AARHUS UNIVERSITY
BUSINESS AND SOCIAL SCIENCES
DEPARTMENT OF ECONOMICS AND BUSINESS

CREATES
Center for Research in Econometric Analysis of Time Series

Bootstrapping realized volatility and realized beta under a local Gaussianity assumption

Ulrich Hounyo

CREATES Research Paper 2013-30

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

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

Bootstrapping realized volatility and realized beta under a local Gaussianity assumption*

Ulrich Hounyo †
Oxford-Man Institute, University of Oxford,
CREATES, University of Aarhus,

September 16, 2013

Abstract

The main contribution of this paper is to propose a new bootstrap method for statistics based on high frequency returns. The new method exploits the local Gaussianity and the local constancy of volatility of high frequency returns, two assumptions that can simplify inference in the high frequency context, as recently explained by Mykland and Zhang (2009).

Our main contributions are as follows. First, we show that the local Gaussian bootstrap is first-order consistent when used to estimate the distributions of realized volatility and realized betas. Second, we show that the local Gaussian bootstrap matches accurately the first four cumulants of realized volatility, implying that this method provides third-order refinements. This is in contrast with the wild bootstrap of Gonçalves and Meddahi (2009), which is only second-order correct. Third, we show that the local Gaussian bootstrap is able to provide second-order refinements for the realized beta, which is also an improvement of the existing bootstrap results in Dovonon, Gonçalves and Meddahi (2013) (where the pairs bootstrap was shown not to be second-order correct under general stochastic volatility). Lastly, we provide Monte Carlo simulations and use empirical data to compare the finite sample accuracy of our new bootstrap confidence intervals for integrated volatility and integrated beta with the existing results.
JEL Classification: C15, C22, C58
Keywords: High frequency data, realized volatility, realized beta, bootstrap, Edgeworth expansions.

1 Introduction

Realized measures of volatility have become extremely popular in the last decade as higher and higher frequency returns are available. Despite the fact that these statistics are measured over large samples, their finite sample distributions are not necessarily well approximated by their asymptotic mixed-Gaussian distributions. This is especially true for realized statistics that are not robust to market microstructure noise since in this case researchers usually face a trade-off between using large sample

*I would like to thank Nour Meddahi for very stimulating discussions, which led to the idea of this paper. I am especially indebted to Sílvia Gonçalves for her valuable comments. I am also grateful to Asger Lunde, Peter Exterkate and Matthew Webb, who provided very useful comments on previous versions of this paper. I acknowledge 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.

†Oxford-Man Institute of Quantitative Finance, University of Oxford. Address: Eagle House, Walton Well Road, Oxford OX2 6ED, UK. Email: ulrich.hounyo@oxford-man.ox.ac.uk.

sizes and incurring in market microstructure biases. This has spurred interest in developing alternative approximations based on the bootstrap. In particular, Gonçalves and Meddahi (2009) have recently proposed bootstrap methods for realized volatility whereas Dovonon, Gonçalves and Meddahi (2013) have studied the application of the bootstrap in the context of realized regressions.

The main contribution of this paper is to propose a new bootstrap method that exploits the local Gaussianity framework described in Mykland and Zhang (2009, 2011). As these authors explain, one useful way of thinking about inference in the context of realized measures is to assume that returns have constant variance and are conditionally Gaussian over blocks of consecutive M observations. Roughly speaking, a high frequency return of a given asset is equal in law to the product of its volatility (the spot volatility) multiplied by a normal standard distribution. Mykland and Zhang (2009) show that this local Gaussianity assumption is useful in deriving the asymptotic theory for the estimators used in this literature by providing an analytic tool to find the asymptotic behaviour without calculations being too cumbersome. This approach also has the advantage of yielding more efficient estimators by varying the size of the block (see Mykland and Zhang (2009) and Mykland, Shephard and Sheppard (2012).

The main idea of this paper is to see how and to what extent this local Gaussianity assumption can be explored to generate a bootstrap approximation. In particular, we propose and analyze a new bootstrap method that relies on the conditional local Gaussianity of intraday returns. The new method (which we term the local Gaussian bootstrap) consists of dividing the original data into non-overlapping blocks of M observations and then generating the bootstrap observations at each frequency within a block by drawing a random draw from a normal distribution with mean zero and variance given by the realized volatility over the corresponding block. Using Mykland and Zhang's (2009) blocking approach, one can act as if the instantaneous volatility is constant over a given block of consecutive observations. In practice, the volatility of asset returns is highly persistent, especially over a daily horizon, implying that it is at least locally nearly constant.

We focus on two realized measures in this paper: realized volatility and realized regression coefficients. The latter can be viewed as a smooth function of the realized covariance matrix. Our proposal in this case is to generate bootstrap observations on the vector that collects the intraday returns that enter the regression model by applying the same idea as in the univariate case. Specifically, we generate bootstrap observations on the vector of variables of interest by drawing a random vector from a multivariate normal distribution with mean zero and covariance matrix given by the realized covariance matrix computed over the corresponding block.

Our findings for realized volatility are as follows. When M is fixed, the local Gaussian bootstrap is asymptotically correct but it does not offer any asymptotic refinements. More specifically, the first four bootstrap cumulants of the t-statistic based on realized volatility and studentized with a variance estimator that is based on a block size of M do not match the cumulants of the original t-statistic to higher order (although they are consistent). Note that when $M = 1$, the new bootstrap method coincides with the wild bootstrap of Gonçalves and Meddahi (2009) based on a $N(0,1)$ external random variable. As Gonçalves and Meddahi (2009) show, this is not an optimal choice, which is in line with our results. Therefore, our result generalizes that of Gonçalves and Meddahi (2009) to the case of a fixed $M > 1$. However, if the block length $M \to \infty$ at rate $o(h^{-1/2})$ (where h^{-1} denotes the sample size), then the local Gaussian bootstrap is able to provide an asymptotic refinement. In particular, we show that the first and third bootstrap cumulants of the t-statistic converge to the corresponding cumulants at the rate $o\left(h^{-1/2}\right)$, which implies that the local Gaussian bootstrap offers a second-order refinement. In this case, the local Gaussian bootstrap is an alternative to the optimal two-point distribution wild bootstrap proposed by Gonçalves and Meddahi (2009). More interestingly, we also show that the local Gaussian bootstrap is able to match the second and fourth order cumulants through order $o(h)$, which implies that this method is able to provide a third-order asymptotic refinement. This is contrast to the optimal wild bootstrap methods of Gonçalves and Meddahi (2009), which can not deliver third-order asymptotic refinements.

For the realized regression estimator proposed by Mykland and Zhang (2009), the local Gaussian bootstrap matches the cumulants of the t-statistics through order $o\left(h^{-1/2}\right)$ when $M \to \infty$ at rate $o\left(h^{-1/2}\right)$. Thus, this method can promise second-order refinements. This is contrast with the pairs bootstrap studied by Dovonon, Gonçalves and Meddahi (2013), which is only first-order correct.

Our Monte Carlo simulations suggest that the new bootstrap method we propose improves upon the first-order asymptotic theory in finite samples and outperforms the existing bootstrap methods.

The rest of this paper is organized as follows. In the next section, we first introduce the setup, our assumptions and describe the local Gaussian bootstrap. In Sections 3 and 4 we establish the consistency of this method for realized volatility and realized beta, respectively. Section 5 contains the higher-order asymptotic properties of the bootstrap cumulants. Section 6 contains simulations, Section 7 contains one empirical application and Section 8 concludes. Three appendices are provided. Appendix A contains the tables with simulation results whereas Appendix B and Appendix C contain the proofs.

2 Framework and the local Gaussian bootstrap

The statistics of interest in this paper can be written as smooth functions of the realized multivariate volatility matrix. Here we describe the theoretical framework for multivariate high frequency returns and introduce the new bootstrap method we propose. Sections 3 and 4 will consider in detail the theoretical properties of this method for the special cases of realized volatility and realized beta, respectively.

We follow Mykland and Zhang (2009) and assume that the log-price process $X_t = \left(X_t^{(1)} \cdots X_t^{(d)}\right)'$ of a d-dimensional vector of assets is defined on a probability space $(\Omega, \mathcal{F}, P)$ equipped with a filtration

$(\mathcal{F}_t)_{t\geq 0}$. We model X as a Brownian semimartingale process that follows the equation,

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

where $\mu = (\mu_t)_{t\geq 0}$ is a d-dimensional predictable locally bounded drift vector, $\sigma = (\sigma_t)_{t\geq 0}$ is an adapted càdlàg $d \times d$ locally bounded spot covolatility matrix and $W = (W_t)_{t\geq 0}$ is d-dimensional Brownian motion.

We follow Barndorff-Nielsen et al. (2006) and assume that the spot covariance matrix $\Sigma_t = \sigma_t \sigma_t'$ is invertible and satisfies the following assumption

$$\Sigma_t = \Sigma_0 + \int_0^t a_s ds + \int_0^t b_s dW_s + \int_0^t v_s dZ_s, \tag{2}$$

where a, b, and v are all adapted càdlàg processes, with a also being predictable and locally bounded, and Z is a vector Brownian motion independent of W.

The representation in (1) and (2) is rather general as it allows for leverage and drift effects. Assumption 2 of Mykland and Zhang (2009) or equation (1) of Mykland and Zhang (2011) also impose a Brownian semimartingale structure on the instantaneous covariance matrix Σ. Equation (2) rules out jumps in volatility, but this can be relaxed (see Assumption H1 of Barndorff-Nielsen et al. (2006) for a weaker assumption on Σ).

Suppose we observe X over a fixed time interval $[0, 1]$ at regular time points ih, for $i = 0, \ldots, 1/h$, from which we compute $1/h$ intraday returns at frequency h,

$$y_i \equiv X_{ih} - X_{(i-1)h} = \int_{(i-1)h}^{ih} \mu_t dt + \int_{(i-1)h}^{ih} \sigma_t dW_t, \quad i = 1, \ldots, \frac{1}{h}, \tag{3}$$

where we will let y_{ki} to denote the i-th intraday return on asset k , $k = 1, \ldots, d$.

As equation (3) shows, the intraday returns y_i depend on the drift μ, unfortunately when carrying out inference for observations in a fixed time interval the process μ_t cannot be consistently estimated. For most purposes it is only a nuisance parameter. To deal with this, Mykland and Zhang (2009) propose to work with a new probability measure which is measure theoretically equivalent to P and under which there is no drift (a statistical risk neutral measure). They pursue the analysis further and propose an approximation measure Q_h defined on the discretized observations X_{ih} only, for which the volatility is constant on each of the $\frac{1}{Mh}$ non overlapping blocks of size M. Since M is the number of high frequency returns within a block, we have that $M \leq \frac{1}{h}$.

Specifically, under the approximate measure Q_h, in each block $j = 1, \ldots, \frac{1}{Mh}$, we have,

$$y_i = \frac{1}{\sqrt{M}} C_{(j)} \eta_{i+(j-1)M},\ \forall i \in ((j-1)M, jM], \tag{4}$$

where $\eta_{i+(j-1)M} \sim i.i.d.N(0, I_d)$, I_d is a $d \times d$ identity matrix and $C_{(j)} = \sqrt{Mh}\sigma_{(j-1)Mh}$, where $C_{(j)}$ is such that $C_{(j)}C_{(j)}' = \Gamma_{(j)} \equiv \int_{(j-1)Mh}^{jMh} \Sigma_u du$ (see Mykland and Zhang (2009), p.1417 for a formal definition of Q_h).

The true distribution is P, but we prefer to work with Q_h since then calculations are much simpler. Afterwards we adjust results back to P using the likelihood ratio (Radon-Nikodym derivative) dQ_h/dP.

Remark 1. As pointed out in Mykland and Zhang's (2009) Theorem 3 and, in Mykland and Zhang's (2011) Theorem 1, the measure P and its approximation Q_h are contiguous on the observables. This is to say that for any sequence $\mathcal{A}_h$ of sets, $P(\mathcal{A}_h) \to 0$ if and only if $Q_h(\mathcal{A}_h) \to 0$ (see Mykland and Zhang (2012) p. 169 for more details). In particular, if an estimator is consistent under Q_h, it is also consistent under P. Rates of convergence (typically $h^{-1/2}$) are also preserved, but the asymptotic distribution may change (for instances of this, see Examples 3 and 5 of Mykland and Zhang (2009). More specifically, when adjusting from Q_h to P, the asymptotic variance of the estimator is unchanged (due to the preservation of quadratic variation under limit operations), while the asymptotic bias may change (see Remark 4 of Mykland and Zhang (2009)). It appears that a given sequence Z_h of martingales will have exactly the same asymptotic distribution under Q_h and P, when the Q_h martingale part of the log likelihood ratio $log(dP/dQ_h)$ has zero asymptotic covariation with Z_h. In this case, we do not need to adjust the distributional result from Q_h to P. Two important examples where this is true are the realized volatility and realized beta which we will study in details in Sections 3 and 4.

Remark 2. In the particular case where the window length M increases with the sample size h^{-1} at rate $o(h^{-1/2})$, there is also no contiguity adjustment (see Remark 2 of Mykland and Zhang (2011)).

Next we introduce a new bootstrap method that exploits the structure of (4). In particular, we mimic the original observed vector of returns, and we use the normality of the data and replace $C_{(j)}$ by its estimate $\hat{C}_{(j)}$, where $\hat{C}_{(j)}$ is such that $\hat{C}_{(j)}\hat{C}'_{(j)} = \sum_{i=1}^{M} y_{i+(j-1)M} y'_{i+(j-1)M} = \hat{\Gamma}_{(j)}$. That is, we follow the main idea of Mykland and Zhang (2009), and assume constant volatility within blocks. Then, inside each block j of size M ($j = 1, \ldots, \frac{1}{Mh}$), we generate the M vector of returns as follows,

$$y^*_{i+(j-1)M} = \frac{1}{\sqrt{M}} \hat{C}_{(j)} \eta_{i+(j-1)M}, \; 1 = 1, \ldots, M, \tag{5}$$

where $\eta_{i+(j-1)M} \sim i.i.d.N(0, I_d)$ across (i, j), and I_d is a $d \times d$ identity matrix.

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_h^*, we write $Z_h^* = o_{P^*}(1)$ in probability, or $Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under P, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{h\to 0} P\left[P^*\left(|Z_h^*| > \delta\right) > \varepsilon\right] = 0$. Similarly, we write $Z_h^* = O_{P^*}(1)$ as $h \to 0$, in probability if for all $\varepsilon > 0$ there exists a $M_\varepsilon < \infty$ such that $\lim_{h\to 0} P\left[P^*\left(|Z_h^*| > M_\varepsilon\right) > \varepsilon\right] = 0$. Finally, we write $Z_h^* \to^{d_*} Z$ as $h \to 0$, in probability under P, if conditional on the sample, Z_h^* weakly converges to Z under P^*, for all samples contained in a set with probability converging to one.

The following result is crucial in obtaining our bootstrap results.

Theorem 2.1. *Let Z_h^* be a sequence of bootstrap statistics. Given the probability measure P and its approximation Q_h, we have that*

$Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under P, if and only if $Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under Q_h.

Proof of Theorem 2.1 For any $\varepsilon > 0$, $\delta > 0$, letting $\mathcal{A}_h \equiv \{P^*(|Z_h^*| > \delta) > \varepsilon\}$, we have that

$Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under P, if for any $\varepsilon > 0$, $\delta > 0$, $\lim_{h\to 0} P(\mathcal{A}_h) = 0$. This is equivalent to $\lim_{h\to 0} Q_h(\mathcal{A}_h) = 0$, since P and Q_h are contiguous (see Remark 1). It follows then that $Z_h^* \to^{P^*} 0$, as $h \to 0$, in probability under Q_h. The inverse follows similarly.

Theorem 2.1 provides a theoretical justification to derive bootstrap consistency results under the approximation measure Q_h as well as under P. This simplifies the bootstrap inference. We will subsequently rely on this theorem to establish the bootstrap consistency results.

3 Results for realized volatility

3.1 Existing asymptotic theory

To describe the asymptotic properties of realized volatility, we need to introduce some notation. For any $q > 0$, define the realized q-th order power variation (cf. Remark 8 of Mykland and Zhang (2009)) as

$$R_q \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{RV_{j,M}}{Mh} \right)^{q/2}.$$

where $RV_{j,M} = \sum_{i=1}^{M} y_{i+(j-1)M}^2$ is the realized volatility over the period $[(j-1)Mh, jMh]$ for $j = 1, \ldots, \frac{1}{Mh}$. Note that when $q = 2$, $R_2 = RV$ (realized volatility). Similarly, for any $q > 0$, define the integrated power variation by

$$\overline{\sigma^q} \equiv \int_0^1 \sigma_u^q du.$$

Mykland and Zhang (2009) show that $\frac{1}{c_{M,q}} R_q \xrightarrow{P} \overline{\sigma^q}$, where $c_{M,q} \equiv E\left(\left(\frac{\chi_M^2}{M} \right)^{q/2} \right)$ with χ_M^2 the standard χ^2 distribution with M degrees of freedom and

$$c_{M,q} = \left(\frac{2}{M} \right)^{q/2} \frac{\Gamma\left(\frac{q+M}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}, \tag{6}$$

where Γ is the Gamma function. Similarly, Mykland and Zhang (2009) provide a CLT result for R_q with M fixed, whereas Mykland and Zhang (2011) allow M to go to infinity with the sample size h^{-1}, provided M is of order $O(h^{-1/2})$. In particular, for $q = 2$, we have that under P and Q_h, as the number of intraday observations increases to infinity,

$$\frac{\sqrt{h^{-1}}\left(R_2 - \overline{\sigma^2}\right)}{\sqrt{V}} \xrightarrow{d} N(0,1), \tag{7}$$

where

$$V = \frac{M\left(c_{M,4} - c_{M,2}^2\right)}{c_{M,2}^2} \int_0^1 \sigma_u^4 du.$$

In practice, this result is infeasible since the asymptotic variance V depends on an unobserved quantity, the integrated quarticity $\int_0^1 \sigma_u^4 du$. Mykland and Zhang (2009) propose a consistent estimator of V ($\widehat{V} = \frac{M\left(c_{M,4}-c_{M,2}^2\right)}{c_{M,2}^2} \frac{1}{c_{M,4}} R_4$), and together with (7), we have the feasible CLT (cf. Remark 8 of Mykland and Zhang (2009)):

$$T_{h,M} \equiv \frac{\sqrt{h^{-1}}\left(R_2 - \overline{\sigma^2}\right)}{\sqrt{\hat{V}}} \xrightarrow{d} N(0,1).$$

Note that, when the block size $M = 1$, this result is equivalent to the CLT for realized volatility derived by Barndorff-Nielsen and Shephard (2002). In particular, $c_{1,2} = E\left(\chi_1^2\right) = 1$, and $c_{1,4} = E\left(\chi_1^2\right)^2 = 3$. Here, when $M > 1$, the realized volatility R_2 using the blocking approach is the same realized volatility studied by Barndorff-Nielsen and Shephard (2002), but the t-statistic is different because $\hat{V}$ changes with M. One advantage of the block-based estimator is to improve efficiency by varying the size of the block (see for e.g. Mykland, Shephard and Sheppard (2012)).

