	78.15	77.94	85.28	51.82	69.41	68.09	64.85	77.46	40.04	60.58	58.69	54.39	70.15
7800	63.85	76.32	74.71	73.86	81.50	45.94	63.92	61.25	57.49	70.07	32.84	52.84	50.79	44.97	60.77
11700	60.28	74.01	71.11	70.08	77.79	40.16	58.24	54.76	50.89	64.43	26.34	45.75	42.75	36.98	52.86
23400	52.45	67.76	63.98	61.37	71.08	28.97	48.05	43.60	38.58	52.35	15.55	32.86	29.88	23.83	38.94

Notes: $\hat{Bias} = \frac{\psi_1}{2n\theta^2\psi_2}\sum_{i=1}^{n} r_i^2$ is a consistent estimator of the bias term in PRV_n under uncorrelated noise. CLT1 and CLT2-intervals based on the Normal using equations (8) and (15), respectively; Boot1, Boot2 and Boot3-intervals based on the wild blocks of blocks bootstrap. Boot1 is for bootstrap percentile interval using equation (22), whereas Boot2 and Boot3 are for percentile-t intervals given in equations (23) and (24), respectively. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 4. Coverage rates of nominal 95% intervals under correlated noise noise based on $\hat{Bias} = \frac{\psi_1}{\theta^2\psi_2}\rho_n^2$

	$\lambda = -0.3$					$\lambda = -0.5$					$\lambda = -0.9$				
n	CLT1	CLT2	Boot1	Boot2	Boot3	CLT1	CLT2	Boot1	Boot2	Boot3	CLT1	CLT2	Boot1	Boot2	Boot3
$\xi^2 = 0.0001$															
195	84.71	85.36	84.48	91.77	92.48	84.70	85.39	84.41	91.76	92.48	84.73	85.42	84.48	91.80	92.34
390	87.15	89.03	88.42	92.54	93.81	87.12	88.99	88.46	92.49	93.79	87.16	88.91	88.39	92.34	93.82
780	88.48	89.55	88.94	94.07	94.66	88.40	89.53	88.95	94.11	94.61	88.48	89.52	89.00	94.12	94.68
1560	89.90	92.66	92.05	95.31	96.57	89.95	92.74	91.92	95.33	96.55	89.99	92.72	91.88	95.38	96.55
4680	91.86	94.92	94.27	95.74	96.57	91.90	94.91	94.32	95.60	96.59	91.92	94.97	94.36	95.69	96.58
7800	92.48	95.07	94.21	95.91	96.46	92.52	95.07	94.26	95.92	96.47	92.45	95.12	94.30	95.87	96.53
11700	92.91	95.26	94.59	95.89	96.36	92.93	95.29	94.55	95.95	96.32	92.93	95.27	94.65	96.00	96.30
23400	93.59	95.66	95.01	96.28	96.36	93.62	95.67	94.99	96.31	96.33	93.65	95.63	94.89	96.20	96.32
$\xi^2 = 0.001$															
195	84.71	85.64	84.62	91.78	92.41	85.00	85.59	84.63	91.83	92.46	85.09	85.60	84.61	91.83	92.49
390	87.16	89.13	88.57	92.51	93.62	87.22	89.16	88.58	92.57	93.68	87.22	89.12	88.57	92.36	93.68
780	88.69	89.77	89.10	94.14	94.68	88.63	89.75	89.16	94.22	94.63	88.66	89.72	89.21	94.28	94.67
1560	89.85	92.79	92.28	95.23	96.39	89.88	92.84	92.26	95.37	96.45	90.00	92.74	91.95	95.43	96.47
4680	91.81	95.00	94.22	95.67	96.50	91.95	95.03	94.21	95.70	96.53	91.95	95.04	94.36	95.72	96.58
7800	92.67	94.98	94.35	95.98	96.52	92.64	95.12	94.26	95.86	96.52	92.52	95.14	94.29	95.98	96.53
11700	92.84	95.15	94.37	95.91	96.27	93.00	95.28	94.38	96.03	96.31	93.00	95.25	94.66	96.05	96.26
23400	93.72	95.68	94.80	96.33	96.27	93.73	95.72	94.90	96.33	96.48	93.75	95.63	94.93	96.23	96.34
$\xi^2 = 0.01$															
195	84.33	84.85	84.02	90.90	91.86	84.56	84.85	83.88	91.13	91.89	84.39	84.62	83.66	91.33	91.89
390	86.53	88.79	88.08	91.83	93.17	86.75	88.86	88.15	92.17	93.09	87.03	88.67	88.05	92.35	93.37
780	88.23	89.15	88.63	93.71	94.34	88.28	89.37	88.77	93.78	94.55	88.40	89.44	89.15	93.69	94.48
1560	89.49	92.41	91.90	94.64	96.15	89.51	92.52	91.82	94.67	95.97	89.85	92.53	91.98	94.82	96.29
4680	91.86	94.66	93.99	95.74	96.37	91.94	94.97	94.19	95.73	96.65	92.07	94.91	94.17	95.77	96.70
7800	92.45	94.94	94.36	95.65	96.25	92.58	95.06	94.26	95.67	96.42	92.51	95.07	94.32	95.96	96.68
11700	92.77	95.05	94.27	95.94	96.31	93.00	95.23	94.36	96.00	96.57	93.30	95.43	94.63	96.28	96.37
23400	93.59	95.68	94.82	96.08	96.19	93.77	95.63	94.70	96.12	96.42	93.54	95.69	94.81	96.32	96.43

Notes: $\hat{Bias} = \frac{\psi_1}{\theta^2\psi_2}\rho_n^2$ is a consistent estimator of the bias term in PRV_n^d under autocorrelated noise. CLT1 and CLT2-intervals based on the Normal using equations (28) and (29), respectively; Boot1, Boot2 and Boot3-intervals based on the wild blocks of blocks bootstrap. Boot1 is for bootstrap percentile interval using equation (30), whereas Boot2 and Boot3 are for percentile-t intervals given in equations (31) and (32), respectively. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 5. Summary statistics

Days	Trans	n	S	$PRV_n^d \cdot 10^3$			q^*	$PRV_n^d \cdot 10^3$
				$q=0$	$q=1$	$q=2$		$q=q^*$
GE								
3 Oct	12613	1402	9	0.903	1.113	1.121	1	1.113
4 Oct	13782	1532	9	1.705	1.734	1.735	1	1.734
5 Oct	10628	1519	7	0.721	0.722	0.723	0	0.721
6 Oct	9991	1428	7	0.688	0.742	0.858	2	0.858
7 Oct	9785	1398	7	0.686	0.687	0.688	0	0.686
10 Oct	10660	1523	7	0.720	0.830	0.951	2	0.951
11 Oct	8588	1432	6	1.498	1.499	1.499	0	1.498
12 Oct	11160	1595	7	0.727	0.727	0.729	0	0.727
13 Oct	8649	1442	6	1.499	1.499	1.499	0	1.499
14 Oct	9261	1544	6	1.556	1.556	1.556	0	1.556
17 Oct	8530	1422	6	1.498	1.499	1.499	0	1.498
18 Oct	8751	1459	6	1.507	1.582	1.584	1	1.582
19 Oct	9023	1504	6	1.545	1.644	1.645	1	1.644
20 Oct	9251	1542	6	1.556	1.557	1.557	0	1.556
21 Oct	12513	1565	8	0.833	0.941	0.942	1	0.941
24 Oct	11642	1456	8	0.791	0.839	0.840	1	0.839
25 Oct	10919	1365	8	0.775	0.776	0.776	0	0.775
26 Oct	9249	1542	6	1.556	1.557	1.557	0	1.556
27 Oct	14598	1622	9	1.776	1.778	1.779	0	1.776
28 Oct	9405	1568	6	1.557	1.633	1.699	4	1.746
31 Oct	8871	1500	6	1.559	1.667	1.669	1	1.667
MSFT								
1 Dec	2177	78	28	0.112	0.124	0.133	0	0.112
2 Dec	1520	77	20	0.079	0.087	0.088	0	0.079
3 Dec	2530	80	32	0.077	0.088	0.088	0	0.077
6 Dec	1717	79	22	0.072	0.097	0.098	1	0.097
7 Dec	1847	81	23	0.063	0.087	0.089	1	0.087
8 Dec	1473	78	19	0.061	0.083	0.084	1	0.083
9 Dec	1851	78	24	0.071	0.083	0.083	0	0.071
10 Dec	1375	77	18	0.084	0.101	0.112	1	0.101
13 Dec	1469	78	19	0.083	0.100	0.106	0	0.083
14 Dec	2558	82	32	0.074	0.090	0.091	0	0.074
15 Dec	2304	80	29	0.101	0.120	0.121	0	0.101
16 Dec	1872	79	24	0.069	0.084	0.088	0	0.069
17 Dec	3385	89	39	0.096	0.114	0.115	0	0.096
20 Dec	3827	93	42	0.174	0.351	0.366	1	0.351
21 Dec	4105	95	44	0.483	0.554	0.556	0	0.483
22 Dec	3742	92	41	0.355	0.400	0.401	0	0.355
23 Dec	3716	93	40	0.318	0.357	0.361	0	0.318
27 Dec	2010	80	26	0.071	0.098	0.113	1	0.098
28 Dec	1676	79	22	0.096	0.120	0.124	0	0.096
29 Dec	1555	78	20	0.079	0.087	0.088	0	0.079
30 Dec	1572	79	20	0.053	0.079	0.085	1	0.079
31 Dec	1887	79	24	0.069	0.080	0.081	0	0.069