3.2 Bootstrap consistency

Here we show that the new bootstrap method we proposed in Section 2 is consistent when applied to realized volatility. Specifically, given (5) with $d = 1$, for $j = 1, \ldots, 1/Mh$, we let

$$y_{i+(j-1)M}^* = \sqrt{\frac{RV_{j,M}}{M}} \eta_{i+(j-1)M}, \; j = 1, \ldots, M, \tag{8}$$

where $\eta_{i+(j-1)M} \sim$ i.i.d.$N(0,1)$ across (i,j). Note that this bootstrap method is related to the wild bootstrap approach proposed by Gonçalves and Meddahi (2009). In particular, when $M = 1$ and $d = 1$, it is equivalent to the wild bootstrap based on a standard normal external random variable.

We define the bootstrap realized volatility estimator as follows

$$R_2^* = \sum_{i=1}^{1/h} y_i^{*2} = \sum_{j=1}^{1/Mh} RV_{j,M}^*,$$

where $RV_{j,M}^* = \sum_{i=1}^{M} y_{i+(j-1)M}^{*2}$. Letting

$$\frac{1}{M}\sum_{i=1}^{M} \eta_{i+(j-1)M}^2 \equiv \frac{\chi_{j,M}^2}{M},$$

it follows that $RV^*_{j,M} = \frac{\chi^2_{j,M}}{M} RV_{j,M}$. We can easily show that

$$E^*\left(R_2^*\right) = c_{M,2} R_2,$$

and

$$V^* \equiv Var^*\left(h^{-1/2} R_2^*\right) = M\left(c_{M,4} - c_{M,2}^2\right) R_4.$$

Hence, we propose the following consistent estimator of V^*:

$$\hat{V}^* = M \frac{c_{M,4} - c_{M,2}^2}{c_{M,4}} R_4^*.$$

The bootstrap analogue of $T_{h,M}$ is given by

$$T^*_{h,M} \equiv \frac{\sqrt{h^{-1}}\left(R_2^* - c_{M,2} R_2\right)}{\sqrt{\hat{V}^*}}.$$

Theorem 3.1. *Suppose (1), (2) and (8) hold. If M is fixed or $M \to \infty$ as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$, then as $h \to 0$,*

$$\sup_{x \in \Re} \left| P^*\left(T^*_{h,M} \leq x\right) - P\left(T_{h,M} \leq x\right) \right| \to 0,$$

in probability under Q_h and under P.

Theorem 3.1 provides a theoretical justification for using the bootstrap distribution of $T^*_{h,M}$ to estimate the distribution of $T_{h,M}$ under the general context studied by Mykland and Zhang (2009) which allow for the presence of drifts and leverage effects under P. This result also justifies the use of the bootstrap for constructing the studentized bootstrap (percentile-t) intervals.

Note that, when $M \to \infty$, such that $M = o(h^{-1/2})$, $V^* \overset{P}{\to} V$, we can also show that bootstrap percentile intervals for integrated volatility are valid. This is in contrast to the optimal two-point wild bootstrap proposed by Gonçalves and Meddahi (2009), which is only valid for percentile-t intervals.

4 Results for realized beta

4.1 Existing asymptotic theory and a new variance estimator

The goal of this section is to describe the realized beta in the context of Mykland and Zhang's (2009) blocking approach. In order to obtain a feasible CLT, we propose a consistent estimator of the variance of the realized beta, which is a new estimator in this literature. To derive this result, we use the approach of Dovonon, Gonçalves and Meddahi (2013) and suppose that σ is independent of W.[1] Note that contrary to Dovonon, Gonçalves and Meddahi (2013), we do not need here to suppose that $\mu_t = 0$ (since under Q_h high frequency returns have mean zero conditionally on σ).

[1] We make the assumption of no leverage for notational simplicity and because this allows us to easily compute the moments of the intraday returns conditionally on the volatility path. The same arguments would follow under the presence of leverage (for instance, by postulating a model for σ_t, as in Meddahi (2002)) but this would unnecessarily complicate the notation without any gain in the intuition.

For simplicity, we consider the bivariate case where $d = 2$ and look at results for assets k and l, whose ith high frequency returns in the jth block will be written as $y_{k,i+(j-1)M}$ and $y_{l,i+(j-1)M}$, respectively, for $i = 1, \ldots, M$ and $j = 1, \ldots, \frac{1}{Mh}$. It follows that under Q_h, $y_{l,i+(j-1)M} = \frac{1}{\sqrt{M}} C_{1(j)} \eta_{1,i+(j-1)M}$ and $y_{k,i+(j-1)M} = \frac{1}{\sqrt{M}} C_{21(j)} \eta_{1,i+(j-1)M} + \frac{1}{\sqrt{M}} C_{2(j)} \eta_{2,i+(j-1)M}$, where

$$C_{(j)} \equiv \begin{pmatrix} C_{1(j)} & 0 \\ C_{21(j)} & C_{2(j)} \end{pmatrix} = \begin{pmatrix} \sqrt{\Gamma_{l(j)}} & 0 \\ \frac{\Gamma_{lk(j)}}{\sqrt{\Gamma_{l(j)}}} & \sqrt{\Gamma_{k(j)} - \frac{\Gamma_{lk(j)}^2}{\Gamma_{l(j)}}} \end{pmatrix},$$

$$\eta_{i+(j-1)M} \equiv \begin{pmatrix} \eta_{1,i+(j-1)M} \\ \eta_{2,i+(j-1)M} \end{pmatrix} \sim i.i.d.N(0, I_2),$$

I_2 is a 2×2 identity matrix, $\Gamma_{lk(j)} = \int_{(j-1)Mh}^{jMh} \Sigma_{lk}(u)\, du$, and when $k = l$, we write $\Gamma_{k(j)} = \Gamma_{kk(j)}$.

Then, conditionally on Σ, we can write

$$y_{li} = \beta_{lki} y_{ki} + u_i, \tag{9}$$

where independently across $i = 1, \ldots, 1/h$, $u_i | y_{ki} \sim N(0, V_i)$, with $V_i \equiv \Gamma_{li} - \frac{\Gamma_{lki}^2}{\Gamma_{ki}}$, and $\beta_{lki} \equiv \frac{\Gamma_{lki}}{\Gamma_{ki}}$, where $\Gamma_{lki} = \int_{(i-1)h}^{ih} \Sigma_{lk}(u)\, du$.

As Dovonon, Gonçalves and Meddahi (2013) argue, the conditional mean parameters of realized regression models are heterogeneous under stochastic volatility. This heterogeneity justifies why the pairs bootstrap method that they studied is not second-order accurate.

Under the approximation measure Q_h for the observables in the jth block ($j = 1, \ldots, \frac{1}{Mh}$), the regression (9) becomes

$$y_{l,i+(j-1)M} = \beta_{lk(j)} y_{k,i+(j-1)M} + u_{i+(j-1)M}, \tag{10}$$

where $u_{i+(j-1)M} | y_{k,i+(j-1)M} \sim$i.i.d.$N\left(0, V_{(j)}\right)$, for $i = 1, \ldots, M$, with $V_{(j)} \equiv \frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma_{lk(j)}^2}{\Gamma_{k(j)}}\right)$, and $\beta_{lk(j)} \equiv \frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}} = \frac{1}{Mh} \int_{(j-1)Mh}^{jMh} \beta_{lk}(u)\, du$. This implies that the integrated beta is $\beta_{lk} = Mh \sum_{j=1}^{1/Mh} \beta_{lk(j)} = \int_0^1 \beta_{lk}(u)\, du$.

Let us denote by $\hat{\beta}_{lk(j)}$ the ordinary least squares (OLS) estimator of $\beta_{lk(j)}$. Mykland and Zhang (2009) proposed to use $\hat{\beta}_{lk}$ defined as follows,

$$\hat{\beta}_{lk} = Mh \sum_{j=1}^{1/Mh} \hat{\beta}_{lk(j)} = Mh \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \right)^{-1} \left(\sum_{i=1}^{M} y_{k,i+(j-1)M} y_{l,i+(j-1)M} \right),$$

to estimate the integrated beta. Note that the realized beta estimator studied by Dovonon, Gonçalves and Meddahi (2013) is a different statistic than ours. Here, the realized beta estimator $\hat{\beta}_{lk}$ is not directly a least squares estimator, but is the result of the average of $\hat{\beta}_{lk(j)}$, the OLS estimators for each block. Since under Q_h, the volatility matrix is constant in each block j, we have that $\beta_{lki} = \beta_{lk(j)}$, for all $i = 1, \ldots, M$, implying consequently that the score is not heterogeneous and has mean zero. This simplifies the asymptotic inference on $\beta_{lk(j)}$, and on β_{lk}. Also note that contrary to what we have observed in the case of realized volatility estimator, here when $M = 1$, the realized beta estimator using the blocking approach become

$$\hat{\beta}_{lk} = h \sum_{i=1}^{1/h} \frac{y_{l,i}}{y_{k,i}},$$

which is a different statistic than the statistic studied by Barndorff-Nielsen and Shephard (2004). But when $M = h^{-1}$, both estimators are equivalent. However, as Mykland and Zhang (2011) pointed out, when $M \to \infty$ with the sample size h^{-1}, the local approximation is good only when $M = O(h^{-1/2})$. It follows then that we are not confortable to contrast Mykland and Zhang (2009) block-based realized beta estimator asymptotic results with those of Barndorff-Nielsen and Shephard (2004a).

Mykland and Zhang (2009) provide a CLT result for β_{lk}. In particular, we have under P and Q_h, as the number of intraday observations increases to infinity (i.e. if $h \to 0$), by using Section 4.2 of Mykland and Zhang (2009),

$$\frac{\sqrt{h^{-1}} \left(\hat{\beta}_{lk} - \beta_{lk} \right)}{\sqrt{V_\beta}} \xrightarrow{d} N(0,1), \tag{11}$$

where

$$V_\beta = \begin{cases} \frac{M}{M-2} \int_0^1 \left(\frac{\Sigma_l(u)}{\Sigma_k(u)} - \beta_{lk}^2(u) \right) du, & \text{if } M = O(1), \text{ as } h \to 0 \text{ such that } M > 2(1+\delta) \text{ for any } \delta > 0, \\ \int_0^1 \left(\frac{\Sigma_l(u)}{\Sigma_k(u)} - \beta_{lk}^2(u) \right) du, & \text{if } M \to \infty \text{ as } h \to 0 \text{ such that } M = o(h^{-1/2}), \end{cases}$$

In practice, this result is infeasible since the asymptotic variance V_β depends on unobserved quantities. Mykland and Zhang (2009) did not provide any consistent estimator of V_β. One of our contributions is to propose a consistent estimator of V_β. To this end, we exploit the special structure of the regression model. To find the asymptotic variance of realized regression estimator $\hat{\beta}_{lk}$, we can write

$$\sqrt{h^{-1}} \left(\hat{\beta}_{lk} - \beta_{lk} \right) = M\sqrt{h} \sum_{j=1}^{1/Mh} \left(\hat{\beta}_{lk(j)} - \beta_{lk(j)} \right).$$

Since $\hat{\beta}_{lk(j)}$ are independent across j, it follows that

$$V_{\beta,h,M} \equiv Var\left(\sqrt{h^{-1}} \left(\hat{\beta}_{lk} - \beta_{lk} \right) \right) = M^2 h \sum_{j=1}^{1/Mh} Var\left(\hat{\beta}_{lk(j)} - \beta_{lk(j)} \right). \tag{12}$$

To compute (12), note that from standard regression theory, we have that under Q_h,

$$Var\left(\hat{\beta}_{lk(j)} - \beta_{lk(j)} \right) = E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \right)^{-1} \right) V_{(j)},$$

which implies that

$$V_{\beta,h,M} = M^2 h \sum_{j=1}^{1/Mh} E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \right)^{-1} \right) V_{(j)}. \tag{13}$$

Note that we can contrast V_β with equation (72) of Mykland and Zhang (2009). In fact, we can write under Q_h, $\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \overset{d}{=} \frac{\Gamma_{k(j)}}{M} \sum_{i=1}^{M} v_{i+(j-1)M}^2 \overset{d}{=} \frac{\Gamma_{k(j)}}{M} \chi_{j,M}^2$, where $v_{i+(j-1)M} \sim i.i.d.N(0,1)$, and $\chi_{j,M}^2$ follow the standard χ^2 distribution with M degrees of freedom, and '$\overset{d}{=}$' denotes equivalence in distribution. Then for any integer $M > 2$ and conditionally on the volatility path, by using the expectation of the inverse of a Chi square distribution we have,

$$E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right) = E\left(\frac{M}{\chi_{j,M}^2}\right)\Gamma_{k(j)}^{-1} = \frac{M}{M-2}\Gamma_{k(j)}^{-1}. \tag{14}$$

It follows then that

$$V_{\beta,h,M} = \frac{M}{M-2}\sum_{j=1}^{1/Mh} Mh\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right). \tag{15}$$

By using the structure of (13), a natural consistent estimator of $V_{\beta,h,M}$ is

$$\hat{V}_{\beta,h,M} \equiv M^2 h \sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\frac{1}{M-1}\sum_{i=1}^{M}\hat{u}_{i+(j-1)M}^2\right), \tag{16}$$

where $\hat{u}_{i+(j-1)M} = y_{l,i+(j-1)M} - \hat{\beta}_{lk(j)} y_{k,i+(j-1)M}$ (see Lemma C.5 and Lemma C.7 in the Appendix). Together with the CLT result (11), we have under P and Q_h the feasible result

$$T_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}(\hat{\beta}_{lk} - \beta_{lk})}{\sqrt{\hat{V}_{\beta,h,M}}} \to^d N(0,1).$$

4.2 Bootstrap consistency

Here we show that the new bootstrap method we proposed in Section 2 is consistent when applied to realized betas. Specifically, given (5) with $d = 2$, for $j = 1, \ldots, 1/Mh$, we generate the M vector of returns as follows. For each $i = 1, \ldots, M$,

$$y_{i+(j-1)M}^* = \begin{pmatrix} y_{l,i+(j-1)M}^* \\ y_{k,i+(j-1)M}^* \end{pmatrix} = \frac{1}{\sqrt{M}}\begin{pmatrix} \sqrt{\hat{\Gamma}_{l(j)}}\eta_{1,i+(j-1)M} \\ \frac{\hat{\Gamma}_{lk(j)}}{\sqrt{\hat{\Gamma}_{l(j)}}}\eta_{1,i+(j-1)M} + \sqrt{\hat{\Gamma}_{k(j)} - \frac{\hat{\Gamma}_{lk(j)}^2}{\hat{\Gamma}_{l(j)}}}\eta_{2,i+(j-1)M} \end{pmatrix}, \tag{17}$$

where $\hat{\Gamma}_{l(j)} = \sum_{i=1}^{M} y_{l,i+(j-1)M}^2$, $\hat{\Gamma}_{k(j)} = \sum_{i=1}^{M} y_{k,i+(j-1)M}^2$, $\hat{\Gamma}_{lk(j)} = \sum_{i=1}^{M} y_{k,i+(j-1)M} y_{l,i+(j-1)M}$, and
$\begin{pmatrix} \eta_{1,i+(j-1)M} \\ \eta_{2,i+(j-1)M} \end{pmatrix} \sim i.i.d.N(0, I_2)$, I_2 is a 2×2 identity matrix.

Let $\hat{\beta}_{lk(j)}^*$ denote the OLS bootstrap estimator from the regression of $y_{l,i+(j-1)M}^*$ on $y_{k,i+(j-1)M}^*$ inside the block j. The bootstrap realized beta estimator is

$$\hat{\beta}_{lk}^* = Mh\sum_{j=1}^{1/Mh}\hat{\beta}_{lk(j)}^*.$$

It is easy to check that $\hat{\beta}_{lk}^*$ converges in probability (under P^*) to

$$\hat{\beta}_{lk} = Mh\sum_{j=1}^{1/Mh} E^*\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^{2*}\right)^{-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^* y_{l,i+(j-1)M}^*\right)\right).$$

The bootstrap analogue of the regression error $u_{i+(j-1)M}$ in model (10) is thus $u_{i+(j-1)M}^* = y_{l,i+(j-1)M}^* - \hat{\beta}_{lk(j)} y_{k,i+(j-1)M}^*$, whereas the bootstrap OLS residuals are defined as $\hat{u}_{i+(j-1)M}^* = y_{l,i+(j-1)M}^* -$

$\hat{\beta}^*_{lk(j)} y^*_{k,i+(j-1)M}$. Thus, conditionally on the observed vector of returns $y_{i+(j-1)M}$, it follows that $u^*_{i+(j-1)M} | y^*_{k,i+(j-1)M} \sim i.i.d.N\left(0, \hat{V}_{(j)}\right)$, for $i = 1, \ldots, M$, where

$$\hat{V}_{(j)} \equiv \frac{1}{M}\left(\hat{\Gamma}_{l(j)} - \frac{\hat{\Gamma}^2_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right) = \frac{1}{M}\left(\sum_{i=1}^{M} y^2_{l,i+(j-1)M} - \frac{\left(\sum_{i=1}^{M} y_{k,i+(j-1)M} y_{l,i+(j-1)M}\right)^2}{\sum_{i=1}^{M} y^2_{k,i+(j-1)M}}\right).$$

We can show that

$$Var^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})\right) = \frac{M-1}{M-2}\hat{V}_{\beta,h,M}.$$

It follows then that a sufficient condition for the bootstrap to provide a consistent estimator of the asymptotic variance of $\sqrt{h^{-1}}(\hat{\beta}_{lk} - \beta_{lk})$ is to allow M to go to infinity. In particular when M increases with h^{-1} but at rate $o(h^{-1/2})$ (so that there is no contiguity adjustment), the bootstrap can be used to approximate the quantiles of the distribution of the root

$$\sqrt{h^{-1}}\left(\hat{\beta}_{lk} - \beta_{lk}\right),$$

thus justifying the construction of bootstrap percentile confidence intervals for β_{lk}. Our next theorem summarizes these results.

Theorem 4.1. *Consider DGP (1), (2) and suppose (17) holds. Then conditionally on σ, as $h \to 0$, under Q_h and P, the following hold*

a)

$$V^*_{\beta,h,M} \equiv Var^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})\right) \xrightarrow{P} \begin{cases} \frac{M-1}{M-2}V_\beta, & \text{if } M = O(1), \text{ as } h \to 0 \text{ such that } M > 2(1+\delta) \text{ for any } \delta > 0, \\ V_\beta, & \text{if } M \to \infty \text{ as } h \to 0 \text{ such that } M = o(h^{-1/2}), \end{cases}$$

b) $\sup_{x\in\mathbb{R}}\left|P^*\left(\sqrt{h^{-1}}\left(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}\right) \le x\right) - P\left(\sqrt{h^{-1}}\left(\hat{\beta}_{lk} - \beta_{lk}\right)\right)\right| \xrightarrow{P} 0$, *as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$.*

Part (a) of Theorem 4.1 shows that the bootstrap variance estimator is not consistent for V_β when the block size M is finite. But when the realized betas become an efficient estimator of integrated betas (i.e. if $M \to \infty$), we can use the bootstrap variance of $\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})$ to consistently estimate the covariance matrix V^*_β. Results in part (b) imply that the bootstrap realized beta estimator has a first order asymptotic normal distribution with mean zero and covariance matrix V_β. This is in line with the existing results in the cross section regression context, where the wild bootstrap and the pairs bootstrap variance estimator of the least squares estimator are robust to heteroskedasticity in the error term.

Bootstrap percentile intervals do not promise asymptotic refinements. 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 $V^*_{\beta,h,M}$:

$$\hat{V}^*_{\beta,h,M} \equiv M^2 h \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M} \right)^{-1} \left(\frac{1}{M-1} \sum_{i=1}^{M} \hat{u}^{*2}_{i+(j-1)M} \right). \quad (18)$$

Our proposal is to use this estimator to construct the bootstrap t-statistic, associated with the bootstrap realized regression coefficient $\hat{\beta}^*_{lk}$,

$$T^*_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}} \left(\hat{\beta}^*_{lk} - \hat{\beta}_{lk} \right)}{\sqrt{\hat{V}^*_{\beta,h,M}}}, \quad (19)$$

the bootstrap analogue of $T_{\beta,h,M}$.

Theorem 4.2. *Consider DGP (1), (2) and suppose (17) holds. Let $M > 4(2+\delta)$ for any $\delta > 0$ such that M is fixed or $M \to \infty$ as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$, conditionally on σ, as $h \to 0$, the following hold.*

$$T^*_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}} \left(\hat{\beta}^*_{lk} - \hat{\beta}_{lk} \right)}{\sqrt{\hat{V}^*_{\beta,h,M}}} \to^{d^*} N(0,1), \text{ in probability, under } Q_h \text{ and } P.$$

Note that when the block size M is finite the bootstrap is also first order asymptotically valid when applied to the t-statistic $T^*_{\beta,h,M}$ (defined in (19)), as our Theorem 4.2 proves. This first order asymptotic validity occurs despite the fact that $V^*_{\beta,h,M}$ does not consistently estimate V_β when M is fixed. The key aspect is that we studentize the bootstrap OLS estimator with $\hat{V}^*_{\beta,h,M}$ (defined in (18)), a consistent estimator of $V^*_{\beta,h,M}$, implying that the asymptotic variance of the bootstrap t-statistic is one.

5 Higher-order properties

In this section, we investigate the asymptotic higher order properties of the bootstrap cumulants. Section 5.1 considers the case of realized volatility whereas Section 5.2 considers realized beta. The ability of the bootstrap to accurately match the cumulants of the statistic of interest is a first step to showing that the bootstrap offers an asymptotic refinement.

The results in this section are derived under the assumption of zero drift and no leverage (i.e. W is assumed independent of Σ). As in Dovonon, Gonçalves and Meddahi (2013), a nonzero drift changes the expressions of the cumulants derived here. The no leverage assumption is mathematically convenient as it allows us to condition on the path of volatility when computing the cumulants of our statistics. Allowing for leverage is a difficult but promising extension of the results derived here.

We introduce some notation. For any statistics T_h and T^*_h, we write $\kappa_j(T_h)$ to denote the j^{th} order cumulant of T_h and $\kappa^*_j(T^*_h)$ to denote the corresponding bootstrap cumulant. For $j = 1$ and 3, κ_j denotes the coefficient of the terms of order $O\left(\sqrt{h}\right)$ of the asymptotic expansion of $\kappa_j(T_h)$, whereas for $j = 2$ and 4, κ_j denotes the coefficients of the terms of order $O(h)$. The bootstrap coefficients $\kappa^*_{j,h}$ are defined similarly.

5.1 Higher order cumulants of realized volatility

Let $\sigma_{q,p} \equiv \frac{\overline{\sigma^q}}{\left(\overline{\sigma^p}\right)^{q/p}}$, for any $q, p > 0$, and $R_{q,p} \equiv \frac{R_q}{(R_p)^{q/p}}$. We make the following assumption.

Assumption H. The log price process follows (1) with $\mu_t = 0$ and σ_t is independent of W_t, where the volatility σ is a càdlàg process, bounded away from zero, and satisfies the following regularity condition:

$$\lim_{h \to 0} h^{(1/2)} \sum_{i=1}^{1/h} \left| \sigma^r_{\eta_i} - \sigma^r_{\xi_1} \right| = 0,$$

for some $r > 0$ and for any η_i and such that $0 \leq \xi_1 \leq \eta_1 \leq h \leq \xi_2 \leq \eta_2 \leq 2h \leq \ldots \leq \xi_{1/h} \leq \eta_{1/h} \leq 1$

Assumption H is stronger than required to prove the central limit theorem for R_q in Mykland and Zhang (2009), but it is a convenient assumption to derive the cumulants expansions of $T_{h,M}$ and $T^*_{h,M}$. Specifically, under Assumption H, Barndorff-Nielsen and Shephard (2004b) show that for any $q > 0$, $\overline{\sigma^q_h} - \overline{\sigma^q} = o(\sqrt{h})$, where $\overline{\sigma^q_h} = h^{1-q/2} \sum_{s=1}^{1/h} \left(\int_{(s-1)h}^{sh} \sigma^2_u du \right)^{q/2}$. Under Q_h, we have shown that for any positive integer $M \geq 1$, $\overline{\sigma^q_{h,M}} \equiv (Mh)^{1-q/2} \sum_{j=1}^{1/Mh} \left(\sigma^2_{j,M} \right)^{q/2} = \overline{\sigma^q_h}$ (see proof of Theorem B.1 in Appendix B). It follows that under Q_h and Assumption H, $\overline{\sigma^q_{h,M}} - \overline{\sigma^q} = o(\sqrt{h})$ and similarly $R_q - c_{M,q}\overline{\sigma^q} = o_P(\sqrt{h})$, (this result also holds under Q_h), a result on which we subsequently rely on to establish the cumulants expansion of $T_{h,M}$ and $T^*_{h,M}$.

The following result states our main findings for realized volatility.

Proposition 5.1. *Consider DGP (1) and suppose (8) holds. Under Assumption H, conditionally on σ and under Q_h, and P, it follows that*

i) $\lim_{h \to 0} \kappa^*_{1,h,M} - \kappa_1 = \left(\frac{c_{M,6}}{\left(c_{M,4}\right)^{3/2}} - 1 \right) \left(-\frac{A_{1,M}}{2} \sigma_{6,4} \right)$, *which is nonzero if M is finite, and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.*

ii)

$$\lim_{h \to 0} \kappa^*_{2,h,M} - \kappa_2 = \left(\frac{c_{M,8}}{\left(c_{M,4}\right)^2} - 1 \right) \left(C_{1,M} - A_{2,M} \right) \sigma_{8,4} + \left(\frac{\left(c_{M,6}\right)^2}{\left(c_{M,4}\right)^3} - 1 \right) \left(\frac{7}{4} A^2_{1,M} \sigma^2_{6,4} \right),$$

which is nonzero if M is finite and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.

iii) $\lim_{h \to 0} \kappa^*_{3,h,M} - \kappa_3 = \left(\frac{c_{M,6}}{\left(c_{M,4}\right)^{3/2}} - 1 \right) \left(B_{1,M} - 3A_{1,M} \right) \sigma_{6,4}$, *which is nonzero if M is finite, and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.*

iv)

$$\lim_{h\to 0} \kappa^*_{4,h,M} - \kappa_4 = \left(\frac{c_{M,6}}{(c_{M,4})^{3/2}} - 1\right)\left(B_{2,M} + 3C_{1,M} - 6A^2_{2,M}\right)\sigma_{8,4} + \left(\frac{(c_{M,6})^2}{(c_{M,4})^3} - 1\right)\left(18{A_{1,M}}^2 - 6A_{1,M}B_{1,M}\right)\sigma^2_{6,4}$$

which is nonzero if M is finite and it is zero if $M = o(h^{-1/2})$, as $h \to 0$.

Here, $A_{1,M}$, $A_{2,M}, C_{1,M}, C_{4,M}, C_{6,M}$, and $C_{8,M}$ are given in Lemma B.2.

Proposition 5.1 shows that the cumulants of $T_{h,M}$ and $T^*_{h,M}$ do not agree when the block size M is fixed, implying that the bootstrap does not provide a higher-order asymptotic refinement for finite values of M. Nevertheless, when $M \to \infty$ at a rate $o(h^{-1/2})$ the bootstrap matches the first and third order cumulants through order $O\left(h^{-1/2}\right)$, which implies that it provides a second-order refinement, i.e. the bootstrap distribution $P^*\left(T^*_{h,M} \leq x\right)$ consistently estimates $P\left(T_{h,M} \leq x\right)$ with an error that vanishes as $o\left(h^{-1/2}\right)$ (assuming the corresponding Edgeworth expansions exist). [2] This is in contrast with the first-order asymptotic Gaussian distribution whose error converges as $O\left(h^{-1/2}\right)$. Note that Gonçalves and Meddahi (2009) also proposed a choice of the external random variable for their wild bootstrap method which delivers second-order refinements. Our results for the bootstrap method based on the local Gaussianity are new. We will compare the two methods in the simulation section.

Parts (ii) and (iv) of Proposition 5.1 show that the new bootstrap method we propose is able to match the second and fourth order cumulants of $T_{h,M}$ when $M \to \infty$ as $h \to 0$ provided $M = o\left(h^{-1/2}\right)$. These results imply that the bootstrap distribution of $\left|T^*_{h,M}\right|$ consistently estimate the distribution of $\left|T_{h,M}\right|$ through order $O\left(h\right)$, in which case the bootstrap offers a third order asymptotic refinement (this again assumes that the corresponding Edgeworth expansions exist, something we have not attempted to prove in this paper). If this is the case, then the local Gaussian bootstrap will deliver symmetric percentile$-t$ intervals for integrated volatility with coverage probabilities that converge to zero at the rate $o\left(h\right)$. In contrast, the coverage probability implied by the asymptotic theory-based intervals converge to the desired nominal level at the rate $O\left(h\right)$. The potential for the local Gaussian bootstrap intervals to yield third-order asymptotic refinements is particularly interesting because Gonçalves and Meddahi (2009) show that their wild bootstrap method is not able to deliver such refinements. Thus, our method is an improvement not only of the Gaussian asymptotic distribution but also of the best existing bootstrap methods.

Remark 3 One reason why the local Gaussian bootstrap is not able to match cumulants when M is finite is that the equation $\frac{c_{M,q}}{\left(c_{M,p}\right)^{q/p}} = 1$ does not always have an integer solution when $q, p \geq 1$.

[2] We do not prove the validity of our Edgeworth expansions. Such a result would be a valuable contribution in itself, which we defer for future research. Here our focus is on using formal expansions to explain the superior finite sample properties of the bootstrap theoretically. See Mammen (1993), Davidson and Flachaire (2001) and Gonçalves and Meddahi (2009) for a similar approach.

For instance, the equation $\frac{c_{M,6}}{(c_{M,4})^{3/2}} = 1$ gives $M = -\frac{1}{4}$ as solution. However, we always have $\lim_{M\to\infty} \frac{c_{M,q}}{(c_{M,p})^{q/p}} = 1$. This is the reason why the local Gaussian bootstrap is able to match cumulants when $M \to \infty$ (but not when M is finite).

5.2 Higher order cumulants of realized beta

In this section, we provide the first and third order cumulants of realized beta. These cumulants enter the Edgeworth expansions of the one-sided distribution functions of $T_{\beta,h,M}$ and $T^*_{\beta,h,M}$, $P^*\left(T^*_{\beta,h,M} \le x\right)$ and $P\left(T_{\beta,h,M} \le x\right)$, respectively.

Proposition 5.2. *Suppose (1), (2) and (17) hold. Conditionally on Σ, under Q_h and P, if M is fixed or $M \to \infty$ as $h \to 0$ such that $M = o\left(h^{-1/2}\right)$, then as $h \to 0$,*

i) $\lim_{h\to 0} \kappa^*_{1,\beta,h,M} - \kappa_{1,\beta} = 0.$

ii) $\lim_{h\to 0} \kappa^*_{3,\beta,h,M} - \kappa_{3,\beta} = 0.$

Proposition 5.2 shows that the cumulants of $T_{\beta,h,M}$ and $T^*_{\beta,h,M}$ agree through order $O\left(\sqrt{h}\right)$, which implies that the error of the bootstrap approximation $P^*\left(T^*_{\beta,h,M} \le x\right)$ to the distribution of $T_{\beta,h,M}$ is of order $o\left(\sqrt{h}\right)$. Since the normal approximation has an error of the order $O\left(\sqrt{h}\right)$, this implies that the local Gaussian bootstrap is second-order correct. This result is an improvement over the bootstrap results in Dovonon, Gonçalves and Meddahi (2013), who showed that the pairs bootstrap is not second-order correct in the general case of stochastic volatility.

6 Monte Carlo results

In this section we assess by Monte Carlo simulation the accuracy of the feasible asymptotic theory approach of Mykland and Zhang (2009). We find that this approach leads to important coverage probability distortions when returns are not sampled too frequently. We also compare the finite sample performance of the new local Gaussian bootstrap method with the existing bootstrap method for realized volatility proposed by Gonçalves and Meddahi (2009).

For integrated volatility, we consider two data generating processes in our simulations. First, following Zhang, Mykland and Aït-Sahalia (2005), we use the one-factor stochastic volatility (SV1F) model of Heston (1993) as our data-generating process, i.e.

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

and

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

where $\nu_t = \sigma_t^2$, B and W are two Brownian motions, and we assume $Corr(B, W) = \rho$. The parameter values are all annualized. In particular, we let $\mu = 0.05/252$, $\kappa = 5/252$, $\alpha = 0.04/252$, $\gamma = 0.05/252$, $\rho = -0.5$.

We also consider the two-factor stochastic volatility (SV2F) model analyzed by Barndorff-Nielsen et al. (2008) and also by Gonçalves and Meddahi (2009), where [3]

$$
\begin{aligned}
dX_t &= \mu dt + \sigma_t dB_t, \\
\sigma_t &= s\text{-}exp\left(\beta_0, +\beta_1 \tau_{1t} + \beta_2 \tau_{2t}\right), \\
d\tau_{1t} &= \alpha_1 \tau_{1t} dt + dB_{1t}, \\
d\tau_{2t} &= \alpha_2 \tau_{2t} dt + \left(1 + \phi \tau_{2t}\right) dB_{2t}, \\
corr\left(dW_t, dB_{1t}\right) &= \varphi_1, \; corr\left(dW_t, dB_{2t}\right) = \varphi_2.
\end{aligned}
$$

We follow Huang and Tauchen (2005) and set $\mu = 0.03$, $\beta_0 = -1.2$, $\beta_1 = 0.04$, $\beta_2 = 1.5$, $\alpha_1 = -0.00137$, $\alpha_2 = -1.386$, $\phi = 0.25$, $\varphi_1 = \varphi_2 = -0.3$. We initialize the two factors at the start of each interval by drawing the persistent factor from its unconditional distribution, $\tau_{10} \sim N\left(0, \frac{-1}{2\alpha_1}\right)$ and by starting the strongly mean-reverting factor at zero.

For integrated beta, the design of our Monte Carlo study follows that of Barndorff-Nielsen and Shephard (2004a), and Dovonon Gonçalves and Meddahi (2013). In particular, we assume that $dX(t) = \sigma(t)\, dW(t)$, with $\sigma(t)\, \sigma'(t) = \Sigma(t)$, where

$$
\Sigma(t) = \begin{pmatrix} \Sigma_{11}(t) & \Sigma_{12}(t) \\ \Sigma_{21}(t) & \Sigma_{22}(t) \end{pmatrix} = \begin{pmatrix} \sigma_1^2(t) & \sigma_{12}(t) \\ \sigma_{21}(t) & \sigma_2^2(t) \end{pmatrix},
$$

and $\sigma_{12}(t) = \sigma_1(t)\, \sigma_2(t)\, \rho(t)$. As Barndorff-Nielsen and Shephard (2004a), we let $\sigma_1^2(t) = \sigma_1^{2(1)}(t) + \sigma_1^{2(2)}(t)$, where for $s = 1, 2$, $d\sigma_1^{2(s)}(t) = -\lambda_s(\sigma_1^{2(s)}(t) - \xi_s)dt + \omega_s \sigma_1^{(s)}(t)\sqrt{\lambda_s} db_s(t)$, where b_i is the i-th component of a vector of standard Brownian motions, independent from W. We let $\lambda_1 = 0.0429$, $\xi_1 = 0.110$, $\omega_1 = 1.346$, $\lambda_2 = 3.74$, $\xi_2 = 0.398$, and $\omega_2 = 1.346$. Our model for $\sigma_2^2(t)$ is the GARCH(1,1) diffusion studied by Andersen and Bollerslev (1998): $d\sigma_2^2(t) = -0.035(\sigma_2^2(t) - 0.636)dt + 0.236\sigma_2^2(t)db_3(t)$. Finally, we follow Barndorff-Nielsen and Shephard (2004), and let $\rho(t) = (e^{2x(t)} - 1)/(e^{2x(t)} + 1)$, where x follows the GARCH diffusion: $dx(t) = -0.03(x(t) - 0.64)dt + 0.118x(t)db_4(t)$.

We simulate data for the unit interval $[0, 1]$. The observed log-price process X is generated using an Euler scheme. We then construct the h-horizon returns $y_i \equiv X_{ih} - X_{(i-1)h}$ based on samples of size $1/h$.

Tables 1 and 2 give the actual rates of 95% confidence intervals of integrated volatility and integrated beta, computed over 10,000 replications. Results are presented for six different samples sizes: $1/h$ 1152, 576, 288, 96, 48, and 12, corresponding to "1.25-minute", "2.5-minute", "5-minute", "15-minute", "half-hour" and "2-hour" returns. In Table 1, for each sample size we have computed the coverage rate by

[3] 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 $x_0 = log(1.5)$.

varying the block size, whereas in Table 2 we summarize results by selecting the optimal block size. We also report results for confidence intervals based on a logarithmic version of the statistic $T_{h,M}$ and its bootstrap version.