"Trans" denotes the number of transactions, n is the sample size used to calculate the pre-averaged realized volatility, we have sampled every Sth transaction price, so the period over which returns are calculated for GE and MSFT are roughly 15 seconds and 5 minutes, respectively. PRV_n^d is the dependent-noise robust version of the pre-averaged realized volatility estimator, q is the order of autocorrelation, q^* is the optimal value of q selected using the decision rule proposed by Hautsch and Podolskij (2013).

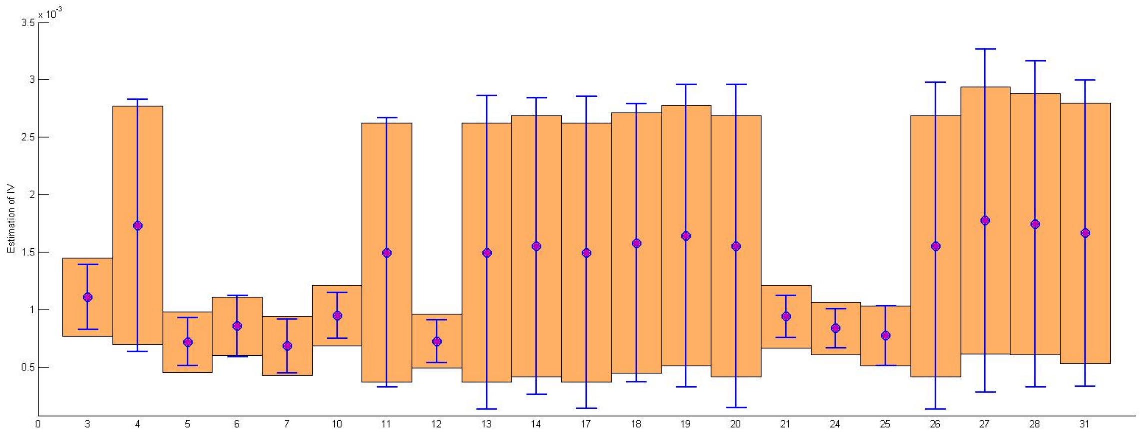

Figure 1: 95% Confidence Intervals (CI's) for the daily IV, for each regular exchange opening days for GE in October 2011, calculated using the asymptotic theory of Jacod et al. (2009) based on (28) (CI's with bars), and the wild blocks of blocks bootstrap method based on (32) (CI's with lines). The pre-averaging realized volatility estimator is the middle of all CI's by construction. Days on the x-axis.

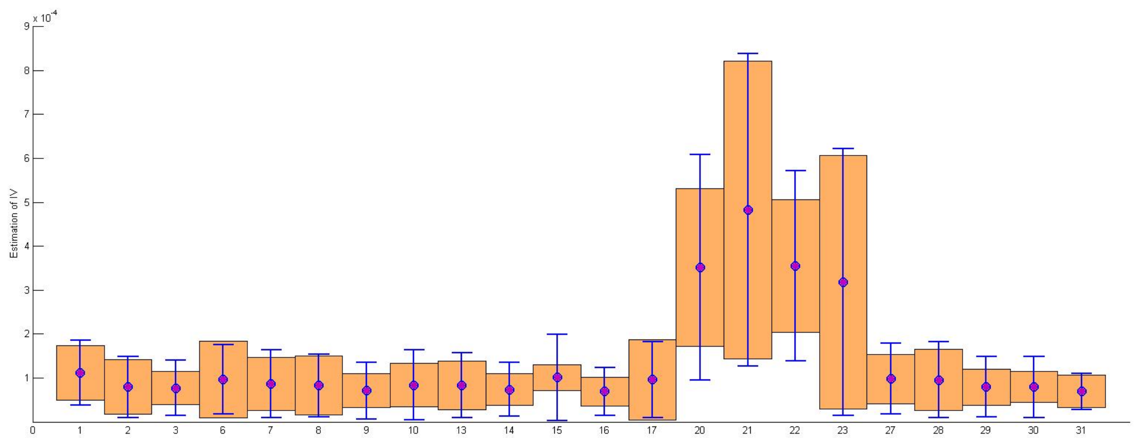

Figure 2: 95% Confidence Intervals (CI's) for the daily IV, for each regular exchange opening days for MSFT in December 2010, calculated using the asymptotic theory of Jacod et al. (2009) based on (28) (CI's with bars), and the wild blocks of blocks bootstrap method based on (32) (CI's with lines). The pre-averaging realized volatility estimator is the middle of all CI's by construction. Days on the x-axis.

Appendix B: Proofs

As in Jacod et al. (2009), we assume throughout this Appendix that the processes a, σ and X are bounded processes satisfying (1) with a and σ adapted càdlàg processes. As Jacod et al. (2009) explain, this assumption simplifies the mathematical derivations without loss of generality (by a standard localization procedure detailed in Jacod (2008)). Formally, we derive our results under the following assumption.

Assumption 4. X satisfies equation (1) with a and σ adapted càdlàg processes such that a, σ, and X are bounded processes (implying that α is also bounded).

Notation

In the following, K denotes a constant which changes from line to line. Moreover, we follow Jacod et al. (2009) and use the following additional notation. We let

$$\bar{X}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(X_{\frac{i+j}{n}} - X_{\frac{i+j-1}{n}}\right), \qquad \bar{\epsilon}_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)\left(\epsilon_{\frac{i+j}{n}} - \epsilon_{\frac{i+j-1}{n}}\right),$$

and note that $\bar{Y}_i = \bar{X}_i + \bar{\epsilon}_i$. In addition, we let

$$\begin{aligned} c_i &= \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)^2 \int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} \sigma_t^2 dt; \\ A_i &= E\left(\bar{\epsilon}_i^2 | X\right) = \sum_{j=0}^{k_n-1}\left(g\left(\frac{j+1}{k_n}\right) - g\left(\frac{j}{k_n}\right)\right)^2 \alpha_{(i+j)/n}; \text{ and} \\ \widetilde{Y}_i &= \bar{Y}_i^2 - A_i - c_i. \end{aligned}$$

Following Jacod et al. (2009), we also introduce the following random variables. For $j = 1, \ldots, J_n$, we let

$$\eta\left(p\right)_j = \frac{1}{\theta\psi_2\sqrt{n}}\zeta(p)_{(j-1)(p+1)k_n}, \text{ with } \zeta\left(p\right)_j = \sum_{i=j}^{j+(p+1)k_n-1} \widetilde{Y}_i,$$

where $p \geq 1$ is a fixed integer; $\eta\left(p\right)_j$ is the normalized sum of squared pre-averaged returns $\widetilde{Y}_i$ over a block of size $b_n = \left(p+1\right)k_n$. Note that $\eta\left(p\right)_j$ is measurable with respect to $\mathcal{F}^n_{j(p+1)k_n}$, the sigma algebra generated by all $\mathcal{F}^0_{j(p+1)k_n/n}$-measurable random variables plus all variables Y_s, with $s < j\left(p+1\right)k_n$. Finally, we let

$$\beta(p)_i = \sup_{s,t\in\left[\frac{i}{n}, \frac{i+(p+1)k_n}{n}\right]}\left(|a_s - a_t| + |\sigma_s - \sigma_t| + |\alpha_s - \alpha_t|\right), \tag{33}$$

and

$$\gamma^2(p)_t = \frac{4}{\psi_2^2}\left(\left(\Phi_{22} + \frac{1}{p+1}\Psi_{22}\right)\theta\sigma_t^4 + 2\left(\Phi_{12} + \frac{1}{p+1}\Psi_{12}\right)\frac{\sigma_t^2\alpha_t}{\theta} + \left(\Phi_{11} + \frac{1}{p+1}\Psi_{11}\right)\frac{\alpha_t^2}{\theta^3}\right). \tag{34}$$

Our bootstrap estimators depend crucially on

$$\bar{B}_j \equiv \frac{1}{b_n}\sum_{i=1}^{b_n}\bar{Y}^2_{i-1+(j-1)b_n} = \frac{1}{b_n}\sum_{i=(j-1)b_n}^{jb_n-1}\bar{Y}_i^2, \text{ for } j = 1, \ldots, J_n,$$

where $J_n = N_n/b_n$ is the number of non-overlapping blocks of size b_n out of $N_n = n - k_n + 2$ observations on pre-averaged returns.

Our first result is instrumental in proving our bootstrap results.

Lemma B.1 *Suppose Assumptions 2 and 4 hold. Then, for all integer* $p \geq 1$, *and each* $q > 0$, *we have that*

a1) $\frac{1}{\sqrt{n}} E\left(\sum_{j=1}^{J_n} \beta(p)^q_{(j-1)(p+1)k_n}\right) \to 0.$

a2) $\frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} \beta(p)^q_{(j-1)(p+1)k_n} \to^P 0.$

a3) $\frac{1}{\sqrt{n}} E\left(\sum_{j=1}^{J_n} E\left(\beta(p)^q_{(j-1)(p+1)k_n} |\mathcal{F}^n_{(j-1)(p+1)k_n}\right)\right) \to 0.$

a4) $\frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} E\left(\beta(p)^q_{(j-1)(p+1)k_n} |\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \to^P 0.$

a5) $\frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} E\left(\beta(2p+1)^q_{(j-1)(p+1)k_n} |\mathcal{F}^n_{(j-1)(p+1)k_n}\right) \to^P 0.$

a6) $\frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} \sqrt{E\left(\beta(p)^2_{(j-1)(p+1)k_n} |\mathcal{F}^n_{(j-1)(p+1)k_n}\right)} \to^P 0.$

a7) $\frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} \sqrt{E\left(\beta(2p+1)^2_{(j-1)(p+1)k_n} |\mathcal{F}^n_{(j-1)(p+1)k_n}\right)} \to^P 0.$

Proof of Lemma B.1. Part a1). Given the definition of $\beta(p)_{(j-1)(p+1)k_n}$ we can write

$$\begin{aligned}
\beta(p)_{(j-1)(p+1)k_n} &\leq \sup_{s,t \in \left[\frac{(j-1)(p+1)k_n}{n}, \frac{(j-1)(p+1)k_n+(p+1)k_n}{n}\right]} (|a_s - a_t|) \\
&\quad + \sup_{s,t \in \left[\frac{(j-1)(p+1)k_n}{n}, \frac{(j-1)(p+1)k_n+(p+1)k_n}{n}\right]} (|\sigma_s - \sigma_t|) \\
&\quad + \sup_{s,t \in \left[\frac{(j-1)(p+1)k_n}{n}, \frac{(j-1)(p+1)k_n+(p+1)k_n}{n}\right]} (|\alpha_s - \alpha_t|) \\
&\equiv \Gamma(a,p)_{(j-1)(p+1)k_n} + \Gamma(\sigma,p)_{(j-1)(p+1)k_n} + \Gamma(\alpha,p)_{(j-1)(p+1)k_n}.
\end{aligned}$$

Given that $\Gamma(a,p)_{(j-1)(p+1)k_n}$, $\Gamma(\sigma,p)_{(j-1)(p+1)k_n}$ and $\Gamma(\alpha,p)_{(j-1)(p+1)k_n}$ are strictly positive, for any $q > 0$, using the c-r inequality, we can write

$$\beta(p)^q_{(j-1)(p+1)k_n} \leq K\left(\Gamma(\sigma,p)^q_{(j-1)(p+1)k_n} + \Gamma(a,p)^q_{(j-1)(p+1)k_n} + \Gamma(\alpha,p)^q_{(j-1)(p+1)k_n}\right).$$

It follows that

$$\begin{aligned}
n^{-1/2} E\left(\sum_{j=1}^{J_n} \beta(p)^q_{(j-1)(p+1)k_n}\right) &\leq K n^{-1/2} E\left(\sum_{j=1}^{J_n} \Gamma(\sigma,p)^q_{(j-1)(p+1)k_n}\right) \\
&\quad + K n^{-1/2} E\left(\sum_{j=1}^{J_n} \Gamma(a,p)^q_{(j-1)(p+1)k_n}\right) \\
&\quad + K n^{-1/2} E\left(\sum_{j=1}^{J_n} \Gamma(\alpha,p)^q_{(j-1)(p+1)k_n}\right) = o(1),
\end{aligned}$$

where we use Lemma 5.3 of Jacod, Podolskij and Vetter (2010) to show that each of the terms above are $o(1)$ (given that a, σ and α are càdlàg bounded processes).

Proof of Lemma B.1. Part a2). Note that given the result of part a1) of Lemma B.1, it is sufficient to show that $\frac{1}{n}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^{q}\right)^2 \to 0$. By the c-r inequality,

$$\frac{1}{n}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^{q}\right)^2 \leq \frac{J_n}{n}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^{2q}\right) \leq K\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^{2q}\right),$$

which is $o\left(1\right)$ by part a1) of Lemma B.1 and given that $J_n = O\left(\sqrt{n}\right)$.

Proof of Lemma B.1. Part a3). Given the law of iterated expectations, the result follows directly from part a1) of Lemma **B.1.**

Proof of Lemma B.1. Part a4). The proof follows similarly as in part a2) of Lemma **B.1**, where we now consider the variable $E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{q}|\mathcal{F}_{(j-1)(p+1)k_n}^{n}\right)$ in place of $\beta\left(p\right)_{(j-1)(p+1)k_n}^{q}$.

Proof of Lemma B.1. Part a5). Given the definition of $\beta(p)_i$, for any $p \geq 1$, such that $b_n = \left(p+1\right)k_n$ we can write

$$\begin{aligned}\frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}E\left(\beta\left(2p+1\right)_{(j-1)b_n}^{q}|\mathcal{F}_{(j-1)b_n}^{n}\right) &= \frac{1}{\sqrt{n}}\sum_{j=1}^{\left[\frac{J_n}{2}\right]}E\left(\beta\left(2p+1\right)_{2(j-1)b_n}^{q}|\mathcal{F}_{2(j-1)b_n}^{n}\right)\\ &+\frac{1}{\sqrt{n}}\sum_{j=1}^{\left[\frac{J_n}{2}\right]}E\left(\beta\left(2p+1\right)_{(2(j-1)+1)b_n}^{q}|\mathcal{F}_{(2(j-1)+1)b_n}^{n}\right),\end{aligned}$$

which is $o_P\left(1\right)$ given part a4) of Lemma B.1.