In our simulations, bootstrap intervals use 999 bootstrap replications for each of the 10,000 Monte Carlo replications. We consider the studentized (percentile-t) symmetric bootstrap confidence interval method computed at the 95% level.

As for all blocking methods, to implement our bootstrap methods, we need to choose the block size M. We follow Politis and Romano (1999) and Hounyo, Gonçalves and Meddahi (2013) and use the Minimum Volatility Method. Here we describe the algorithm we employ for a two-sided confidence interval.

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

(i) For $M = M_{small}$ to $M = M_{big}$ compute a bootstrap interval for the parameter of interest (integrated volatility or integrated beta) at the desired confidence level, this resulting in endpoints $IC_{M,low}$ and $IC_{M,up}$.

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

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

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

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

One might ask what is a selection of reasonable M_{small} and M_{big}? In our experience, for a sample size $1/h = 1152$, the choices $M_{small} = 1$ and $M_{big} = 12$ usually suffice, for the samples sizes : $1/h =$ 1152, 576, 288, 96, and 48, we have used $M_{small} = 1$ and $M_{big} = 12$. For results in Table 2, we used $l = 2$ in our simulations. Some initial simulations (not recorded here) showed that the actual coverage rate of the confidence intervals using the bootstrap is not sensitive to reasonable choice of l, in particular, for $l = 1, 2, 3$.

Starting with integrated volatility, the Monte Carlo results in Tables 1 and 2 show that for both models (SV1F and SV2F), the asymptotic intervals tend to undercover. The degree of undercoverage is especially large, when sampling is not too frequent. It is also larger for the raw statistics than for the log-based statistics. The SV2F model exhibits overall larger coverage distortions than the

SV1F model, for all sample sizes. When $M = 1$, the Gaussian bootstrap method is equivalent to the wild bootstrap of Gonçalves and Meddahi (2009) that uses the normal distribution as external random variable. One can see that the bootstrap replicates their simulations results. In particular, the Gaussian bootstrap intervals tend to overcover across all models. The actual coverage probabilities of the confidence intervals using the Gaussian bootstrap are typically monotonically decreasing in M, and does not tend to decrease very fast in M for larger values of sample size.

A comparison of the local Gaussian bootstrap with the best existing bootstrap methods for realized volatility[4] shows that, for smaller samples sizes, the confidence intervals based on Gaussian bootstrap are conservative, yielding coverage rates larger than 95% for the SV1F model. The confidence intervals tend to be closer to the desired nominal level for the SV2F than the best bootstrap proposed by Gonçalves and Meddahi (2009). For instance, for SV1F model, the Gaussian bootstrap covers 96.51% of the time when $h^{-1} = 12$ whereas the best bootstrap of Gonçalves and Meddahi (2009) does only 87.42%. These rates decrease to 93.21% and 80.42% for the SV2F model, respectively.

We also consider intervals based on the i.i.d. bootstrap studied by Gonçalves and Meddahi (2009). Despite the fact that the i.i.d. bootstrap does not theoretically provide an asymptotic refinement for two-sided symmetric confidence intervals, it performs well.

While none of the intervals discussed here (bootstrap or asymptotic theory-based) allow for $M = h^{-1}$, we have also studied this setup which is nevertheless an obvious interest in practice. For the SV1F model, results are not very sensitive to the choice of the block size, whereas for the SV2F model coverage rates for intervals using a very large value of block size ($M = h^{-1}$) are systematically much lower than 95% even for the largest sample sizes. When $M = h^{-1}$, the realized volatility R_2 using the blocking approach is the same realized volatility studied by Barndorff-Nielsen and Shephard (2002), but the estimator of integrated quarticity using the blocking approach is $\frac{h^{-1}+2}{h^{-1}} R_2^2$. This means that asymptotically we replace $\int_0^1 \sigma_t^4 dt$ by $\left(\int_0^1 \sigma_t^2 dt\right)^2$, which is only valid under constant volatility. By Cauchy-Schwarz inequality, we have $\left(\int_0^1 \sigma_t^2 dt\right)^2 \leq \int_0^1 \sigma_t^4 dt$, it follows then that we underestimated the asymptotic variance of the realized volatility estimator. This explains the poor performance of the theory based on the blocking approach when the block size is too large. This also confirms the theoretical prediction, which require $M = O(\sqrt{h^{-1}})$ for a good approximation for the probability measure P.

For realized beta, we see that intervals based on the feasible asymptotic procedure using Mykland and Zhang's (2009) blocking approach and the bootstrap tend to be similar for larger sample sizes whereas, at the smaller sample sizes, intervals based on the asymptotic normal distribution are quite severely distorted. For instance, the coverage rate for the feasible asymptotic theory of Mykland and Zhang (2009) when $h^{-1} = 12$ (cf. $h^{-1} = 48$) is only equal to 88.49% (92.86%), whereas it is equal to 95.17% (94.84%), for the Gaussian bootstrap (the corresponding symmetric interval based on the

[4] The wild bootstrap based on Proposition 4.5 of Gonçalves and Meddahi (2009).

pairs bootstrap of Dovonon Gonçalves and Meddahi (2013) yields a coverage rate of 93.59% (93.96%), better than Mykland and Zhang (2009) but worse than the Gaussian bootstrap interval). Our Monte Carlo results also confirm that for a good approximation, a very large block size is not recommended.

Overall, all methods behave similarly for larger sample sizes, in particular the coverage rate tends to be closer to the desired nominal level. The Gaussian bootstrap performance is quite remarkable and outperforms the existing methods, especially for smaller samples sizes ($h^{-1} = 12$ and 48).

7 Empirical results

As a brief illustration, in this section we implement the local Gaussian bootstrap method with real high-frequency financial intraday data, and compare it to the existing feasible asymptotic procedure of Mykland and Zhang (2009). The data consists of transaction log prices of General Electric (GE) shares carried out on the New York Stock Exchange (NYSE) in August 2011. Before analyzing the data we have cleaned the data. For each day, we consider data from the regular exchange opening hours from time stamped between 9:30 a.m. till 4 p.m. Our procedure for cleaning data is exactly identical to that used by Barndorff-Nielsen et al. (2008). We detail in Appendix A the cleaning we carried out on the data.

We implemented the realized volatility estimator of Mykland and Zhang (2009) on returns recorded every S transactions, where S is selected each day so that there are 96 observations a day. This means that on average these returns are recorded roughly every 15 minutes. Table 3 in the Appendix provides the number of transactions per day, and the sample size used. Typically each interval corresponds to about 131 transactions.

This choice is motivated by the empirical study of Hansen and Lunde (2006), who investigate 30 stocks of the Dow Jones Industrial Average, in particular they have presented detailed work for the GE shares. They suggest to use 10 to 15 minutes horizon for liquid assets to avoid the market microstructure noise effect.

Hence the main assumptions underlying the validity of the Mykland and Zhang (2009) block-based method and our new bootstrap method are roughly satisfied and we feel comfortable to implement them on this data.

To implement the realized volatility estimator, we need to choose the block size M. We use the Minimum Volatility Method described above to choose M.

We consider bootstrap percentile-t intervals, computed at the 95% level. The results are displayed in Figure 1 in the appendix in terms of daily 95% confidence intervals (CIs) for integrated volatility. Two types of intervals are presented: our proposed new local Gaussian bootstrap method , and the the feasible asymptotic theory using Mykland and Zhang (2009) blocking approach. The realized volatility estimate R_2 is in the center of both confidence intervals by construction. A comparison of the local Gaussian bootstrap intervals with the intervals based on the feasible asymptotic theory using Mykland

and Zhang (2009) block-based approach suggests that the both types of intervals tend to be similar. The width of these intervals varies through time. However there are instances where the bootstrap intervals are wider than the asymptotic theory-based interval. These days often correspond to days with large estimate of volatility. We have asked whether it will be due to jumps. At this end we have implemented the jumps test using blocked bipower variation of Mykland, Shephard and Sheppard (2012). We have found no evidence of jumps at 5% significance level for these two days. The figures also show a lot of variability in the daily estimate of integrated volatility.

8 Conclusion

This paper proposes a new bootstrap method for statistics that are smooth functions of the realized multivariate volatility matrix based on Mykland and Zhang's (2009) blocking approach. We show how and to what extent the local Gaussianity assumption can be explored to generate a bootstrap approximation. We use Monte Carlo simulations and derive higher order expansions for cumulants to compare the accuracy of the bootstrap and the normal approximations at estimating confidence intervals for integrated volatility and integrated beta. Based on these expansions, we show that at second order the bootstrap matches the cumulants of realized betas-based t-statistics whereas it provides a third-order asymptotic refinement for realized volatility. This is an improvement of the existing bootstrap results. Our new bootstrap method also generalizes the wild bootstrap of Gonçalves and Meddahi (2009). Monte Carlo simulations suggest that the Gaussian bootstrap improves upon the first-order asymptotic theory in finite samples and outperform the existing bootstrap methods for realized volatility and realized betas. An important extension is to prove the validity of the Edgeworth expansions derived here. Another promising extension is to use the bootstrap method for volatility estimator (multipower variation) using the blocking approach in presence of jumps.

Appendix A

This appendix is organized as follows. First, we details the cleaning we carried out on the data. Second, we report simulation results. Finally we report empirical results.

Data Cleaning

In line with Barndorff-Nielsen et al. (2009) we perform the following data cleaning steps:

(i) Delete entries outside the 9:30pm and 4pm time window.

(ii) Delete entries with a quote or transaction price equal to be zero.

(iii) Delete all entries with negative prices or quotes.

(iv) Delete all entries with negative spreads.

(v) Delete entries whenever the price is outside the interval $[bid - 2 * spread \; ; \; ask + 2 * spread]$.

(vi) Delete all entries with the spread greater or equal than 50 times the median spread of that day.

(vii) Delete all entries with the price greater or equal than 5 times the median mid-quote of that day.

(viii) Delete all entries with the mid-quote greater or equal than 10 times the mean absolute deviation from the local median mid-quote.

(ix) Delete all entries with the price greater or equal than 10 times the mean absolute deviation from the local median mid-quote.

We report in Table 1 below, the actual coverage rates for the feasible asymptotic theory approach and for our bootstrap methods. In Table 2 we summarize results using the optimal block size by minimizing confidence interval volatility. Table 3 provides some statistics of GE shares in August 2011.

Table 1: Coverage rates of nominal 95% CI for integrated volatility and integrated beta

	Integrated volatility									Integrated beta	
	SV1F				SV2F						
	Raw		Log		Raw		Log			Raw	
M	CLT	Boot	CLT	Boot	CLT	Boot	CLT	Boot	M	CLT	Boot
					$1/h = 12$						
1	85.44	98.49	90.08	97.86	80.38	96.62	86.17	96.24	2	83.66	95.88
2	85.56	97.31	90.31	96.80	80.43	94.70	86.27	94.73	3	87.63	95.03
3	85.71	96.46	90.84	96.08	80.34	93.77	85.89	93.70	4	89.14	94.83
4	85.88	96.20	90.97	95.93	80.34	92.88	85.52	92.89	6	90.67	94.49
12	86.11	94.84	91.27	94.87	77.66	88.89	81.65	86.97	12	90.44	93.63
					$1/h = 48$						
1	92.04	98.55	93.51	97.71	88.28	97.09	90.93	96.67	3	92.40	95.88
2	92.10	97.28	93.59	96.50	88.13	95.63	91.08	95.48	4	92.69	95.34
4	92.20	96.40	93.80	95.80	88.16	94.55	91.10	94.53	8	92.93	94.69
8	92.33	95.60	93.88	95.18	87.89	93.32	90.33	93.20	12	92.67	93.78
48	92.74	95.06	94.22	95.04	81.83	86.63	82.92	84.57	48	91.63	92.43
					$1/h = 96$						
1	93.35	97.94	94.09	97.10	90.20	97.06	92.10	96.66	3	92.62	95.57
2	93.43	96.78	93.99	96.06	90.37	95.84	92.24	95.67	4	93.13	95.00
4	93.47	95.78	94.03	95.61	90.46	94.70	92.09	94.83	8	93.83	94.84
8	93.50	95.26	94.09	95.32	90.07	93.81	91.75	94.01	12	93.77	94.57
96	93.42	94.80	94.35	94.87	81.93	84.61	82.79	83.60	96	91.94	92.35
					$1/h = 288$						
1	94.57	97.09	94.61	96.25	93.39	97.44	93.96	96.76	3	93.87	95.79
2	94.56	96.00	94.61	95.67	93.51	96.35	93.95	95.95	4	94.72	95.64
4	94.62	95.48	94.67	95.36	93.50	95.57	93.98	95.28	8	94.95	95.43
8	94.55	95.26	94.81	95.19	93.43	95.06	93.82	94.75	12	94.66	94.99
288	94.46	94.78	94.84	94.99	82.43	83.86	83.34	83.53	288	90.04	90.32
					$1/h = 576$						
1	94.53	96.12	94.75	95.84	94.19	96.96	94.49	96.52	3	93.94	95.62
2	94.57	95.53	94.68	95.41	94.17	96.23	94.52	95.78	4	94.46	95.40
4	94.74	95.15	94.70	95.16	94.32	95.59	94.56	95.45	8	94.58	94.87
8	94.67	95.08	94.72	94.96	94.22	95.38	94.46	95.16	12	94.53	94.88
576	94.58	94.85	94.76	94.92	82.01	82.37	82.05	82.32	576	87.07	87.07
					$1/h = 1152$						
1	95.06	96.06	95.16	95.70	94.51	96.52	94.47	95.95	3	94.78	95.93
2	95.13	95.68	95.20	95.65	94.53	95.79	94.47	95.42	4	94.92	95.48
4	95.05	95.49	95.20	95.31	94.42	95.21	94.50	95.11	8	94.88	95.13
8	95.15	95.47	95.18	95.20	94.39	95.03	94.47	94.85	12	94.95	94.87
1152	94.86	94.97	94.83	94.91	82.60	82.73	82.85	82.89	1152	81.68	81.62

Notes: CLT-intervals based on the Normal; Boot-intervals based on our proposed new local Gaussian bootstrap; M is the block size used to compute confidence intervals. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 2: Coverage rates of nominal 95% intervals for integrated volatility and integrated beta using the optimal block size

Integrated volatility

	SV1F									SV2F								
		Raw				Log					Raw				Log			
n	M^*	CLT	iidB	WB	Boot	CLT	iidB	WB	Boot	M^*	CLT	iidB	WB	Boot	CLT	iidB	WB	Boot
12	3.75	86.44	93.66	87.42	96.51	90.79	96.11	88.35	96.07	3.84	80.42	90.82	79.41	93.21	86.70	93.43	80.41	93.35
48	5.37	90.89	94.62	93.98	95.76	92.31	95.45	94.71	95.35	5.75	85.40	92.77	89.98	93.63	87.59	94.15	90.82	93.80
96	5.86	93.01	94.67	94.38	95.29	93.68	95.48	94.85	95.11	5.81	88.94	94.05	92.01	94.28	90.48	94.28	93.11	94.21
288	5.66	94.49	94.70	94.98	95.05	94.60	94.86	94.81	94.96	5.94	93.32	94.85	94.18	95.02	93.62	94.99	94.36	94.90
576	5.82	94.50	94.54	94.51	94.95	94.58	94.65	94.66	94.98	6.06	94.08	94.89	94.47	95.19	94.38	94.90	94.81	95.03
1152	6.01	95.05	94.87	95.13	95.04	95.15	94.85	95.14	95.02	6.25	94.41	94.83	94.56	94.92	94.38	94.92	94.76	94.83

Integrated beta

n	M^*	CLT	PairsB	Boot
12	3.64	88.49	93.59	95.17
48	4.97	92.86	93.96	94.84
96	5.60	93.70	94.98	94.62
288	5.76	94.55	94.75	94.98
576	5.78	94.26	94.72	94.64
1152	5.92	94.77	94.67	94.78

Notes: CLT-intervals based on the Normal; iidB-intervals based on the i.i.d. bootstrap of Gonçalves and Meddahi (2009); WB-wild bootstrap based on Proposition 4.5 of Gonçalves and Meddahi (2009); Boot-intervals based on our proposed new local Gaussian bootstrap; PairsB-intervals based on the pairs bootstrap of Dovonon Gonçalves and Meddahi (2013); M^* is the optimal block size selected by using the Minimum Volatility method. 10,000 Monte Carlo trials with 999 bootstrap replications each.

Table 3: Summary statistics

Days	Trans	n	S
1 Aug	11303	96	118
2 Aug	13873	96	145
3 Aug	13205	96	138
4 Aug	16443	96	172
5 Aug	16212	96	169
8 Aug	18107	96	189
9 Aug	18184	96	190
10 Aug	15826	96	165
11 Aug	15148	96	158
12 Aug	12432	96	130
15 Aug	12042	96	126
16 Aug	10128	96	106
17 Aug	9104	96	95
18 Aug	15102	96	158
19 Aug	11468	96	120
22 Aug	10236	96	107
23 Aug	11518	96	120
24 Aug	10429	96	109
25 Aug	9794	96	102
26 Aug	9007	96	94
29 Aug	10721	96	112
30 Aug	9131	96	96
31 Aug	10724	96	112

"Trans" denotes the number of transactions, n the sample size used to compute the realized volatility, and sampling of every S'th transaction price, so the period over which returns are calculated is roughly 15 minutes.

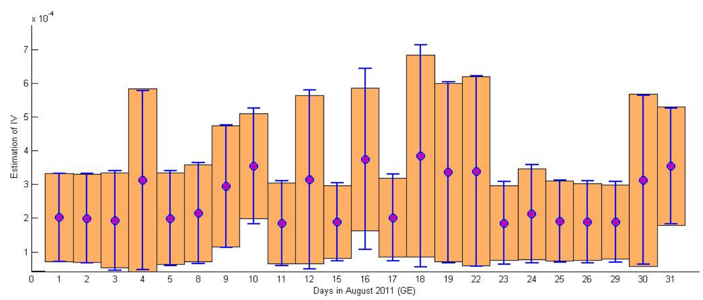

Figure 1: 95% Confidence Intervals (CI's) for the daily $\overline{\sigma^2}$, for each regular exchange opening days in August 2011, calculated using the asymptotic theory of Mykland and Zhang (CI's with bars), and the new wild bootstrap method (CI's with lines). The realized volatility estimator is the middle of all CI's by construction. Days on the x-axis.

Appendix B

This appendix concerns only the case where $d = 1$ (i.e. when the parameter of interest is the integrated volatility). We organized this appendix as follows. First, we introduce some notation. Second, we state Lemmas B.1 and B.2, Theorems B.1 and B.2 and their proofs useful for proofs for the theorem 3.1 and proposition 5.1 presented in the main text. These results are used to obtain the formal Edgeworth expansions through order $O(h)$ for realized volatility. Finally, we prove the Theorem 3.1 and the Propositions 5.1.

Notation

To make for greater comparability, and in order to use some existing results, we have kept the notation from Gonçalves and Meddahi (2009) whenever possible. We introduce some notation, recall that, for any $q > 0$, $\overline{\sigma^q} \equiv \int_0^1 \sigma_u^q du$, and let $\overline{\sigma_{h,M}^q} \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{\sigma_{j,M}^2}{Mh}\right)^{q/2}$, where $\sigma_{j,M}^2 \equiv \int_{(j-1)Mh}^{jMh} \sigma_u^2 du$. We let $\sigma_{q,p} \equiv \frac{\overline{\sigma^q}}{\left(\overline{\sigma^p}\right)^{q/p}}$, when $\overline{\sigma^q}$ is replaced with $\overline{\sigma_{h,M}^q}$ we write $\sigma_{q,p,h,M}$, and $R_q \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{RV_{j,M}}{Mh}\right)^{q/2}$, where $RV_{j,M} = \sum_{i=1}^{M} y_{i+(j-1)M}^2$. We also let $R_{q,p} \equiv \frac{R_q}{(R_p)^{q/p}}$. Recall that $c_{M,q} \equiv E\left(\left(\frac{\chi_M^2}{M}\right)^{q/2}\right)$ with χ_M^2 the standard χ^2 distribution with M degrees of freedom. Note that $c_{M,2} = 1$, $c_{M,4} = \frac{M+2}{M}$, $c_{M,6} = \frac{(M+2)(M+4)}{M^2}$ and $c_{M,8} = \frac{(M+2)(M+4)(M+6)}{M^3}$. It follows by using the definition of $c_{M,q}$ gives in equation (6) and this property of the Gamma function, for all $x > 0$, $\Gamma(x+1) = x\Gamma(x)$.
We follow Gonçalves and Meddahi (2009) and we write

$$T_{h,M} = S_{h,M} \left(\frac{\hat{V}}{V_{h,M}}\right)^{-1/2} = S_{h,M} \left(1 + \sqrt{h} U_{h,M}\right)^{-1/2},$$

where

$$S_{h,M} = \frac{\sqrt{h^{-1}}\left(R_2 - c_{M,2}\overline{\sigma^2}\right)}{\sqrt{V_{h,M}}} \text{ and } U_{h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{V} - V_{h,M}\right)}{V_{h,M}},$$

and $V_{h,M} = Var\left(\sqrt{h^{-1}} R_2\right) = M\left(c_{M,4} - c_{M,2}^2\right)\overline{\sigma_{h,M}^4}$. The proof of Lemma B.1 below relies heavily on the fact that, for any $q > 0$, $\left|RV_{j,M}\right|^{q/2} - c_{M,q}\left|\sigma_{j,M}^2\right|^{q/2}$ are conditionally on σ independent with zero mean since $RV_{j,M} = \sigma_{j,M}^2 \frac{\chi_{j,M}^2}{M}$ where $\frac{\chi_{j,M}^2}{M} \equiv \frac{\sum_{i=1}^{M} \eta_{(j-1)M+i}^2}{M}$ and $\eta_i \sim i.i.d.\ N(0,1)$. We rewrite $R_2 - c_{M,2}\overline{\sigma^2}$ and $\hat{V} - V_{h,M}$ as follows

$$R_2 - c_{M,2}\overline{\sigma^2} = \sum_{j=1}^{1/Mh} \left(RV_{j,M} - c_{M,2}\sigma_{j,M}^2\right),$$

$$\hat{V} - V_{h,M} = M\left(\frac{c_{M,4} - c_{M,2}^2}{c_{M,4}}\right)(Mh)^{-1} \sum_{j=1}^{1/Mh} \left(RV_{j,M}^2 - c_{M,4}\sigma_{j,M}^4\right).$$

Similarly for the bootstrap, we let $T_{h,M}^* = S_{h,M}^* \left(1 + \sqrt{h} U_{h,M}^*\right)^{-1/2}$, where $S_{h,M}^* = \frac{\sqrt{h^{-1}}\left(R_2^* - c_{M,2} R_2\right)}{\sqrt{V^*}}$, $U_{h,M}^* \equiv \frac{\sqrt{h^{-1}}\left(\hat{V}^* - V^*\right)}{V^*}$ and $V^* = Var^*\left(n^{1/2} R_2^*\right)$.

Finally, note that throughout we will use $\sum_{i\neq j\neq k} = \sum_{i\neq j, i\neq k, j\neq k}$, to denote a sum where all indices differ, e.g.

Lemma B.1. *Suppose (1) holds, conditionally on σ, and under Q_h for any $q > 0$, and any $M \geq 1$ such that $M \approx cn^{-\alpha}$ with $\alpha \in [0, 1/2)$, we have*

a1) $E\left(\left|RV_{j,M}\right|^{q/2}\right) = c_{M,q}\left|\sigma^2_{j,M}\right|^{q/2}$,

a2) $V_{h,M} \equiv Var\left(\sqrt{h^{-1}}R_2\right) = M\left(c_{M,4} - c^2_{M,2}\right)\overline{\sigma^4_{h,M}}$,

a3) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^3\right] = (Mh)^2\left(c_{M,6} - 3c_{M,2}c_{M,4} + 2c^3_{M,2}\right)\overline{\sigma^6_{h,M}}$,

a4)

$$\begin{aligned} E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^4\right] &= 3(Mh)^2\left(c_{M,4} - c^2_{M,2}\right)^2\left(\overline{\sigma^4_{h,M}}\right)^2 \\ &+ (Mh)^3\left(c_{M,8} - 3c_{M,2}c_{M,6} + 12c^2_{M,2}c_{M,4} - 6c^4_{M,2} - 3c^2_{M,4}\right)\overline{\sigma^8_{h,M}}, \end{aligned}$$

a5)

$$E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)\left(\hat{V} - V_{h,M}\right)\right] = M\left(\frac{c_{M,4} - c^2_{M,2}}{c_{M,4}}\right)(Mh)\left(c_{M,6} - c_{M,2}c_{M,4}\right)\overline{\sigma^6_{h,M}},$$

a6) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^2\left(\hat{V} - V_{h,M}\right)\right] = M\left(\frac{c_{M,4}-c^2_{M,2}}{c_{M,4}}\right)(Mh)^2\left[\begin{array}{c} c_{M,8} - c^2_{M,4} \\ -2c_{M,2}c_{M,6} + c^2_{M,2}c_{M,4}\end{array}\right]\overline{\sigma^8_{h,M}}$,

a7)

$$\begin{aligned} E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^3\left(\hat{V} - V_{h,M}\right)\right] &= 3M(Mh)^2\frac{\left(c_{M,4} - c^2_{M,2}\right)^2\left(c_{M,6} - c_{M,2}c_{M,4}\right)}{c_{M,4}}\left(\overline{\sigma^4_{h,M}}\right)\left(\overline{\sigma^4_{h,M}}\right) \\ &+ 384h^3\overline{\sigma^{10}_{h,M}} \\ &= 3M(Mh)^2\frac{\left(c_{M,4} - c^2_{M,2}\right)^2\left(c_{M,6} - c_{M,2}c_{M,4}\right)}{c_{M,4}}\left(\overline{\sigma^4_{h,M}}\right)\left(\overline{\sigma^4_{h,M}}\right) \\ &+ O\left(h^3\right) \text{ as } h \to 0, \end{aligned}$$

a8) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^4\left(\hat{V} - V_{h,M}\right)\right] =$

$$(Mh)^3 M\frac{c_{M,4}-c^2_{M,2}}{c_{M,4}}\left[\begin{array}{c} 4\left(c_{M,6} - 3c_{M,2}c_{M,4} + 2c^3_{M,2}\right)\left(c_{M,6} - c_{M,2}c_{M,4}\right)\left(\overline{\sigma^6_{h,M}}\right)^2 \\ +6\left(c_{M,8} - c^2_{M,4} - 2c_{M,2}c_{M,6} + 2c^2_{M,2}c_{M,4}\right)\left(c_{M,4} - c^2_{M,2}\right)\left(\overline{\sigma^4_{h,M}}\right)\left(\overline{\sigma^8_{h,M}}\right)\end{array}\right]$$

$+ O\left(h^4\right)$ *as* $h \to 0$,

a9) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)\left(\hat{V} - V_{h,M}\right)^2\right] = O\left(h^2\right)$ *as* $h \to 0$,

a10) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^2\left(\hat{V} - V_{h,M}\right)^2\right] =$

$(Mh)^2 M \frac{\left(c_{M,4}-c_{M,2}^2\right)^2}{c_{M,4}^2}\left(\left(c_{M,4} - c_{M,2}^2\right)\left(c_{M,8} - c_{M,4}^2\right)\left(\overline{\sigma_{h,M}^4}\right)\left(\overline{\sigma_{h,M}^8}\right) + 2\left(c_{M,6} - c_{M,2}c_{M,4}\right)^2\left(\overline{\sigma_{h,M}^6}\right)^2\right)$

$+ O\left(h^3\right)$ *as* $h \to 0$,

a11) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^3\left(\hat{V} - V_{h,M}\right)^2\right] = O\left(h^3\right)$ *as* $h \to 0$,

a12) $E\left[\left(R_2 - c_{M,2}\overline{\sigma^2}\right)^4\left(\hat{V} - V_{h,M}\right)^2\right] =$

$(Mh)^3 M \frac{\left(c_{M,4}-c_{M,2}^2\right)^2}{c_{M,4}^2}\left[\begin{array}{c} 3\left(c_{M,4} - c_{M,2}^2\right)^2\left(c_{M,8} - c_{M,4}^2\right)\left(\overline{\sigma_{h,M}^4}\right)^2\left(\overline{\sigma_{h,M}^8}\right) \\ +12\left(c_{M,4} - c_{M,2}^2\right)^2\left(c_{M,6} - c_{M,2}c_{M,4}\right)^2\left(\overline{\sigma_{h,M}^6}\right)^2\left(\overline{\sigma_{h,M}^4}\right) \end{array}\right]$

$+ O\left(h^4\right)$ *as* $h \to 0$.

Lemma B.2. *Suppose (1) holds, conditionally on* σ, *and under* Q_h *for any* $M \geq 1$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in (1/2, 1]$, *we have*

$$\begin{aligned}
E\left(S_{h,M}\right) &= 0, \\
E\left(S_{h,M}^2\right) &= 1, \\
E\left(S_{h,M}^3\right) &= \sqrt{h}B_{1,M}\sigma_{6,4,h,M}, \\
E\left(S_{h,M}^4\right) &= 3 + hB_{2,M}\sigma_{8,4,h,M}, \\
E\left(S_{h,M}U_{h,M}\right) &= A_{1,M}\sigma_{6,4,h,M}, \\
E\left(S_{h,M}^2U_{h,M}\right) &= \sqrt{h}A_{2,M}\sigma_{8,4,h,M},
\end{aligned}$$

and as $h \to 0$ *we have,*

$$\begin{aligned}
E\left(S_{h,M}^3U_{h,M}\right) &= A_{3,M}\sigma_{6,4,h,M} + O\left(h\right), \\
E\left(S_{h,M}^4U_{h,M}\right) &= \sqrt{h}[D_{1,M}\sigma_{8,4,h,M} + D_{2,M}\sigma_{6,4,h,M}^2] + O\left(h^{3/2}\right), \\
E\left(S_{h,M}U_{h,M}^2\right) &= O\left(h^{1/2}\right), \\
E\left(S_{h,M}^3U_{h,M}^2\right) &= O\left(h^{1/2}\right), \\
E\left(S_{h,M}^2U_{h,M}^2\right) &= [C_{1,M}\sigma_{8,4,h,M} + C_{2,M}\sigma_{6,4,h,M}^2] + O\left(h\right), \\
E\left(S_{h,M}^4U_{h,M}^2\right) &= [E_{1,M}\sigma_{8,4,h,M} + E_{2,M}\sigma_{6,4,h,M}^2] + O\left(h\right),
\end{aligned}$$

where,

$$
\begin{aligned}
A_{1,M} &= \frac{1}{\sqrt{M}} \frac{c_{M,6} - c_{M,2}c_{M,4}}{c_{M,4}\left(c_{M,4} - c_{M,2}^2\right)^{1/2}} = \frac{2\sqrt{2}}{M}, \\
B_{1,M} &= \sqrt{M} \frac{\left(c_{M,6} - 3c_{M,2}c_{M,4} + 2c_{M,2}^3\right)}{\left(c_{M,4} - c_{M,2}^2\right)^{3/2}} = 2\sqrt{2}, \\
A_{2,M} &= \frac{c_{M,8} - c_{M,4}^2 - 2c_{M,2}c_{M,6} + 2c_{M,2}^2 c_{M,4}}{c_{M,4}\left(c_{M,4} - c_{M,2}^2\right)} = \frac{12}{M}, \\
B_{2,M} &= M \frac{c_{M,8} - 4c_{M,2}c_{M,6} + 12c_{M,2}^2 c_{M,4} - 6c_{M,2}^4 - 3c_{M,4}^2}{\left(c_{M,4} - c_{M,2}^2\right)^2} = 12, \\
C_{1,M} &= \frac{c_{M,8} - c_{M,4}^2}{c_{M,4}^2 M} = \frac{24 + 8M}{M^2(2+M)},
\end{aligned}
$$

with $A_{3,M} = 3A_{1,M}$, $C_{2,M} = 2A_{1,M}^2$, $D_{1,M} = 6A_{2,M}$, $D_{2,M} = 4A_{1,M}B_{1,M}$, $E_{1,M} = 3C_{1,M}$ *and* $E_{2,M} = 12A_{1,M}^2$.

Remark 4 The bootstrap analogue of Lemma B.1 replaces $RV_{j,M}$ with $RV_{j,M}^*$, $\sigma_{j,M}^2$ with $RV_{j,M}$, and $\overline{\sigma_{h,M}^q}$ with R_q, the bootstrap analogue of Lemma B.2 replaces $\sigma_{q,p,h,M}$ with $R_{q,p}$.

Theorem B.1. ***(Cumulants of*** $T_{h,M}$***)*** *Consider DGP (1) and suppose assumption H holds. Then for any* $q > 0$, *and any* $M \geq 1$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$, $\overline{\sigma_{h,M}^q} = \overline{\sigma_h^q}$ *and* $\overline{\sigma^q} - \overline{\sigma_{h,M}^q} = o_P(\sqrt{h})$, *conditionally on* σ *and under* Q_h, *it follows that as* $h \to 0$,

$$
\begin{aligned}
\kappa_1(T_{h,M}) &= \sqrt{h}\kappa_1 + o(h) \text{ with } \kappa_1 = -\frac{A_{1,M}}{2}\sigma_{6,4}; \\
\kappa_2(T_{h,M}) &= 1 + h\kappa_2 + o(h) \text{ with } \kappa_2 = (C_{1,M} - A_{2,M})\sigma_{8,4} + \frac{7}{4}A_{1,M}^2\sigma_{6,4}^2; \\
\kappa_3(T_{h,M}) &= \sqrt{h}\kappa_3 + o(h) \text{ with } \kappa_3 = (B_{1,M} - 3A_{1,M})\sigma_{6,4}; \\
\kappa_4(T_{h,M}) &= h\kappa_4 + o(h) \text{ with } \kappa_4 = \left(B_{2,M} + 3C_{1,M} - 6A_{2,M}^2\right)\sigma_{8,4} + \left(18{A_{1,M}}^2 - 6A_{1,M}B_{1,M}\right)\sigma_{6,4}^2.
\end{aligned}
$$

Note that $A_{1,M}$, $A_{2,M}$, $B_{1,M}$, $B_{2,M}$, *and* $C_{1,M}$, *are as in Lemma B.2, and* $A_{3,M} = 3A_{1,M}$, $C_{2,M} = 2A_{1,M}^2$, $D_{1,M} = 6A_{2,M}$, $D_{2,M} = 4A_{1,M}B_{1,M}$, $E_{1,M} = 3C_{1,M}$ *and* $E_{2,M} = 12A_{1,M}^2$.

Theorem B.2. ***(Bootstrap Cumulants of*** $T_{h,M}^*$***)*** *Consider DGP (1) and suppose (5) holds. Let* $M \geq 1$ *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$, *under assumption H, conditionally on* σ, *it follows that as* $h \to 0$

$$
\begin{aligned}
\kappa_1^*\left(T_{h,M}^*\right) &= \sqrt{h}\kappa_{1,h,M}^* + o(h) \text{ with } \kappa_{1,h,M}^* = -\frac{A_{1,M}}{2}R_{6,4}; \\
\kappa_2^*\left(T_{h,M}^*\right) &= 1 + h\kappa_{2,h,M}^* + o(h) \text{ with } \kappa_{2,h,M}^* = (C_{1,M} - A_{2,M})R_{8,4} + \frac{7}{4}A_{1,M}^2R_{6,4}^2; \\
\kappa_3^*\left(T_{h,M}^*\right) &= \sqrt{h}\kappa_{3,h,M}^* + o(h) \text{ with } \kappa_{3,h,M}^* = (B_{1,M} - 3A_{1,M})R_{6,4}; \\
\kappa_4^*\left(T_{h,M}^*\right) &= h\kappa_{4,h,M}^* + o(h) \text{ with } \kappa_{4,h,M}^* = \left(B_{2,M} + 3C_{1,M} - 6A_{2,M}^2\right)R_{8,4} + \left(18{A_{1,M}}^2 - 6A_{1,M}B_{1,M}\right)R_{6,4}^2.
\end{aligned}
$$

Note that $A_{1,M}$, $A_{2,M}$, B_1, M, $B_{2,M}$, *and* C_1, M, *are as in Lemma B.2, and* $A_{3,M} = 3A_{1,M}$, $C_{2,M} = 2A_{1,M}^2$, $D_{1,M} = 6A_{2,M}$, $D_{2,M} = 4A_{1,M}B_{1,M}$, $E_{1,M} = 3C_{1,M}$ *and* $E_{2,M} = 12A_{1,M}^2$.