Proof of Lemma B.1. Part a6). Here, the proof contains two steps. Step 1. We show show that $\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}|\mathcal{F}_{(j-1)(p+1)k_n}^{n}\right)}\right) \to 0$. Step 2. We show show that $\frac{1}{n}Var\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}|\mathcal{F}_{(j-1)(p+1)k_n}^{n}\right)}\right) \to 0$. Note that using the first expression in equation (5.47) of Jacod et al. (2009), the result of step 1 follows directly. Given this result, to show step 2, it is sufficient to show that $\frac{1}{n}E\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}|\mathcal{F}_{(j-1)(p+1)k_n}^{n}\right)}\right)^2 \to 0$. We have that

$$\begin{aligned}\frac{1}{n}\left(\sum_{j=1}^{J_n}\sqrt{E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}|\mathcal{F}_{(j-1)(p+1)k_n}^{n}\right)}\right)^2 &\leq \frac{J_n}{n}\sum_{j=1}^{J_n}E\left(E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}|\mathcal{F}_{(j-1)(p+1)k_n}^{n}\right)\right)\\ &= \frac{J_n}{n}\sum_{j=1}^{J_n}E\left(\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}\right) \leq K\frac{1}{\sqrt{n}}E\left(\sum_{j=1}^{J_n}\beta\left(p\right)_{(j-1)(p+1)k_n}^{2}\right),\end{aligned}$$

which is $o\left(1\right)$ given equation (5.47) of Jacod et al. (2009) and the fact that $J_n = O\left(\sqrt{n}\right)$ under our assumptions.

Proof of Lemma B.1. Part a7). The proof follows similarly as part a5) and therefore we omit the details.

Our next result is crucial to the proofs of Lemmas 3.1 and 4.1.

Lemma B.2 *Under Assumptions 1, 2, and 4, if* $b_n = (p+1)\, k_n$ *where* $p \geq 1$ *is fixed, then*

$$\frac{\sqrt{n} b_n^2}{k_n^2 \psi_2^2} \sum_{j=1}^{J_n} \bar{B}_j^2 \to^P V_p + \theta\,(p+1) \int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_s \right)^2 ds.$$

Proof of Lemma B.2. Given the definition of $\bar{B}_j$, we have that

$$\bar{B}_j = \frac{1}{b_n} \sum_{i=(j-1)b_n}^{jb_n - 1} \bar{Y}_i^2 = \frac{1}{b_n} \sum_{i=(j-1)b_n}^{jb_n - 1} \underbrace{\left(\bar{Y}_i^2 - A_i - c_i \right)}_{\equiv \tilde{Y}_i} + \frac{1}{b_n} \sum_{i=(j-1)b_n}^{jb_n - 1} (A_i + c_i)$$

where $A_i \equiv E\left(\bar{\epsilon}_i^2 | X\right)$ and $c_i = \sum_{j=1}^{k_n} g\left(\frac{j}{k_n}\right)^2 \int_{\frac{i+j-1}{n}}^{\frac{i+j}{n}} \sigma_t^2 dt$. It follows that

$$\frac{\sqrt{n} b_n^2}{k_n^2 \psi_2^2} \sum_{j=1}^{J_n} \bar{B}_j^2 = \mathcal{B}_{1n} + \mathcal{B}_{2n} + \mathcal{B}_{3n},$$

where

$$\begin{aligned}
\mathcal{B}_{1n} &\equiv \sqrt{n} \sum_{j=1}^{J_n} \left(\frac{1}{\theta \psi_2 \sqrt{n}} \sum_{i=(j-1)b_n}^{jb_n - 1} \widetilde{Y}_i \right)^2 = \sqrt{n} \sum_{j=1}^{J_n} \eta\,(p)_j^2, \\
\mathcal{B}_{2n} &\equiv \frac{2}{\theta \psi_2} \sum_{j=1}^{J_n} \eta\,(p)_j \sum_{i=(j-1)b_n}^{jb_n - 1} (A_i + c_i)\,; \text{ and} \\
\mathcal{B}_{3n} &\equiv \frac{1}{\theta^2 \psi_2^2 \sqrt{n}} \sum_{j=1}^{J_n} \left(\sum_{i=(j-1)b_n}^{jb_n - 1} (A_i + c_i) \right)^2.
\end{aligned}$$

We show that (1) $\mathcal{B}_{1n} \to^P \int_0^1 \gamma_t^2(p)\, dt$; (2) $\mathcal{B}_{2n} \to^P 0$, and that (3) $\mathcal{B}_{3n} \to^P (p+1)\,\theta \int_0^1 \left(\sigma_t^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_t \right)^2 dt$. Starting with (1), write

$$\begin{aligned}
\sqrt{n} \sum_{j=1}^{J_n} \eta\,(p)_j^2 - \int_0^1 \gamma_t^2(p)\, dt &= \mathcal{B}_{1.1n} + \mathcal{B}_{1.2n} + \mathcal{B}_{1.3n}, \quad \text{with} \\
\mathcal{B}_{1.1n} &= \sqrt{n} \sum_{j=1}^{J_n} \left(\eta\,(p)_j^2 - E\left(\eta\,(p)_j^2 \,|\mathcal{F}^n_{(j-1)(p+1)k_n} \right) \right), \\
\mathcal{B}_{1.2n} &= \sqrt{n} \sum_{j=1}^{J_n} E\left(\eta\,(p)_j^2 \,|\mathcal{F}^n_{(j-1)(p+1)k_n} \right) - \frac{N_n}{n} \frac{1}{J_n} \sum_{j=1}^{J_n} \gamma(p)^2_{\frac{j-1}{J_n}}, \\
\mathcal{B}_{1.3n} &= \frac{N_n}{n} \frac{1}{J_n} \sum_{j=1}^{J_n} \gamma(p)^2_{\frac{j-1}{J_n}} - \int_0^1 \gamma_t^2(p)\, dt.
\end{aligned}$$

We show that each of $\mathcal{B}_{1.\ell n} \to^P 0$ for $\ell = 1, 2, 3$. For $\ell = 1$, by Lenglart's inequality (see e.g. Lemma 4.4 of Vetter (2008)), it is sufficient to show that $n \sum_{j=1}^{J_n} E\left(\eta\,(p)_j^4 \,|\mathcal{F}^n_{(j-1)(p+1)k_n} \right) \to^P 0$, which follows immediately by using equation (5.57) of Jacod et al. (2009). Next, to show that $\mathcal{B}_{1.2n} \to^P 0$, note that

$$\begin{aligned}
\mathcal{B}_{1.2n} &\leq \sum_{j=1}^{J_n} \left| \sqrt{n} E\left(\eta\left(p\right)_j^2 | \mathcal{F}^n_{(j-1)(p+1)k_n} \right) - \frac{N_n}{n} \frac{1}{J_n} \gamma(p)^2_{\frac{j-1}{J_n}} \right| \\
&= \sum_{j=1}^{J_n} \left| \sqrt{n} E\left(\frac{1}{\theta^2 \psi_2^2 n} \zeta^2(p)_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n} \right) - \frac{1}{n}\left(p+1\right) \theta \sqrt{n} \gamma(p)^2_{\frac{j-1}{J_n}} \right| \\
&= \frac{\sqrt{n}}{\theta^2 \psi_2^2 n} \sum_{j=1}^{J_n} \left| E\left(\zeta^2(p)_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n} \right) - \theta^3 \psi_2^2 \left(p+1\right) \gamma(p)^2_{\frac{j-1}{J_n}} \right| \\
&\leq \frac{K}{\theta^2 \psi_2^2 \sqrt{n}} \sum_{j=1}^{J_n} \chi(p)_{(j-1)(p+1)k_n}
\end{aligned}$$