Proof of Lemma B.1 $a1)$ follows from $RV_{j,M}=\sum_{i=1}^{M}y_{(j-1)M+i}^{2}=\frac{\chi_{j,M}^{2}}{M}\sigma_{j,M}^{2}$, where $\frac{\chi_{j,M}^{2}}{M}\equiv\frac{\sum_{i=1}^{M}\eta_{(j-1)M+i}^{2}}{M}$ and $\eta_i\sim i.i.d.\ N(0,1)$. For $a2)$ note that $R_2=\sum_{j=1}^{1/Mh}RV_{j,M}$, where $RV_{j,M}$ is conditional on σ independent with $Var\left(RV_{j,M}\right)=\left(c_{M,4}-c_{M,2}^{2}\right)\sigma_{j,M}^{4}$, with $\sigma_{j,M}^{4}=\left(\sigma_{j,M}^{2}\right)^{2}$. It follows that,

$$
\begin{aligned}
Var\left(\sqrt{h^{-1}}R_2\right) &= h^{-1}\sum_{j=1}^{1/Mh}Var\left(RV_{j,M}\right)\\
&= h^{-1}\left(c_{M,4}-c_{M,2}^{2}\right)\sum_{j=1}^{1/Mh}\sigma_{j,M}^{4}\\
&= M\left(c_{M,4}-c_{M,2}^{2}\right)\overline{\sigma_{h,M}^{4}}.
\end{aligned}
$$

To prove the remaining results we follow the same structure of proofs as Gonçalves and Meddahi (2009). Here $c_{M,4}$ plays the role of $\mu_q=E\left(|\eta|^q\right)$, where $\eta\sim i.i.d.\ N(0,1)$ in Gonçalves and Meddahi (2009) and $RV_{j,M}$ plays the role of r_i^2 in Gonçalves and Meddahi (2009).

Proof of Lemma B.2 Results follow immediately by using Lemma B.1 given the definitions of $S_{h,M}$, $U_{h,M}$.

Proof of Theorem B.1 The first four cumulants of $T_{h,M}$ are given by (e.g., Hall, 1992, p.42):

$$
\begin{aligned}
\kappa_1\left(T_{h,M}\right) &= E\left(T_{h,M}\right),\\
\kappa_2\left(T_{h,M}\right) &= E\left(T_{h,M}^2\right)-\left(E\left(T_{h,M}\right)\right)^2,\\
\kappa_3\left(T_{h,M}\right) &= E\left(T_{h,M}^3\right)-3E\left(T_{h,M}^2\right)E\left(T_{h,M}\right)+2(E\left(T_{h,M}\right))^3,\\
\kappa_4\left(T_{h,M}\right) &= E\left(T_{h,M}^4\right)-4E\left(T_{h,M}^3\right)E\left(T_{h,M}\right)-3(E\left(T_{h,M}^2\right))^2+12E\left(T_{h,M}^2\right)\left(E\left(T_{h,M}\right)\right)^2\\
&- 6(E\left(T_{h,M}\right))^4.
\end{aligned}
$$

Our goal is to identify the terms of order up to $O(h)$ in the asymptotic expansions of these four cumulants. We will first provide asymptotic expansions through order $O(h)$ for the first four moments of $T_{h,M}$ by using a Taylor expansion. For a fixed value k, a second-order Taylor expansion of $f(x)=(1+x)^{-k/2}$ around 0 yields $f(x)=1-\frac{k}{2}x+\frac{k}{4}(\frac{k}{2}+1)x^2+O(x^3)$. We have that for any fixed integer k,

$$
\begin{aligned}
T_{h,M}^k &= S_{h,M}^k\left(1+\sqrt{h}U_{h,M}\right)^{-k/2}+O(h^{3/2}),\\
&= S_{h,M}^k-\frac{k}{2}\sqrt{h}S_{h,M}^kU_{h,M}+\frac{k}{4}(\frac{k}{2}+1)hS_{h,M}U_{h,M}^2+O(h^{3/2}).
\end{aligned}
$$

For $k=1,\cdots,4$, the moments of $T_{h,M}^k$ up to order $O(h^{3/2})$ are given by

$$
\begin{aligned}
E\left(T_{h,M}\right) &= 0-\frac{\sqrt{h}}{2}E\left(S_{h,M}U_{h,M}\right)+\frac{3}{8}hE\left(S_{h,M}U_{h,M}^2\right)\\
E\left(T_{h,M}^2\right) &= 1-\sqrt{h}E\left(S_{h,M}^2U_{h,M}\right)+hE\left(S_{h,M}^2U_{h,M}^2\right)\\
E\left(T_{h,M}^3\right) &= E\left(S_{h,M}^3\right)-\sqrt{h}\frac{3}{2}E\left(S_{h,M}^3U_{h,M}\right)+\frac{15}{8}hE\left(S_{h,M}^3U_{h,M}^2\right)\\
E\left(T_{h,M}^4\right) &= E\left(S_{h,M}^4\right)-2\sqrt{h}E\left(S_{h,M}^4U_{h,M}\right)+3hE\left(S_{h,M}^4U_{h,M}^2\right).
\end{aligned}
$$

where we used $E\left(S_{h,M}\right)=0$, and $E\left(S_{h,M}^2\right)=1$. By Lemma B.2 in Appendix B, we have that

$$
\begin{aligned}
E\left(T_{h,M}\right) &= \sqrt{h}\left(-\frac{A_{1,M}}{2}\sigma_{6,4,h}\right) + O(h^{3/2}),\\
E\left(T_{h,M}^2\right) &= 1 + \sqrt{h}\left(\left(C_{1,M} - A_{2,M}\right)\sigma_{8,4} + C_{2,M}\sigma_{6,4,h}^2\right) + O(h^2)\\
E\left(T_{h,M}^3\right) &= \sqrt{h}\left(\left(B_{1,M} - \frac{3}{2}A_{3,M}\right)\sigma_{6,4,h}\right) + O(h^{3/2})\\
E\left(T_{h,M}^4\right) &= 3 + h\left(\left(B_{2,M} - 2D_{1,M} + 3E_{1,M}\right)\sigma_{8,4,h} + \left(3E_{2,M} - 2D_{2,M}\right)\sigma_{6,4,h}^2\right) + O(h^2).
\end{aligned}
$$

Thus $\kappa_1\left(T_{h,M}\right) = \sqrt{h}\left(-\frac{A_{1,M}}{2}\sigma_{6,4,h}\right) + O(h^{3/2}) = \sqrt{h}\left(-\frac{A_{1,M}}{2}\sigma_{6,4}\right) + o(h^{3/2})$, since under Assumption H, Barndorff-Nielsen and Shephard (2004b) showed that $\overline{\sigma^q} - h^{1-q/2}\sum_{s=1}^{1/h}\left(\int_{(s-1)h}^{sh}\sigma_u^2 du\right)^{q/2} = o_P(\sqrt{h})$.

Next we show that under Q_h, and given the defininion of $\overline{\sigma_{h,M}^q}$, we have $\overline{\sigma_{h,M}^q} = h^{1-q/2}\sum_{s=1}^{1/h}\left(\int_{(s-1)h}^{sh}\sigma_u^2 du\right)^{q/2}$.

Note that, for any positive integer M, given the definitions of $\overline{\sigma_{h,M}^q}$ and $\sigma_{j,M}^2$, we can write

$$
\begin{aligned}
\overline{\sigma_{h,M}^q} &= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}\left(\sigma_{j,M}^2\right)^{q/2}\\
&= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}\left(\int_{(j-1)Mh}^{jMh}\sigma_u^2 du\right)^{q/2},
\end{aligned}
$$

using the fact that under Q_h, we have $\sigma_{j,M}^2 \equiv \int_{(j-1)Mh}^{jMh}\sigma_u^2 du = Mh\sigma_{(j-1)Mh}^2 > 0$, it follows that

$$
\begin{aligned}
\overline{\sigma_{h,M}^q} &= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}\left(Mh\sigma_{(j-1)Mh}^2\right)^{q/2}\\
&= (Mh)^{1-q/2}\sum_{j=1}^{1/Mh}M^{q/2}\left(h\sigma_{(j-1)Mh}^2\right)^{q/2}\\
&= h^{1-q/2}\sum_{j=1}^{1/Mh}M\left(h\sigma_{(j-1)Mh}^2\right)^{q/2}\\
&= h^{1-q/2}\sum_{j=1}^{1/Mh}\sum_{i=1}^{M}\left(h\sigma_{(j-1)Mh}^2\right)^{q/2}\\
&= h^{1-q/2}\sum_{j=1}^{1/Mh}\sum_{i=1}^{M}\left(\int_{((j-1)M+i-1)h}^{((j-1)M+i)h}\sigma_u^2 du\right)^{q/2}\\
&= h^{1-q/2}\sum_{s=1}^{1/h}\left(\int_{(s-1)h}^{sh}\sigma_u^2 du\right)^{q/2}.
\end{aligned}
$$

Thus $\overline{\sigma^q_{h,M}} = \overline{\sigma^q_h}$, this proves the first result. The remaining results follow similarly.
Proof of Theorem B.2 See the proof of Theorem B.1 and Remark 4.

Proofs of Theorem 3.1, and Proposition 5.1.

Proof of Theorem 3.1 Given that $T_{h,M} \overset{d}{\to} N(0,1)$, it suffices that $T^*_{h,M} \overset{d^*}{\to} N(0,1)$ in probability under Q_h. Let

$$H^*_{h,M} = \frac{\sqrt{h^{-1}}\left(R^*_2 - E^*\left(R^*_2\right)\right)}{\sqrt{V^*}},$$

and note that

$$T^*_{h,M} = H^*_{h,M}\sqrt{\frac{V^*}{\hat{V}^*}}.$$

The proof contains two steps.
Step 1 We show that $H^*_{h,M} \overset{d^*}{\to} N(0,1)$ in probability under Q_h.
Step 2 We show that $\hat{V}^* \overset{P^*}{\to} V^*$ in probability under Q_h.
For step 1, we can write

$$H^*_{h,M} = \sum_{j=1}^{1/Mh} z^*_j,$$

where

$$z^*_j = \frac{\sqrt{h^{-1}}\left(RV^*_{j,M} - E^*\left(RV^*_{j,M}\right)\right)}{\sqrt{V^*}}$$

with $E^*\left(\sum_{j=1}^{1/Mh} z^*_j\right) = 0$, and $Var^*\left(\sum_{j=1}^{1/Mh} z^*_j\right) = 1$.

Since $z^*_1, \ldots, z^*_{1/Mh}$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta > 0$ and for some constant $C > 0$ (which changes from line to line),

$$\sup_{x\in\Re}\left|P^*\left(H^*_{h,M} \leq x\right) - \Phi(x)\right| \leq C\sum_{j=1}^{1/Mh} E^*\left|z^*_j\right|^{2+\delta},$$

which converges to zero in probability for any $M \geq 1$ such that $M \approx ch^{-\alpha}$ with $\alpha \in [0, 1/2)$, as $h \to 0$. Indeed, we have that

$$\begin{aligned}
\sum_{j=1}^{1/Mh} E^*\left|z^*_j\right|^{2+\delta} &= \sum_{j=1}^{1/Mh} E^*\left|\frac{\sqrt{h^{-1}}\left(RV^*_{j,M} - E^*\left(RV^*_{j,M}\right)\right)}{\sqrt{V^*}}\right|^{2+\delta} \\
&\leq 2V^{*-\frac{(2+\delta)}{2}} h^{\frac{-(2+\delta)}{2}} \sum_{j=1}^{1/Mh} E^*\left|RV^*_{j,M}\right|^{2+\delta} \\
&= 2V^{*-\frac{(2+\delta)}{2}} h^{\frac{-(2+\delta)}{2}} E^*\left|\frac{\sum_{i=1}^{M}\eta^2_{(j-1)M+i}}{M}\right|^{2+\delta} \sum_{j=1}^{1/Mh}\left|RV_{j,M}\right|^{2+\delta},
\end{aligned}$$

where the inequality follows from the C_r and the Jensen inequalities. Then, given the definitions of $c_{M,2(2+\delta)}$ and $R_{2(2+\delta)}$, we can write

$$\begin{aligned}\sum_{j=1}^{1/Mh} E^*\left|z_j^*\right|^{2+\delta} &\leq 2V^{*-\frac{(2+\delta)}{2}}c_{M,2(2+\delta)}M^{1+\delta}h^{\frac{\delta}{2}}R_{2(2+\delta)}\\ &\leq CV^{*-\frac{(2+\delta)}{2}}c_{M,2(2+\delta)}^2h^{\frac{\delta}{2}-\alpha(1+\delta)}\frac{1}{c_{M,2(2+\delta)}}R_{2(2+\delta)}\\ &= O_p\left(h^{\frac{\delta}{2}-\alpha(1+\delta)}c_{M,2(2+\delta)}^2\right)\\ &= o_p(1).\end{aligned}$$

Note that for any $\delta > 0$ and $\alpha \in [0, 1/2)$ we have $\frac{\delta}{2} - \alpha(1+\delta) > 0$. Results follow since as $h \to 0$, $V^* \xrightarrow{P} 2\overline{\sigma^4} > 0$, and we have $\frac{1}{c_{M,2(2+\delta)}}R_{2(2+\delta)} \xrightarrow{P} \overline{\sigma^{2(2+\delta)}} = O(1)$, and $c_{M,2(2+\delta)} \to 1$.

For step 2, we show that $Bias^*\left(\widehat{V^*}\right) \xrightarrow{Q_h} 0$ and $Var^*\left(\widehat{V^*}\right) \xrightarrow{Q_h} 0$.
We have that

$$\begin{aligned}Bias^*\left(\widehat{V^*}\right) &= E^*\left(\widehat{V^*}\right) - V^*\\ &= M\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)(Mh)^{-1}\sum_{j=1}^{1/Mh}E^*\left(RV_{j,M}^{2*} - c_{M,4}RV_{j,M}^2\right)\\ &= 0,\end{aligned}$$

we also have,

$$\begin{aligned}Var^*\left(\widehat{V^*}\right) &= E^*\left(\widehat{V^*}-V^*\right)^2 - \left(E^*\left(\widehat{V^*}-V^*\right)\right)^2\\ &= M^2\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)^2(Mh)^{-2}E^*\left(\sum_{j=1}^{1/Mh}\left(RV_{j,M}^{2*}-c_{M,4}RV_{j,M}^2\right)\right)^2\\ &= M^2\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)^2(Mh)^{-2}\sum_{j=1}^{1/Mh}RV_{j,M}^4E^*\left(\left(\frac{\chi_{j,M}^2}{M}\right)^2 - c_{M,4}\right)^2,\end{aligned}$$

then, given the definitions of $c_{M,2}, c_{M,4}$, $c_{M,8}$ and R_8, we can write

$$\begin{aligned}Var^*\left(\widehat{V^*}\right) &= M^2\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)^2(Mh)^{-2}\left(c_{M,8}-c_{M,4}^2\right)\sum_{j=1}^{1/Mh}RV_{j,M}^4\\ &= M^2\left(\frac{c_{M,4}-c_{M,2}^2}{c_{M,4}}\right)^2(Mh)\left(c_{M,8}-c_{M,4}^2\right)R_8\\ &= h\left(\frac{2M}{M+2}\right)^2\frac{(M+2)(M+4)(M+6)-M(M+2)^2}{M^2}R_8\\ &= O_{Q_h}(Mh)\\ &= o_{Q_h}(1)\text{ as } Mh \to 0.\end{aligned}$$

Finally results follow in probability under P, by using Theorem 2.1.

Proof of Proposition 5.1 This follows from Theorem B.1 and B.2, given that conditionally on σ for any $q > 0$, $\frac{1}{c_{M,q}}R_q \to \overline{\sigma^q}$ in probability under Q_h and P (see Section 4.1 of Myklang and Zhang

(2009)). For any $p, q > 1$, $\lim_{M\to\infty} \frac{c_{M,q}}{(c_{M,p})^{q/p}} = 1$.

Appendix C

This appendix concerns the multivariate case where the parameter of interest is the integrated beta.

Appendix C.1. Asymptotic expansions of the cumulants of $T_{\beta,h,M}$

Notation

We introduce some notation.

$$T_{\beta,h,M} = S_{\beta,h,M}\left(\frac{\hat{V}_\beta}{V_{\beta,h,M}}\right)^{-1/2} = S_{\beta,h,M}\left(1+\sqrt{h}U_{h,M}\right)^{-1/2},$$

where

$$S_{\beta,h,M} = \frac{\sqrt{h^{-1}}\left(\hat{\beta}_{lk}-\beta_{lk}\right)}{\sqrt{V_{\beta,h,M}}} \text{ and } U_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{V}_{\beta,h,M}-V_{\beta,h,M}\right)}{V_{\beta,h,M}},$$

and $V_{\beta,h,M} \equiv Var\left(\sqrt{h^{-1}}\left(\hat{\beta}_{lk}-\beta_{lk}\right)\right) = \frac{M}{M-2}\sum_{j=1}^{1/Mh} Mh\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)$. We let

$$U_{1,\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\sum_{j=1}^{1/Mh}\left(V_{1,(j)} - E\left(V_{1,(j)}\right)\right)}{V_{\beta,h,M}} \text{ and } U_{2,\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\sum_{j=1}^{1/Mh}\left(V_{2,(j)} - E\left(V_{2,(j)}\right)\right)}{V_{\beta,h,M}},$$

where

$$V_{1,(j)} = \frac{M^2h}{M-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right), \text{ and}$$

$$V_{2,(j)} = \frac{M^2h}{M-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^{2}.$$

We also let for any $q > M$, $R_{\beta,q} \equiv Mh\sum_{j=1}^{1/Mh}\left(\frac{M}{M-1}\right)^{\frac{q}{2}}\frac{1}{b_{M,q}c_{M-1,q}}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{q}{2}}$, where the definition of $c_{M,q}$ is given in equation (6), and for any $q > M$, we have $b_{M,q} \equiv E\left(\left(\frac{M}{\chi_M^2}\right)^{\frac{q}{2}}\right) = \left(\frac{M}{2}\right)^{\frac{q}{2}}\frac{\Gamma\left(\frac{M}{2}-\frac{q}{2}\right)}{\Gamma\left(\frac{M}{2}\right)}$, where χ_M^2 is the standard χ^2 distribution with M degrees of freedom. Note that $b_{M,2} = \frac{M}{M-2}$, $b_{M,4} = \frac{M^2}{(M-2)(M-4)}$, and $b_{M,6} = \frac{M^3}{(M-2)(M-4)(M-6)}$. It follows by using the definition of $b_{M,q}$ and this property of the Gamma function, for all $x > 0$, $\Gamma(x+1) = x\Gamma(x)$. Finally we denote by $y_{k(j)} = \left(y_{k,1+(j-1)M}, \cdots, y_{k,Mj}\right)'$, the M returns of asset k observed within the block j.