where we use the fact that $N_n/J_n = (p+1)\,k_n$ with $k_n = \theta\sqrt{n}$ and rely on equation (5.41) of Jacod et al. (2009) to bound the term in absolute value, where

$$\chi(p)_{(j-1)(p+1)k_n} = n^{-1/4} + \sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n} \right)}$$

and $\beta\left(p\right)_i$ is as defined in (33). It follows that

$$\frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} \chi(p)_{(j-1)(p+1)k_n} \leq \frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} n^{-1/4} + \frac{1}{\sqrt{n}} \sum_{j=1}^{J_n} \sqrt{E\left(\beta\left(p\right)^2_{(j-1)(p+1)k_n} | \mathcal{F}^n_{(j-1)(p+1)k_n} \right)} \to^P 0,$$

where the first term is of order $O\left(n^{-1/4}\right)$ and the second term is $o_P\left(1\right)$ given part a6) of Lemma B.1. Finally, $\mathcal{B}_{1.3n} \to^P 0$ follows immediately by Riemann's integrability of σ, the fact that $\frac{N_n}{n} \to 1$ and $J_n \to \infty$ as $n \to \infty$.

To show (2), let $\varphi_j \equiv \sum\limits_{i=(j-1)b_n}^{jb_n-1} \left(A_i + c_i\right)$ and $\zeta\left(X,p\right)_j = \sum\limits_{i=(j-1)b_n}^{jb_n-1} \left(\bar{X}_i^2 - c_i\right)$. We can write

$$\begin{aligned}
\mathcal{B}_{2n} &= \frac{2}{\theta\psi_2} \sum_{j=1}^{J_n} \varphi_j \cdot \eta\left(p\right)_j = \mathcal{B}_{2.1n} + \mathcal{B}_{2.2n}, \quad \text{with} \\
\mathcal{B}_{2.1n} &= \frac{2}{\theta\psi_2} \sum_{j=1}^{J_n} \left(\varphi_j \eta\left(p\right)_j - E\left(\varphi_j \eta\left(p\right)_j | \mathcal{F}^n_{(j-1)(p+1)k_n} \right) \right), \quad \text{and} \\
\mathcal{B}_{2.2n} &= \frac{2}{\theta\psi_2} \sum_{j=1}^{J_n} E\left(\varphi_j \eta\left(p\right)_j | \mathcal{F}^n_{(j-1)(p+1)k_n} \right).
\end{aligned}$$

We show that each of $\mathcal{B}_{2.\ell n} \to^P 0$ for $\ell = 1, 2$. Note that given the definitions of A_i, c_i, and the fact that $k_n = \theta\sqrt{n}$, Assumption 4 implies that $A_i + c_i \leq K/\sqrt{n}$ uniformly in i. Given that $b_n = (p+1)\,k_n$, it follows that $\varphi_j \leq K$ uniformly in j. Starting with $\ell = 1$, by Lenglart's inequality, it is sufficient to show that $\sum_{j=1}^{J_n} E\left(\varphi_j^2 \eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) \to^P 0$. We can write

$$\begin{aligned}
\sum_{j=1}^{J_n} E\left(\varphi_j^2 \eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) &\leq K\sum_{j=1}^{J_n} E\left(\eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) \\
&= K\Bigg(\frac{1}{\sqrt{n}}\Bigg(\sqrt{n}\sum_{j=1}^{J_n} E\left(\eta\left(p\right)_j^2 | \mathcal{F}_{(j-1)(p+1)k_n}^n\right) - \frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\gamma(p)^2_{\frac{j-1}{J_n}}\Bigg)\Bigg) \\
&\quad +K\Bigg(\frac{1}{\sqrt{n}}\left(\frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\gamma(p)^2_{\frac{j-1}{J_n}} - \int_0^1 \gamma_t^2\left(p\right)dt\right) + \frac{1}{\sqrt{n}}\int_0^1 \gamma_t^2\left(p\right)dt\Bigg) \\
&\equiv K\Bigg(\frac{1}{\sqrt{n}}\mathcal{B}_{1.2n} + \frac{1}{\sqrt{n}}\mathcal{B}_{1.3n} + \frac{1}{\sqrt{n}}\int_0^1 \gamma_t^2\left(p\right)dt\Bigg) \\
&= \frac{1}{\sqrt{n}}o_P\left(1\right) + \frac{1}{\sqrt{n}}o_P\left(1\right) + O_P\left(\frac{1}{\sqrt{n}}\right) = o_P\left(1\right),
\end{aligned}$$

where in particular we use the fact that $\mathcal{B}_{1.2n}=o_P\left(1\right)$ and $\mathcal{B}_{1.3n} = o_P\left(1\right)$, and $\int_0^1 \gamma_t^2\left(p\right)dt = O_P\left(1\right)$. It follows that $\mathcal{B}_{2.1n} \to^P 0$. Next, to show that $\mathcal{B}_{2.2n} \to^P 0$, note that we can write

$$\mathcal{B}_{2.2n} \leq \frac{2K}{\theta\psi_2}\frac{1}{n^{1/4}}\left(n^{1/4}\sum_{j=1}^{J_n} E\left(\eta\left(p\right)_j | \mathcal{F}_{(j-1)(p+1)k_n}^n\right)\right) = O_P\left(n^{-1/4}\right) o_P\left(1\right) = o_P\left(1\right),$$

given that $\varphi_j \leq K$, and given equation (5.49) of Jacod et al. (2009).

Finally, to show (3), note that given the definitions of A_i and c_i, and by using equations (5.23) and (5.36) of Jacod et al. (2009), we can write

$$\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i + c_i\right) = \sum_{i=(j-1)b_n}^{jb_n-1}\left(\frac{\psi_1}{\theta\sqrt{n}}\alpha_{(j-1)b_n/n} + \frac{\theta\psi_2}{\sqrt{n}}\sigma^2_{(j-1)b_n/n}\right) + O\left(\frac{p}{\sqrt{n}} + p\beta(p)_{(j-1)b_n}\right). \tag{35}$$

It follows that

$$\mathcal{B}_{3n} \equiv \frac{1}{\theta^2\psi_2^2\sqrt{n}}\sum_{j=1}^{J_n}\left(\sum_{i=(j-1)b_n}^{jb_n-1}\left(A_i + c_i\right)\right)^2 = L_n + R_n,$$

where the leading term is

$$L_n = \left(p+1\right)\theta\frac{N_n}{n}\frac{1}{J_n}\sum_{j=1}^{J_n}\left(\frac{\psi_1}{\theta^2\psi_2}\alpha_{(j-1)b_n/n} + \sigma^2_{(j-1)b_n/n}\right)^2 \to^P (p+1)\theta\int_0^1\left(\sigma_t^2 + \frac{\psi_1}{\theta^2\psi_2}\alpha_t\right)^2 dt. \tag{36}$$

The remainder is such that

$$R_n = K\cdot O_P\left(\frac{1}{\sqrt{n}} + \frac{1}{\sqrt{n}}\sum_{j=1}^{J_n}\beta(p)^2_{(j-1)b_n}\right) \to^P 0$$

by using Lemma (5.4) of Jacod et al. (2009).