Similarly for the bootstrap, we let $T^*_{\beta,h,M} = S^*_{\beta,h,M}\left(1+\sqrt{h}U^*_{\beta,h,M}\right)^{-1/2}$, where $S^*_{\beta,h,M} = \frac{\sqrt{h^{-1}}(\hat{\beta}^*_{lk}-\hat{\beta}_{lk})}{\sqrt{V^*_\beta}}$,

$U^*_{\beta,h,M} \equiv \frac{\sqrt{h^{-1}}\left(\hat{V}^*_{\beta,h,M} - V^*_{\beta,h,M}\right)}{V^*_{\beta,h,M}}$ and $V^*_{\beta,h,M} = Var^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})\right)$. We also let

$$U^*_{1,\beta,h,M} \equiv \frac{\sqrt{h^{-1}} \sum_{j=1}^{1/Mh} \left(V^*_{1,(j)} - E^*\left(V^*_{1,(j)}\right)\right)}{V^*_{\beta,h,M}} \quad \text{and} \quad U^*_{2,\beta,h,M} \equiv \frac{\sqrt{h^{-1}} \sum_{j=1}^{1/Mh} \left(V^*_{2,(j)} - E\left(V^*_{2,(j)}\right)\right)}{V^*_{\beta,h,M}}$$

where

$V^*_{1,(j)} \equiv \frac{M^2 h}{M-1}\left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-1}\left(\sum_{i=1}^{M} u^{*2}_{i+(j-1)M}\right)$ and

$V^*_{2,(j)} \equiv \frac{M^2 h}{M-1}\left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-2}\left(\sum_{i=1}^{M} y^*_{k,i+(j-1)M} u^*_{i+(j-1)M}\right)^2$.

Finally we let $y^*_{k(j)} = \left(y^*_{k,1+(j-1)M}, \cdots, y^*_{k,Mj}\right)'$.

Auxiliary Lemmas

Lemma C.3. *Suppose (1) and (2) hold. Then, we have that*

$$\hat{V}_{\beta,h,M} = \sum_{j=1}^{1/Mh} V_{1,(j)} - \sum_{j=1}^{1/Mh} V_{2,(j)}.$$

Lemma C.4. *Suppose (1) and (2) hold with W independent of Σ. Then, conditionally on Σ, and under Q_h, we have for any integer M such that $M \approx ch^{-\alpha}$ with $\alpha \in [0, 1/2)$,*

a1) $E\left(V_{1,(j)}\right) = \frac{M^3 h}{(M-1)(M-2)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)$, *for $M > 2$;*

a2) $E\left(V^2_{1,(j)}\right) = \frac{M^5(M+2)}{(M-1)^2(M-2)(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a3) $E\left(V_{2,(j)}\right) = \frac{M^2 h}{(M-1)(M-2)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)$, *for $M > 2$;*

a4) $E\left(V^2_{2,(j)}\right) = \frac{3M^4}{(M-1)^2(M-2)(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a5) $E\left(V_{1,(j)} V_{2,(j)}\right) = \frac{M^4(M+2)}{(M-1)^2(M-2)(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a6) $Var\left(V_{1,(j)}\right) = \frac{4M^5}{(M-1)(M-2)^2(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a7) $Var\left(V_{2,(j)}\right) = \frac{2M^4}{(M-1)(M-2)^2(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a8) $Cov\left(V_{1,(j)}, V_{2,(j)}\right) = \frac{4M^4}{(M-1)(M-2)^2(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$;*

a9) $Var\left(V_{1,(j)} - V_{2,(j)}\right) = \frac{2M^5(2M-3)}{(M-1)(M-2)^2(M-4)} h^2 \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2$, *for $M > 4$.*

Lemma C.5. *Suppose (1) and (2) hold with W independent of Σ. Then, conditionally on Σ, and under Q_h, let $M > 4$ such that $M \approx ch^{-\alpha}$ with $\alpha \in [0, 1/2)$,*

a1) $E\left(\hat{V}_{\beta,h,M}\right) = V_{\beta,h,M}$;

a2) $Var\left(\hat{V}_{\beta,h,M}\right) = \frac{2M^4(2M-3)}{(M-1)(M-2)^2(M-4)} h \left(Mh \sum_{j=1}^{1/Mh} \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}} \right)^2 \right)^2 \right)$;

a3) $\hat{V}_{\beta,h,M} - V_{\beta,h,M} \to 0$ *in probability;*

a4) $V_{\beta,h,M} \to V_\beta$.

Lemma C.6. *Suppose (1) and (2) hold with* W *independent of* Σ. *Then, conditionally on* Σ, *and under* Q_h, *we have for any integer* M *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$,

a1) $E\left(S_{\beta,h,M}\right) = 0$;

a2) $E\left(S^2_{\beta,h,M}\right) = 1$;

a3) $E\left(S^3_{\beta,h,M}\right) = 0$;

a4) $E\left(S_{\beta,h,M} U_{1,\beta,h,M}\right) = 0$;

a5) $E\left(S_{\beta,h,M} U_{2,\beta,h,M}\right) = 0$;

a6) $E\left(S^3_{\beta,h,M} U_{1,\beta,h,M}\right) = 0$;

a7) $E\left(S^3_{\beta,h,M} U_{2,\beta,h,M}\right) = 0$.

Lemma C.7. *Suppose (1) and (2) hold with* W *independent of* Σ. *Then, conditionally on* Σ, *and under* Q_h, *we have for any integer* M *such that* $M \approx ch^{-\alpha}$ *with* $\alpha \in [0, 1/2)$,

a1) $(\hat{\Gamma}_{k(j)})^{-1} \sum_{i=1}^{M} \hat{u}^2_{i+(j-1)M} = \frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}} \right)^2$;

a2) $E\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}} \right)^2 \right)^{\frac{q}{2}} = \left(\frac{M-1}{M}\right)^{\frac{q}{2}} b_{M,q} c_{M-1,q} \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}} \right)^2 \right)^{\frac{q}{2}}$, *for* $M > q$;

a3) $R_{\beta,q} \equiv Mh \sum_{j=1}^{1/Mh} \left(\frac{M^{\frac{q}{2}}}{(M-1)^{\frac{q}{2}} b_{M,q} c_{M-1,q}} \right) \left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}} \right)^2 \right)^{\frac{q}{2}} - Mh \sum_{j=1}^{1/Mh} \left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}} \right)^2 \right)^{\frac{q}{2}} \to 0$ *in probability under* Q_h *and* P, *for any* $M > q(1+\delta)$, *for some* $\delta > 0$;

a4) $\hat{V}_{\beta,h,M} - V_{\beta,h,M} \to 0$ *in probability under* Q_h *and* P, *for any* $M > 2(1+\delta)$, *for some* $\delta > 0$.

Proof of Lemma C.3. Given the definition of $\hat{V}_{\beta,h,M}$ in the text (see Equation (16)), and the definition of $\hat{u}_{i+(j-1)M} = y_{l,i+(j-1)M} - \hat{\beta}_{lk(j)} y_{k,i+(j-1)M}$, we can write

$$\begin{aligned} \hat{V}_{\beta,h,M} &= M^2 h \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M} \right)^{-1} \left(\frac{1}{M-1} \sum_{i=1}^{M} \left(y_{l,i+(j-1)M} - \hat{\beta}_{lk(j)} y_{k,i+(j-1)M} \right)^2 \right) \\ &= \frac{M^2 h}{M-1} \sum_{j=1}^{1/Mh} \left(\sum_{i=1}^{M} y^2_{k,i+(j-1)M} \right)^{-1} \left(\sum_{i=1}^{M} \left(u_{i+(j-1)M} - \left(\hat{\beta}_{lk(j)} - \beta_{lk(j)} \right) y_{k,i+(j-1)M} \right)^2 \right), \end{aligned}$$

where we used the definition of $y_{l,i+(j-1)M}$ see Equation (16). Adding and subtracting appropriately, it follows that

$$\begin{aligned}
\hat{V}_{\beta,h,M} &= \frac{M^2 h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)+\left(\hat{\beta}_{lk(j)}-\beta_{lk(j)}\right)^2\right)\\
&- 2\frac{M^2 h}{M-1}\sum_{j=1}^{1/Mh}\left(\hat{\beta}_{lk(j)}-\beta_{lk(j)}\right)\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\\
&= \frac{M^2 h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)\\
&- \frac{M^2 h}{M-1}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^2\\
&= \sum_{j=1}^{1/Mh} V_{1,(j)}-\sum_{j=1}^{1/Mh} V_{2,(j)},
\end{aligned}$$

where we used $\left(\hat{\beta}_{lk(j)}-\beta_{lk(j)}\right)=\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)$.

Proof of Lemma C.4 part a1). Given the definition of $V_{1,(j)}$, the law of iterated expectations and the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, we can write

$$\begin{aligned}
E\left(V_{1,(j)}\right) &= E\left(E\left(V_{1,(j)}|y_{k(j)}\right)\right)\\
&= \frac{M^2 h}{M-1}E\left(E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)\right)|y_{k(j)}\right)\\
&= \frac{M^2 h}{M-1}E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} E\left(u_{i+(j-1)M}^2|y_{k(j)}\right)\right)\right)\\
&= \frac{M^3 h}{M-1}V_{(j)}E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right),
\end{aligned}$$

then given equation (14) in the text and by replacing $V_{(j)}$ by $\frac{1}{M}\left(\Gamma_{l(j)}-\frac{\Gamma_{lk(j)}^2}{\Gamma_{k(j)}}\right)$, we have that

$$E\left(V_{1,(j)}\right)=\frac{M^3 h}{(M-1)(M-2)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right).$$

Proof of Lemma C.4 part a2). Given the definition of $V_{1,(j)}$ and the law of iterated expectations, we can write

$$\begin{aligned}
E\left(V_{1,(j)}^2\right) &= E\left(E\left(V_{1,(j)}^2|y_{k(j)}\right)\right)\\
&= \frac{M^4 h^2}{(M-1)^2}E\left(E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)^2\right)|y_{k(j)}\right)\\
&= \frac{M^4 h}{(M-1)^2}V_{(j)}^2 E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}E\left(\sum_{i=1}^{M}\left(\frac{u_{i+(j-1)M}}{\sqrt{V_{(j)}}}\right)^2|y_{k(j)}\right)^2\right).
\end{aligned}$$

Note that since $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, $E\left(\sum_{i=1}^{M}\left(\frac{u_{i+(j-1)M}}{\sqrt{V_{(j)}}}\right)^2 |y_{k(j)}\right)^2 = E\left(\chi_{j,M}^2\right)^2 = M\left(M+2\right)$ where $\chi_{j,M}^2$ follow the standard χ^2 distribution with M degrees of freedom. Then we have

$$E\left(V_{1,(j)}^2\right) = \frac{M^5\left(M+2\right)h}{(M-1)^2}V_{(j)}^2 E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\right),$$

then given the fact that $\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 \overset{d}{=} \frac{\Gamma_{k(j)}}{M}\chi_{j,M}^2$, where '$\overset{d}{=}$' denotes equivalence in distribution, by using the second moment of an inverse of χ^2 distribution, we have $E\left(\frac{1}{\chi_{j,M}^2}\right)^2 = \frac{1}{(M-2)(M-4)}$, and by replacing $V_{(j)}$ by $\frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma_{lk(j)}^2}{\Gamma_{k(j)}}\right)$ it follows that

$$E\left(V_{1,(j)}^2\right) = \frac{M^5\left(M+2\right)}{\left(M-1\right)^2\left(M-2\right)\left(M-4\right)}h^2\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^2.$$

Proof of Lemma C.4 part a3). Given the definition of $V_{1,(j)}$, the law of iterated expectations and the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$. we can write

$$\begin{aligned}
E\left(V_{2,(j)}\right) &= E\left(E\left(V_{2,(j)}|y_{k(j)}\right)\right) \\
&= \frac{M^2 h}{M-1}E\left(E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^2\right)|y_{k(j)}\right) \\
&= \frac{M^2 h}{M-1}E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2 E\left(u_{i+(j-1)M}^2|y_{k(j)}\right)\right)\right) \\
&= \frac{M^2 h}{M-1}V_{(j)}E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\right),
\end{aligned}$$

then using equation (14) in the text and replacing $V_{(j)}$ by $\frac{1}{M}\left(\Gamma_{l(j)} - \frac{\Gamma_{lk(j)}^2}{\Gamma_{k(j)}}\right)$ yields

$$E\left(V_{2,(j)}\right) = \frac{M^2 h}{\left(M-1\right)\left(M-2\right)}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right).$$

Proof of Lemma C.4 part a4). Given the definition of $V_{2,(j)}$ and the law of iterated expectations, we can write

$$\begin{aligned}
E\left(V_{2,(j)}^2\right) &= E\left(E\left(V_{2,(j)}^2|y_{k(j)}\right)\right) \\
&= \frac{M^4 h^2}{(M-1)^2}E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-4} E\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^4 |y_{k(j)}\right) \\
&\equiv \frac{M^4 h^2}{(M-1)^2}E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-4} A\right).
\end{aligned}$$

Then using the conditional independence and mean zero property of $y_{k,i+(j-1)M}u_{i+(j-1)M}$ we have that

$$\begin{aligned}
A &\equiv E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^4 |y_{k(j)}\right)\\
&= \sum_{i=1}^{M} E\left(y_{k,i+(j-1)M}^4 u_{i+(j-1)M}^4 |y_{k(j)}\right)\\
&+ 3\sum_{i\neq s} E\left(y_{k,i+(j-1)M}^2 u_{i+(j-1)M}^2 |y_{k(j)}\right) E\left(y_{k,s+(j-1)M}^2 u_{s+(j-1)M}^2 |y_{k(j)}\right)\\
&= 3V_{(j)}^2\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^4 + \sum_{i\neq s} y_{k,i+(j-1)M}^2 y_{k,s+(j-1)M}^2\right)\\
&= 3V_{(j)}^2\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^4\right)^2,
\end{aligned}$$

thus we can write

$$E\left(V_{2,(j)}^2\right) = \frac{M^4h^2}{(M-1)^2}3V_{(j)}^2 E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\right),$$

result follows similarly where we use the same arguments as in the proof of Lemma C.4 part a2).

Proof of Lemma C.4 part a5). The proof follows similarly as parts a2) and a4) of Lemma C.4 and therefore we omit the details.

Proof of Lemma C.5 part a1). Given the definitions of $\hat{V}_{\beta,h,M}$, $V_{1,(j)}$, $V_{2,(j)}$ and by using Lemma C.3 and part 1 of Lemma C.4, we can write

$$\begin{aligned}
E\left(\hat{V}_{\beta,h,M}\right) &= E\left(\sum_{j=1}^{1/Mh} V_{1,(j)}\right) - E\left(\sum_{j=1}^{1/Mh} V_{2,(j)}\right)\\
&= \sum_{j=1}^{1/Mh} E\left(V_{1,(j)}\right) - \sum_{j=1}^{1/Mh} E\left(V_{1,(j)}\right)\\
&= \frac{M^3h}{(M-1)(M-2)}\sum_{j=1}^{1/Mh}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right) - \frac{M^2h}{(M-1)(M-2)}\sum_{j=1}^{1/Mh}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}} - \left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)\\
&= \frac{M}{M-1}V_{\beta,h,M} - \frac{1}{M-1}V_{\beta,h,M}\\
&= V_{\beta,h,M}.
\end{aligned}$$

Proof of Lemma C.5 part a2). Given the definitions of $\hat{V}_{\beta,h,M}$, $V_{1,(j)}$, $V_{2,(j)}$ and Lemma C.3, we can write

$$Var\left(\hat{V}_{\beta,h,M}\right) = Var\left(\sum_{j=1}^{1/Mh} V_{1,(j)}\right) + Var\left(\sum_{j=1}^{1/Mh} V_{2,(j)}\right) - 2Cov\left(\sum_{j=1}^{1/Mh} V_{1,(j)}, \sum_{j=1}^{1/Mh} V_{2,(j)}\right),$$

given the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, we have $V_{1,(j)}$ and $V_{2,(j)}$ are conditionally independent given $y_{k(j)}$, $V_{1,(j)}$ and $V_{2,(t)}$ are conditionally independent for all $t \neq j$ given $y_{k(j)}$. It follows that

$$
\begin{aligned}
Var\left(\hat{V}_{\beta,h,M}\right) &= \sum_{j=1}^{1/Mh}\left(E\left(V_{1,(j)}^2\right)-E\left(V_{1,(j)}\right)^2\right)+\left(E\left(V_{2,(j)}^2\right)-E\left(V_{2,(j)}\right)^2\right)\\
&- 2\sum_{j=1}^{1/Mh}\left(E\left(V_{2,(j)}V_{2,(j)}\right)-E\left(V_{1,(j)}\right)E\left(V_{2,(j)}\right)\right),
\end{aligned}
$$

finally results follow given Lemma C.4.

Proof of Lemma C.5 part a3). Results follow directly given Lemma C.4 parts a1) and a2) since $E\left(\hat{V}_{\beta,h,M}-V_{\beta,h,M}\right)=0$ and $Var\left(\hat{V}_{\beta,h,M}-V_{\beta,h,M}\right)\to 0$ as $h\to 0$ provide that $Mh\to 0$

Proof of Lemma C.5 part a4). This result follows from the boundedness of $\Sigma_k(u)$, $\Sigma_l(u)$ and the Reimann integrable of $\Sigma_{kl}(u)$ for any $k,l=1\cdots d$.

Proof of Lemma C.6 part a1). Given the definition of $S_{\beta,h,M}$ we can write

$$
\begin{aligned}
E\left(S_{\beta,h,M}\right) &= \frac{\sqrt{h^{-1}}M}{\sqrt{V_{\beta,h,M}}}\sum_{i=1}^{M}E\left(\left(\hat{\beta}_{lk(j)}-\beta_{lk}(j)\right)\right)\\
&= 0,
\end{aligned}
$$

where the last equality use the unbiased property of OLS estimator $\hat{\beta}_{lk(j)}$.

Proof of Lemma C.6 part a2). Given the definitions of $S_{\beta,h,M}$ and $V_{\beta,h,M}$ we can have that

$$
\begin{aligned}
Var\left(S_{\beta,h,M}\right) &= \frac{1}{V_{\beta,h,M}}Var\left(\sqrt{h}\left(\hat{\beta}_{lk}-\beta_{lk}\right)\right)\\
&= 1.
\end{aligned}
$$

Proof of Lemma C.6 part a3). Given the definition of $S_{\beta,h,M}$ and the fact that we can write $\sqrt{h}\left(\hat{\beta}_{lk}-\beta_{lk}\right)$ as follows

$$
\sqrt{h}\left(\hat{\beta}_{lk}-\beta_{lk}\right)=M\sqrt{h}\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right),
$$

it follows that

$$
E\left(S_{\beta,h,M}^3\right)=\frac{M^3h^{3/2}}{V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)\right)^3,
$$

then using the fact that $u_{i+(j-1)M}|y_{k(j)}\sim i.i.d.N\left(0,V_{(j)}\right)$, we have that

$$
\begin{aligned}
E\left(S_{\beta,h,M}^3\right) &= \frac{M^3h^{3/2}}{V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^2\right)^{-3}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}u_{i+(j-1)M}\right)^3\right)\\
&= \frac{M^3h^{3/2}}{V_{\beta,h,M}^{3/2}}E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^2\right)^{-3}\left(\sum_{i=1}^{M}y_{k,i+(j-1)M}^3E\left(u_{i+(j-1)M}^3|y_{k(j)}\right)\right)\right)\\
&= 0.
\end{aligned}
$$

Proof of Lemma C.6 part a4). We start the proof by introducing this notation, which is relevant only for part a4) of Lemma C.6. We let $B=E\left(S_{\beta,h,M}U_{1,\beta,h,M}\right)$, then given the definitions of $S_{\beta,h,M}$,

$U_{1,\beta,h,M}$ and using the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, we can write

$$\begin{aligned} B &= \frac{M^3 h}{(M-1)V_{\beta,h,M}^{3/2}} \sum_{j=1}^{1/Mh} E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-2}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M} u_{i+(j-1)M}\right)\left(\sum_{i=1}^{M} u_{i+(j-1)M}^2\right)\right) \\ &- \frac{M}{V_{\beta,h,M}^{3/2}} \sum_{j=1}^{1/Mh} E\left(V_{1,(j)}\right) E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M} u_{i+(j-1)M}\right)\right), \end{aligned}$$

using the law of iterated expectations and again the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, results follow.