Proof of Lemma 3.1. Part a) Given the definition of V_n^*, we can write

$$V_n^* = V_{1n}^* - \frac{\sqrt{n} N_n b_n}{(N_n - b_n + 1)^2} V_{2n}^*,$$

where

$$\begin{aligned} V_{1n}^* &= \frac{1}{b_n} \sum_{t=0}^{b_n - 1} v_{1n,t}^*, \text{ with } v_{1n,t}^* \equiv \frac{\sqrt{n}}{(N_n - b_n + 1) N_n} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \left(\sum_{i=t+1}^{b_n + t} Z_{i+(j-1)b_n} \right)^2, \text{ and} \\ V_{2n}^* &= \frac{1}{b_n} \sum_{t=0}^{b_n - 1} v_{2n,t}^{*2}, \text{ with } v_{2n,t}^* \equiv \frac{1}{N_n} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \sum_{i=t+1}^{b_n + t} Z_{i+(j-1)b_n}. \end{aligned}$$

We now proceed in two steps. In Step 1, we show that $v_{1n,t}^* \to^P V_p + \theta (p+1) \int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_s \right)^2 ds$ uniformly in t. In Step 2, we show that $v_{2n,t}^{*2} \to^P \left(\int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_s \right) ds \right)^2$, also uniformly in t. This together with the fact that $\frac{\sqrt{n} N_n b_n}{(N_n - b_n + 1)^2} \to (p+1)\theta$ as $n \to \infty$ when $b_n = (p+1) k_n$ and k_n satisfies Assumption 2 imply the result. Proof of Step 1. For $t = 0, \ldots, b_n - 1$ and $j = 1, \ldots, \left[\frac{N_n - t}{b_n}\right]$, let

$$\bar{B}_{j,t} \equiv \frac{1}{b_n} \sum_{i=1}^{b_n} \bar{Y}_{i-1+t+(j-1)b_n}^2 = \frac{k_n \psi_2}{N_n} \frac{1}{b_n} \sum_{i=1}^{b_n} Z_{i+t+(j-1)b_n},$$

where $Z_i \equiv \frac{N_n}{k_n} \frac{1}{\psi_2} \bar{Y}_{i-1}^2$ and note that the $\bar{B}_{j,t}$ are averages of non-overlapping blocks for given t. With this notation, we have that

$$v_{1n,t}^* = \frac{N_n^2}{(N_n - b_n + 1) N_n} \frac{\sqrt{n} b_n^2}{k_n^2 \psi_2^2} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \bar{B}_{j,t}^2,$$

where we can show that $\frac{N_n^2}{(N_n - b_n + 1) N_n} \to 1$ under the condition that $b_n = (p+1) k_n$. Using arguments similar to those used to prove Lemma B.2, we can show that

$$\frac{\sqrt{n} b_n^2}{k_n^2 \psi_2^2} \sum_{j=1}^{\left[\frac{N_n - t}{b_n}\right]} \bar{B}_{j,t}^2 \to^P V_p + \theta (p+1) \int_0^1 \left(\sigma_s^2 + \frac{\psi_1}{\theta^2 \psi_2} \alpha_s \right)^2 ds$$

uniformly in t. The proof of Step 2 relies on the consistency result in Theorem 1 of Christensen, Kinnebrock and Podolskij (2010). Indeed $v_{2n,t}^*$ is the main term in Jacod et al. (2009) pre-averaged realized volatility estimator without the bias corrected term, with starting point t. **Part b).** Follows directly from part a) of Lemma 3.1 when replacing σ_t by a constant for all t. **Part c).** Follows directly from part a) of Lemma 3.1.

Proof of Lemma 4.1. Given the definition of V_n^*, we can write

$$V_n^* = Var^* \left(n^{1/4} \widetilde{PRV}_n^* \right) = \frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2} \sum_{j=1}^{J_n - 1} \left(\bar{B}_j - \bar{B}_{j+1} \right)^2 Var^* \left(\eta_j^2 \right).$$

Let,

$$\Xi_j = \frac{b_n}{\psi_2 k_n} \bar{B}_j,$$

by adding and substracting appropriately and given $Var^*(\eta) = 1/2$, it follows that

$$V_n^* = \underbrace{n^{1/2}\left(\sum_{j=1}^{J_n-1} \Xi_j^2 - \sum_{j=1}^{J_n-1} \Xi_j \Xi_{j+1}\right)}_{\equiv \widetilde{L}_n} + \frac{n^{1/2}}{2}\left(\Xi_{J_n}^2 - \Xi_1^2\right). \tag{37}$$

Note that given the definition of $\bar{B}_j$ and Ξ_j we can write

$$\begin{aligned} \frac{n^{1/2}}{2}\left(\Xi_{J_n}^2 - \Xi_1^2\right) &= \frac{n^{1/2} b_n^2}{\psi_2^2 k_n^2}\left(\bar{B}_1^2 + \bar{B}_{J_n}^2\right) \\ &= O_P\left(\left(\frac{b_n}{n^{3/4}}\right)^2\right) \\ &= o_P(1), \end{aligned}$$

where the second equality follows since $\bar{B}_j = O_P(1/\sqrt{n})$ uniformly in j, and the last equality holds so long as $\delta < 3/4$, which is verify under our assumptions. Thus, given (37) the rest of the proof can be reduced to $\widetilde{L}_n \to^P V$. The proof of this claim follows closey that for Theorem 4.1 of Christensen et al. (2013), however for completeness, we present here the relevant details.

Following Christensen et al. (2013), we introduce two approximating version of Ξ_j first, namely

$$\widetilde{\mathcal{Z}}_j = \frac{1}{\psi_2 k_n} \sum_{i=1}^{b_n} \widetilde{Y}_{i-1+(j-1)b_n}^2,$$

$$\widehat{\mathcal{Z}}_j = \frac{1}{\psi_2 k_n} \sum_{i=1}^{b_n} \widetilde{Y}_{i-1+jb_n}^2,$$

where we have set $\tilde{Y}_i = \bar{\epsilon}_i + \sigma_{\frac{jb_n}{N_n}} \bar{W}_i$, with $\bar{W}_i = \sum_{t=1}^{k_n} g\left(\frac{t}{k_n}\right)\left(W_{\frac{i+t}{n}} - W_{\frac{i+t-1}{n}}\right)$, for $jb_n \leq i \leq (j+1)b_n - 1$. Indeed we will show that the error due to replacing $\bar{Y}_i$ by $\tilde{Y}_i$ is small and will not affect our theoretical results, since σ is assumed to be an Ito semimartingale itself. We have that, for $jb_n \leq i \leq (j+1)b_n - 1$

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

Note also that $E\left(|\mathcal{Z}_j|\right) \leq K\frac{b_n}{n}$, thus it follows that

$$
\begin{aligned}
E\left(\left|\mathcal{Z}_j - \widetilde{\mathcal{Z}}_j\right|\right) &\leq Kb_n\left(\frac{(k_n b_n)^{1/2}}{n}\left(\frac{1}{\sqrt{k_n}}\right)^{-\frac{1}{2}}\right) \\
&\leq K\left(\frac{b_n}{n}\right)^{3/2},
\end{aligned}
$$

similarly for $\widehat{\mathcal{Z}}_j$, we have $E\left(\left|\mathcal{Z}_j - \widehat{\mathcal{Z}}_j\right|\right) \leq K\left(\frac{b_n}{n}\right)^{3/2}$. So by using the fact that $\delta < 2/3$ we obtain $\widetilde{L}_n - \widehat{L}_n = o_P(1)$, where

$$
\widehat{L}_n = \sqrt{n}\sum_{j=1}^{J_n-1}\left(\widehat{\mathcal{Z}}_j^2 - \widehat{\mathcal{Z}}_j\widetilde{\mathcal{Z}}_{j+1}\right).
$$

Then it is simple to deduce that

$$
\begin{aligned}
\sqrt{n}\left|\sum_{\alpha=1}^{J_n-1} E\left(\widehat{\mathcal{Z}}_j^2 - E\left(\widehat{\mathcal{Z}}_j^2|\mathcal{F}^n_{\frac{(j-1)b_n}{n}}\right)\right)\right| &\leq K\frac{b_n^{3/2}}{n}, \\
\sqrt{n}\left|\sum_{j=1}^{J_n-1} E\left(\widehat{\mathcal{Z}}_j\widetilde{\mathcal{Z}}_{j+1} - E\left(\widehat{\mathcal{Z}}_j\widetilde{\mathcal{Z}}_{j+1}|\mathcal{F}^n_{\frac{(j-1)b_n}{n}}\right)\right)\right| &\leq K\frac{b_n^{3/2}}{n},
\end{aligned}
$$

by conditional independence, and now we are left with

$$
\widehat{L}_n = \sqrt{n}\sum_{j=1}^{J_n-1} E\left(\widehat{\mathcal{Z}}_j^2 - \widehat{\mathcal{Z}}_j\widetilde{\mathcal{Z}}_{j+1}|\mathcal{F}^n_{\frac{(j-1)b_n}{n}}\right) + o_P(1).
$$

From the same arguments as in Lemma 7.3 and Lemma 7.5 of Christensen et al. (2013) plus using $\delta > 1/2$, we obtain

$$
\begin{aligned}
&\sqrt{n}E\left(\widehat{\mathcal{Z}}_j^2 - \widehat{\mathcal{Z}}_j\widetilde{\mathcal{Z}}_{j+1}|\mathcal{F}^n_{\frac{(j-1)b_n}{n}}\right) \\
&= \int_{\frac{(\alpha-1)b_n}{n}}^{\frac{\alpha b_n}{n}} \varsigma(s)\,ds + o\left(\frac{b_n}{n}\right),
\end{aligned}
$$

uniformly in j, where we use

$$
V = \int_0^1 \varsigma(s)\,ds, \text{ with } \varsigma(s) = \frac{4}{\psi_2^2}\left(\Phi_{22}\theta\sigma_s^4 + 2\Phi_{12}\frac{\sigma_s^2\alpha_s}{\theta} + \Phi_{11}\frac{\alpha_s^2}{\theta^3}\right),
$$

thus we have

$$
\widetilde{L}_n = \int_0^1 \varsigma(s)\,ds + o_P(1)
$$

and the proof is complete.

Proof of Theorem 4.1 Let $S_n^* = n^{1/4}\left(\widetilde{PRV}^*_n - E^*\left(\widetilde{PRV}^*_n\right)\right) = \frac{b_n}{\psi_2 k_n}\sum_{j=1}^{J_n} z_j^*$, where $z_j^* =$

$n^{1/4}\frac{b_n}{\psi_2 k_n}\left(\bar{B}_j{}^* - E^*\left(\bar{B}_j{}^*\right)\right)$. It follows that $E^*\left(\sum_{j=1}^{J_n} z_j^*\right) = 0$, and

$$V_n^* \equiv Var^*\left(\sum_{j=1}^{J_n} z_j^*\right) \xrightarrow{P} V.$$

Since $z_1^*, \cdots, z_{J_n}^*$ are conditionally independent, by the Berry-Esseen bound, for some small $\varepsilon > 0$ and for some constant $C > 0$ which changes from line to line,

$$\sup_{x\in\Re}\left|P^*\left(S_n^* \leq x\right) - \Phi\left(x/\sqrt{V}\right)\right| \leq C\sum_{j=1}^{J_n} E^*\left|z_j^*\right|^{2+\varepsilon},$$

which converges to zero in probability as $n \to \infty$. We have

$$\begin{aligned}
\sum_{j=1}^{J_n} E^*\left|z_j^*\right|^{2+\varepsilon} &= \sum_{j=1}^{J_n} E^*\left|n^{1/4}\frac{b_n}{\psi_2 k_n}\left(\bar{B}_j^* - E^*\left(\bar{B}_j^*\right)\right)\right|^{2+\varepsilon} \\
&\leq 2n^{\frac{(2+\varepsilon)}{4}}\left(\frac{b_n}{\psi_2 k_n}\right)^{2+\varepsilon}\sum_{j=1}^{J_n} E^*\left|\bar{B}_j^*\right|^{2+\varepsilon} \\
&\leq CE^*\left|\eta_1\right|^{2+\varepsilon} n^{\frac{(2+\varepsilon)}{4}}\left(n^{\frac{(2+\varepsilon)}{4}} k_n^{-(2+\varepsilon)} b_n^{(2+\varepsilon)}\right)\sum_{j=1}^{J_n}\left|\bar{B}_j\right|^{2+\varepsilon} \\
&\leq CE^*\left|\eta_1\right|^{2+\varepsilon}\left(n^{\frac{(2+\varepsilon)}{4}} k_n^{-(2+\varepsilon)} b_n^{(2+\varepsilon)}\right)\frac{1}{b_n^{(2+\varepsilon)}}\sum_{j=1}^{J_n}\left(b_n^{(2+\varepsilon)-1}\sum_{i=1}^{b_n}\bar{Y}_{i-1+(j-1)b_n}^{2(2+\varepsilon)}\right) \\
&\leq \underbrace{CE^*\left|\eta_1\right|^{2+\varepsilon}}_{=O(1)}\underbrace{\left(n^{\frac{(2+\varepsilon)}{4}-\frac{\varepsilon}{2}} k_n^{-(2+\varepsilon)} b_n^{(1+\varepsilon)}\right)}_{\propto n^{\frac{(2+\varepsilon)}{4}+(\delta_1-1)(1+\varepsilon)}}\underbrace{\left(n^{\frac{\varepsilon}{2}}\sum_{i=1}^{N_n}\bar{Y}_i^{2(2+\varepsilon)}\right)}_{=O_p(1)} \\
&= O_p\left(n^{\frac{(2+\varepsilon)}{4}+(\delta_1-1)(1+\varepsilon)}\right) = o_p(1),
\end{aligned}$$

since for any $\varepsilon \geq 2$, so long as $\delta < 2/3$, we have $\frac{2+\varepsilon}{4} + (\delta - 1)(1+\varepsilon) < 0$, and given that by Theorem 3.3 of Jacod, Podolskij and Vetter (2010)

$$n^{\frac{\varepsilon}{2}}\sum_{i=1}^{N_n}\bar{Y}_i^{2(2+\varepsilon)} \to^P \mu_{2(2+\varepsilon)}\int_0^1\left(\theta\psi_2\sigma_s^2 + \frac{1}{\theta}\psi_1\alpha_s\right)^{2+\varepsilon} ds,$$

which is bounded given Assumption 3, and $E^*\left|\eta_j\right|^{2+\varepsilon} \leq \Delta < \infty$. It follows that $n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)$ $N(0,V)$ in probability.

Proof of Theorem 4.2 Given that $T_n \xrightarrow{d} N(0,1)$, it suffices to show that $T_n^* \xrightarrow{d^*} N(0,1)$ in probability. Let

$$H_n^* = \frac{n^{1/4}\left(\widetilde{PRV}_n^* - E^*\left(\widetilde{PRV}_n^*\right)\right)}{\sqrt{V_n^*}},$$

and note that

$$T_n^* = H_n^* \sqrt{\frac{V_n^*}{\hat{V}_n^*}},$$

where $\hat{V}_n^*$ is defined in the main text. Theorem 4.1 proved that $H_n^* \xrightarrow{d^*} N(0,1)$ in probability. Thus, it suffices to show that $\hat{V}_n^* - V_n^* \xrightarrow{P^*} 0$ in probability. In particular, we show that (1) $Bias^*\left(\hat{V}_n^*\right) = 0$, and (2) $Var^*\left(\hat{V}_n^*\right) \xrightarrow{P} 0$. It is easy to verify that (1) holds by the definition of $\hat{V}_n^*$ and V_n^*. To prove (2), note that

$$\begin{aligned}
Var^*\left(\hat{V}_n^*\right) &= E^*\left(\hat{V}_n^* - V_n^*\right)^2 = \left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\right)^2 E^*\left(\sum_{j=1}^{J_n-1}\left(\bar{B}_j - \bar{B}_{j+1}\right)^2\left(\eta_j^2 - E^*\left(\eta^2\right)\right)\right)^2 \\
&= \left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\right)^2 E^*\left(\eta^2 - E^*\left(\eta^2\right)\right)^2 \sum_{j=1}^{J_n-1}\left(\bar{B}_j - \bar{B}_{j+1}\right)^4 \\
&\leq 2^3\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\right)^2 E^*\left(\eta^2 - E^*\left(\eta^2\right)\right)^2 \left(\sum_{j=1}^{J_n-1}\bar{B}_j^4 + \sum_{j=1}^{J_n-1}\bar{B}_{j+1}^4\right) \\
&= 2^3\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\right)^2 E^*\left(\eta^2 - E^*\left(\eta^2\right)\right)^2 \left(2\sum_{j=1}^{J_n}\bar{B}_j^4 - \left(\bar{B}_1^4 + \bar{B}_{J_n}^4\right)\right) \\
&\leq 2^3\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\right)^2 E^*\left(\eta^2 - E^*\left(\eta^2\right)\right)^2 \left(2\left(\frac{1}{b_n}\sum_{j=1}^{J_n}\left(\sum_{i=1}^{b_n}\bar{Y}_{i-1+(j-1)b_n}^8\right)\right) - \left(\bar{B}_1^4 + \bar{B}_{J_n}^4\right)\right) \\
&\leq \underbrace{2^3\left(\left(\frac{n^{1/2}b_n^2}{\psi_2^2 k_n^2}\frac{Var^*(\eta)}{E^*(\eta^2)}\right)^2 E^*\left(\eta^2 - E^*\left(\eta^2\right)\right)^2\right)}_{=O\left(\frac{b_n^4}{n}\right)}\left(\frac{1}{b_n}\frac{2}{n}\underbrace{\left(n\sum_{i=1}^{N_n}\bar{Y}_i^8\right)}_{=O_P(1)} - \underbrace{\left(\bar{B}_1^4 + \bar{B}_{J_n}^4\right)}_{=O_P\left(\frac{1}{n^2}\right)}\right) \\
&= O_P\left(\frac{b_n^3}{n^2}\right) + O_P\left(\frac{b_n^4}{n^3}\right) = o_P(1),
\end{aligned}$$

where we have used the fact that η_j is i.i.d. to justify the third equality and Theorem 3.3 of Jacod, Podolskij and Vetter (2010) to justify the fact that $n\sum_{i=1}^{N_n}\bar{Y}_i^8 = O_P(1)$.

References

[1] Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). "The Distribution of Realized Exchange Rate Volatility," *Journal of the American Statistical Association*, 96, 42-55.

[2] Bandi, F., and J. Russell (2008). "Microstructure noise, realized variance, and optimal sampling," *Review of Economic Studies*, 75(2), 339-369.

[3] Bandi, F., and J. Russell (2011). "Market microstructure noise, integrated variance estimators, and the accuracy of asymptotic approximations," *Journal of Econometrics*, 160, 145-159.

[4] Barndorff-Nielsen, O. and N. Shephard (2001). “Non-Gaussian Ornstein-Uhlenbeck- Based Models And Some Of Their Uses In Financial Economics,” *Journal of the Royal Statistical Society*, B, 63, 167-241.

[5] Barndorff-Nielsen, O. and N. Shephard (2002). “Econometric analysis of realized volatility and its use in estimating stochastic volatility models,” *Journal of the Royal Statistical Society*, Series B, 64, 253-280.

[6] Barndorff-Nielsen, O. and N. Shephard, (2006). “Econometrics of testing for jumps in financial economics using bipower variation,” *Journal of Financial Econometrics*, 4, 1-30.

[7] Barndorff-Nielsen, O., P. Hansen, A. Lunde, and N. Shephard (2008). “Designing realised kernels to measure the ex-post variation of equity prices in the presence of noise,” *Econometrica*, 76, 1481-1536.

[8] Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N. (2009). “Realised kernels in practice: trades and quotes,” *Econometrics Journal* 12, C1-C32.

[9] Bühlmann, P. and H. R. Künsch (1995). “The blockwise bootstrap for general parameters of a stationary time series,” *Scandinavian Journal of Statistics*, 22(1), 35-54.

[10] Christensen, K., R. Oomen and M. Podolskij (2010). “Realised quantile-based estimation of the integrated variance,” *Journal of Econometrics 159*, 74-98.

[11] 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, 116-133.

[12] Christensen, K., M. Podolskij, and M. Vetter (2013). “On covariation estimation for multivariate continuous Ito semimartingales with noise in non-synchronous observation schemes,” *Journal of Multivariate Analysis*, 120, 59-84.

[13] Gloter, A. and J. Jacod (2001). “Diffusions with measurement errors. ii - optimal estimators,” *ESAIM*, 5, 243-260.

[14] Gonçalves, S. and N. Meddahi (2008). “Edgeworth corrections for realized volatility,” *Econometric Reviews*, 27, 139-162.

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

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

[17] Gonçalves, S., and H. White (2002). “The bootstrap of the mean for dependent heterogeneous arrays,” *Econometric Theory*, 18, 1367-1384.

[18] Hansen, P.R. and A. Lunde (2006). “Realized variance and market microstructure noise,” *Journal of Business and Economics Statistics*, 24, 127-161.

[19] Hautsch N., and Podolskij, M., (2013). “Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, Implementation, and Empirical Evidence,” *Journal of Business and Economic Statistics*, 31(2), 165-183.

[20] Heston, S. (1993). "Closed-form solution for options with stochastic volatility with applications to bonds and currency options," *Review of Financial Studies*, 6, 327-343.

[21] Huang, X., and G. Tauchen (2006). "The relative contribution of jumps to total price variance," *Journal of Financial Econometrics*, 3, 456-499.

[22] Jacod, J. (2008). "Asymptotic properties of realized power variations and related functionals of semimartingales," *Stochastic Processes and Their Applications*, 118, 517-559.

[23] Jacod, J., Y. Li, P. Mykland, M. Podolskij, and M. Vetter (2009). "Microstructure noise in the continuous case: the pre-averaging approach," *Stochastic Processes and Their Applications*, 119, 2249-2276.

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

[25] Jacod, J. and P. Protter, (1998). "Asymptotic error distributions for the Euler method for stochastic differential equations,"*Annals of Probability* 26, 267-307.

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

[27] Künsch, H.R. (1989). "The jackknife and the bootstrap for general stationary observations," *Annals of Statistics* 17, 1217-1241.

[28] Liu, R.Y. (1988). "Bootstrap procedure under some non-i.i.d. models," *Annals of Statistics* 16, 1696-1708.

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

[30] Mykland, P. and L. Zhang (2014). "Assessment of Uncertainty in High Frequency Data: The Observed Asymptotic Variance," *Manuscript, The University of Chicago.*

[31] Phillips, P.C.B., and H.R. Moon. (1999). "Linear Regression Limit Theory for Nonstationary Panel Data," *Econometrica*, 67, 1057-1111.

[32] Podolskij, M., and M. Vetter (2009). "Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps," *Bernoulli*, 15(3), 634-658.

[33] Politis, D. N. and Romano, J. P. (1992). "A general resampling scheme for triangular arrays of α-mixing random variables," *Annals of Statistics* 20, 1985-2007.

[34] Politis, D.N., Romano, J.P., and Wolf, M. (1999). "Subsampling," Springer-Verlag, New York.

[35] Vetter, M. (2008). "Estimation methods in noisy diffusion models," *Ph.D. Thesis. Bochum University.*

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

[37] Xiu, D. (2010). "Quasi-maximum likelihood estimation of volatility with high frequency data," *Journal of Econometrics*, 159, 235-250.

[38] Zhang, L. (2006). "Efficient estimation of stochastic volatility using noisy observations: a multi-scale approach," *Bernoulli*, 12, 1019-1043.

[39] 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, 1394-1411.

[40] Zhang, L., Mykland, P. and Y. Aït-Sahalia (2011). "Edgeworth expansions for realized volatility and related estimators," *Journal of Econometrics*, 160, 190-203.