Proof of Lemma C.3 part a5). Given the definitions of $S_{\beta,h,M}$, $U_{2,\beta,h,M}$ and using the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$, we can write

$$\begin{aligned} E\left(S_{\beta,h,M} U_{2,\beta,h,M}\right) &= \frac{M^3 h^{3/2}}{(M-1)V_{\beta,h,M}^{3/2}} E\left(\sum_{j=1}^{1/Mh}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-3}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M} u_{i+(j-1)M}\right)^3\right) \\ &- \frac{M}{V_{\beta,h,M}^{3/2}} \sum_{j=1}^{1/Mh} E\left(V_{2,(j)}\right) E\left(\left(\sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right)^{-1}\left(\sum_{i=1}^{M} y_{k,i+(j-1)M} u_{i+(j-1)M}\right)\right), \end{aligned}$$

then results follow by using the law of iterated expectations and again the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$.

The proof of the remaining results (Lemma C.6 part a6) and part a7)) follow similarly and therefore we omit the details.

Proof of Lemma C.7 part a1). Given the definition of $\hat{u}_{i+(j-1)M}$, we can write

$$\begin{aligned} (\hat{\Gamma}_{k(j)})^{-1} \sum_{i=1}^{M} \hat{u}_{i+(j-1)M}^2 &= \frac{1}{\hat{\Gamma}_{k(j)}} \sum_{i=1}^{M}\left(y_{l,i+(j-1)M} - \hat{\beta}_{lk(j)} y_{k,i+(j-1)M}\right)^2 \\ &= \frac{1}{\hat{\Gamma}_{k(j)}} \sum_{i=1}^{M}\left(y_{l,i+(j-1)M}^2 - 2\hat{\beta}_{lk(j)} y_{l,i+(j-1)M} y_{k,i+(j-1)M} + \hat{\beta}_{lk(j)}^2 y_{k,i+(j-1)M}^2\right) \\ &= \frac{1}{\hat{\Gamma}_{k(j)}}\left(\sum_{i=1}^{M} y_{l,i+(j-1)M}^2 - 2\hat{\beta}_{lk(j)} \sum_{i=1}^{M} y_{l,i+(j-1)M} y_{k,i+(j-1)M} + \hat{\beta}_{lk(j)}^2 \sum_{i=1}^{M} y_{k,i+(j-1)M}^2\right), \end{aligned}$$

thus results follow by replacing $\hat{\beta}_{lk(j)} = \frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}$.

Proof of Lemma C.7 part a2). Given the definitions of $\hat{\Gamma}_{l(j)}$, $\hat{\Gamma}_{k(j)}$ and $\hat{\Gamma}_{lk(j)}$ and using part a1) of

Lemma C.7, we can write

$$\begin{aligned}
E\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{q}{2}} &= E\left(\frac{\sum_{i=1}^{M}\hat{u}^2_{i+(j-1)M}}{\hat{\Gamma}_{k(j)}}\right)^{\frac{q}{2}}\\
&= E\left(\left(\sum_{i=1}^{M}y^2_{k,i+(j-1)M}\right)^{-\frac{q}{2}}E\left(\sum_{i=1}^{M}\hat{u}^2_{i+(j-1)M}\right)^{\frac{q}{2}}|y_{k(j)}\right)\\
&= E\left(\left(\sum_{i=1}^{M}y^2_{k,i+(j-1)M}\right)^{-\frac{q}{2}}V_{(j)}^{\frac{q}{2}}E\left(\left(\sum_{i=1}^{M}\frac{\hat{u}^2_{i+(j-1)M}}{V_{(j)}}\right)^{\frac{q}{2}}|y_{k(j)}\right)\right),
\end{aligned}$$

where we use the law of iterated expectations and the fact that $u_{i+(j-1)M}|y_{k(j)} \sim i.i.d.N\left(0, V_{(j)}\right)$. Then given the definition of $c_{M,q}$, we can write

$$\begin{aligned}
E\left(\left(\sum_{i=1}^{M}\frac{\hat{u}^2_{i+(j-1)M}}{V_{(j)}}\right)^{\frac{q}{2}}|y_{k(j)}\right) &= E\left(\left(\chi^2_{j,M}\right)^{\frac{q}{2}}\right)\\
&= (M-1)^{\frac{q}{2}}c_{M-1,q},
\end{aligned}$$

it follows then that,

$$\begin{aligned}
E\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{q}{2}} &= E\left(\frac{\sum_{i=1}^{M}\hat{u}^2_{i+(j-1)M}}{\hat{\Gamma}_{k(j)}}\right)^{\frac{q}{2}}\\
&= (M-1)^{\frac{q}{2}}c_{M-1,q}V_{(j)}^{\frac{q}{2}}E\left(\left(\sum_{i=1}^{M}y^2_{k,i+(j-1)M}\right)^{-\frac{q}{2}}\right)\\
&= (M-1)^{\frac{q}{2}}c_{M-1,q}V_{(j)}^{\frac{q}{2}}\Gamma_{k(j)}^{-\frac{q}{2}}E\left(\left(\frac{M}{\chi^2_{j,M}}\right)^{\frac{q}{2}}\right)\\
&= \left(\frac{M-1}{M}\right)^{\frac{q}{2}}b_{M,q}c_{M-1,q}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^{\frac{q}{2}};
\end{aligned}$$

where $b_{M,q} = E\left(\left(\frac{M}{\chi^2_{j,M}}\right)^{\frac{q}{2}}\right)$, for $M > q$.

Proof of Lemma C.7 part a3). We verify the moments conditions of the Weak Law of Large Numbers for independent and nonidentically distributed on $z_j \equiv \frac{M^{\frac{q}{2}}}{(M-1)^{\frac{q}{2}}b_{M,q}c_{M-1,q}}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{q}{2}}$, $j = 1, \ldots, \frac{1}{Mh}$. By using part a2) of Lemma C.7, for any $\delta > 0$, and conditionally on σ, we can write

$$E\left|z_j\right|^{1+\delta} = \left(\frac{M-1}{M}\right)^{\frac{\delta q}{2}}\frac{b_{M,(1+\delta)q}}{b_{M,q}}\frac{c_{M-1,(1+\delta)q}}{c_{M-1,q}}\left(\frac{\Gamma_{l(j)}}{\Gamma_{k(j)}}-\left(\frac{\Gamma_{lk(j)}}{\Gamma_{k(j)}}\right)^2\right)^{\frac{(1+\delta)q}{2}} < \infty$$

since Σ is an adapted càdlàg spot covolatility matrix and locally bounded and invertible (in particular, $\Gamma_{k(j)} > 0$), moreover in the case where $M \to \infty$, as $h \to 0$ (i.e. $M \approx ch^{-\alpha}$ with $\alpha \in (0, 1/2)$) we have

$\left(\frac{M-1}{M}\right)^{\frac{\delta q}{2}} \frac{b_{M,(1+\delta)q}}{b_{M,q}} \frac{c_{M-1,(1+\delta)q}}{c_{M-1,q}} \to 1.$

Proof of Lemma C.7 part a4). Result follows directly given the definition of $\hat{V}_{\beta,h,M}, V_{\beta,h,M}$ and part a3) of Lemma C.7, where we let $q = 2$.

Remark 5 The bootstrap analogue of Lemma C.3 and C.4 replace $V_{1(j)}$ with $V^*_{1(j)}$, $V_{2(j)}$, the bootstrap analogue of Lemma C.5 replaces $\hat{V}_{\beta,h,M}$ with $\hat{V}^*_{\beta,h,M}$, $V_{\beta,h,M}$ with $V^*_{\beta,h,M}$, $\Gamma_{l(j)}$ with $\hat{\Gamma}_{l(j)}$, $\Gamma_{k(j)}$ with $\hat{\Gamma}_{k(j)}$, and $\Gamma_{lk(j)}$ with $\hat{\Gamma}_{lk(j)}$; whereas the bootstrap analogue of Lemma C.6 replaces $S_{\beta,h,M}$ with $S^*_{\beta,h,M}$, $U_{1,\beta,h,M}$ with $U^*_{1,\beta,h,M}$ and $U_{2,\beta,h,M}$ with $U^*_{2,\beta,h,M}$.

Lemma C.8. *Suppose (1) and (2) hold with W independent of Σ. Then, conditionally on Σ, we have for any integer M such that $M \approx ch^{-\alpha}$ with $\alpha \in [0, 1/2)$, and for some small $\delta > 0$,*

a1) $E^*\left(\sum_{i=1}^M y^{2*}_{k,i+(j-1)M}\right)^{-2(2+\delta)} = b_{M,4(2+\delta)}\hat{\Gamma}^{-2(2+\delta)}_{k(j)}$, *for $M > 4(2+\delta)$;*

a2) $E^*\left(\left|\sum_{i=1}^M y^*_{k,i+(j-1)M} u^*_{i+(j-1)M}\right|^{2(2+\delta)}\right) \leq \mu^2_{2(2+\delta)} M^{2+\delta}\hat{\Gamma}^{2+\delta}_{k(j)}\hat{V}^{2+\delta}_{(j)}$;

Proof of Lemma C.8 part a1). Given the definition of $y^*_{k,i+(j-1)M}$, we can write, $\sum_{i=1}^M y^{2*}_{k,i+(j-1)M} \stackrel{d}{=} \frac{\hat{\Gamma}_{k(j)}}{M}\sum_{i=1}^M v^2_{i+(j-1)M} \stackrel{d}{=} \frac{\hat{\Gamma}_{k(j)}}{M}\chi^2_{j,M}$, where $v_{i+(j-1)M} \sim i.i.d.N(0,1)$, and $\chi^2_{j,M}$ follow the standard χ^2 distribution with M degrees of freedom. Then for any integer $M > 4(2+\delta)$, we have that,

$$E\left(\sum_{i=1}^M y^2_{k,i+(j-1)M}\right)^{-2(2+\delta)} = E\left(\frac{M}{\chi^2_{j,M}}\right)^{2(2+\delta)} \hat{\Gamma}^{-2(2+\delta)}_{k(j)} = b_{M,4(2+\delta)}\hat{\Gamma}^{-2(2+\delta)}_{k(j)}.$$

Proof of Lemma C.8 part a2). Indeed by using the C_r inequality, the law of iterated expectations and the fact that $u^*_{i+(j-1)M}|y^*_{k(j)} \sim i.i.d.N\left(0, \hat{V}_{(j)}\right)$, we can write for any $\delta > 0$,

$$\begin{aligned} E^*\left(\left|\sum_{i=1}^M y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right|^{2(2+\delta)}\right) &\leq M^{3+2\delta}\sum_{i=1}^M E^*\left|y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right|^{2(2+\delta)} \\ &= M^{3+2\delta}\sum_{i=1}^M E^*\left(y^{*2(2+\delta)}_{k,i+(j-1)M}E^*\left(u^{*2(2+\delta)}_{i+(j-1)M}|y^*_{k(j)}\right)\right) \\ &= \mu^2_{2(2+\delta)}M^{2+\delta}\hat{\Gamma}^{2+\delta}_{k(j)}\hat{V}^{2+\delta}_{(j)}, \end{aligned}$$

where the last equality follows since $y^{2*}_{k,i+(j-1)M} \stackrel{d}{=} \frac{\Gamma_{k(j)}}{M}v^2_{i+(j-1)M}$, where $v_{i+(j-1)M} \sim i.i.d.N(0,1)$ and $\mu_{2(2+\delta)} = E|v|^{2(2+\delta)}$.

Proof of proposition 5.2. As in Theorem B.1, the first and third cumulants of $T_{\beta,h,M}$ are given by

$$\begin{aligned} \kappa_1(T_{\beta,h,M}) &= E(T_{\beta,h,M}), \\ \kappa_3(T_{\beta,h,M}) &= E(T^3_{\beta,h,M}) - 3E(T^2_{\beta,h,M})E(T_{\beta,h,M}) + 2[E(T_{\beta,h,M})]^3. \end{aligned}$$

Here, our goal is to identify the terms of order up to $O(\sqrt{h})$ of the asymptotic expansions of these two cumulants. We will first provide asymptotic expansions through order $O(\sqrt{h})$ for the first three moments of $T_{\beta,h,M}$. Note that for a given fixed value of k, a first-order Taylor expansion of $f(x) = (1+x)^{-k/2}$ around 0 yields $f(x) = 1 - \frac{k}{2}x + O(x^2)$. We have for any fixed integer k, We have that for any fixed integer k,

$$
\begin{aligned}
T^k_{\beta,h,M} &= S^k_{\beta,h,M}\left(1+\sqrt{h}U_{\beta,h,M}\right)^{-k/2}, \\
&= S^k_{\beta,h,M} - \frac{k}{2}\sqrt{h}S^k_{\beta,h,M}U_{\beta,h,M} + O(h) \\
&= S^k_{\beta,h,M} - \frac{k}{2}\sqrt{h}S^k_{\beta,h,M}U_{1,\beta,h,M} + \frac{k}{2}\sqrt{h}S^k_{\beta,h,M}U_{2,\beta,h,M} + O(h).
\end{aligned}
$$

For $k = 1, \cdots, 3$, the moments of $T^k_{\beta,h,M}$ up to order $O(h)$ are given by

$$
\begin{aligned}
E\left(T_{\beta,h,M}\right) &= E\left(S_{\beta,h,M}\right) - \frac{\sqrt{h}}{2}E\left(S_{\beta,h,M}U_{1,\beta,h,M}\right) + \frac{\sqrt{h}}{2}E\left(S_{\beta,h,M}U_{2,\beta,h,M}\right) \\
E\left(T^2_{\beta,h,M}\right) &= E\left(S^2_{\beta,h,M}\right) - \sqrt{h}E\left(S^2_{\beta,h,M}U_{1,\beta,h,M}\right) + \sqrt{h}E\left(S^2_{\beta,h,M}U_{2,\beta,h,M}\right) \\
E\left(T^3_{\beta,h,M}\right) &= E\left(S^3_{\beta,h,M}\right) - \sqrt{h}\frac{3}{2}E\left(S^3_{\beta,h,M}U_{1,\beta,h,M}\right) + \sqrt{h}\frac{3}{2}E\left(S^3_{\beta,h,M}U_{2,\beta,h,M}\right).
\end{aligned}
$$

Given Lemma C.6, we have that

$$
\begin{aligned}
E\left(T_{\beta,h,M}\right) &= 0 \\
E\left(T^2_{\beta,h,M}\right) &= 1 - \sqrt{h}E\left(S^2_{\beta,h,M}U_{1,\beta,h,M}\right) + \sqrt{h}E\left(S^2_{\beta,h,M}U_{2,\beta,h,M}\right) \\
E\left(T^3_{\beta,h,M}\right) &= 0.
\end{aligned}
$$

It follows that $\kappa_1(T_{\beta,h,M}) = 0$ and $\kappa_3(T_{\beta,h,M}) = 0$.

Proof of Theorem 4.1 For part a), the proof follows the same steps as the proof of $V_{\beta,h,M}$ which we explain in the main text, in particular, given the definition of $\hat{\beta}^*_{lk}$, we have that

$$
\begin{aligned}
V^*_{\beta,h,M} &= Var^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})\right) \\
&= M^2 h \sum_{j=1}^{1/Mh} Var^*\left(\hat{\beta}^*_{lk(j)} - \hat{\beta}_{lk(j)}\right) \\
&= M^2 h \sum_{j=1}^{1/Mh} E^*\left(\left(\sum_{i=1}^{M} y^{2*}_{k,i+(j-1)M}\right)^{-1}\right)\hat{V}_{(j)} \\
&= \frac{M^2 h}{M-2}\sum_{j=1}^{1/Mh}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}} - \left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right) \\
&= \frac{M-1}{M-2}\hat{V}_{\beta,h,M},
\end{aligned}
$$

then results follows, given Lemma C.5 or part a4) of Lemma C.7**.**

For part b), we have $\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk}) = \sum_{j=1}^{1/Mh} z^*_{j,\beta}$, where

$$
z^*_{j,\beta} = M\sqrt{h}\left(\sum_{i=1}^{M} y^{*2}_{k,i+(j-1)M}\right)^{-1}\left(\sum_{i=1}^{M} y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right).
$$

Note that $E^*\left(z^*_{j,\beta}\right)=0$, and that

$$Var^*\left(\sum_{j=1}^{1/Mh} z_j^*\right)=V^*_{\beta,h,M}\xrightarrow{P} V_\beta,$$

by part a) moreover, since $z_1^*,\ldots,z^*_{1/Mh}$ are conditionally independent, by the Berry-Esseen bound, for some small $\delta>0$ and for some constant $C>0$,

$$\sup_{x\in\Re}\left|P^*\left(\sqrt{h^{-1}}(\hat{\beta}^*_{lk}-\hat{\beta}_{lk})\leq x\right)-\Phi\left(\frac{x}{V_\beta}\right)\right|\leq C\sum_{j=1}^{1/Mh}E^*\left|z_j^*\right|^{2+\delta},$$

Next, we show that $\sum_{j=1}^{1/Mh}E^*\left|z_j^*\right|^{2+\delta}=o_p(1)$. We have that

$$\begin{aligned}\sum_{j=1}^{1/Mh}E^*\left|z^*_{j,\beta}\right|^{2+\delta} &= \left(M\sqrt{h}\right)^{2+\delta}\sum_{j=1}^{1/Mh}E^*\left(\left(\sum_{i=1}^{M}y^{*2}_{k,i+(j-1)M}\right)^{-(2+\delta)}\left|\sum_{i=1}^{M}y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right|^{2+\delta}\right)\\ &\equiv \left(M\sqrt{h}\right)^{2+\delta}\sum_{j=1}^{1/Mh}E^*\left(A_j^*B_j^*\right),\end{aligned}$$

it follows then by using Cauchy-Schwarz inequality that

$$\begin{aligned}E^*\left(A_j^*B_j^*\right) &\leq \sqrt{\left(\sum_{i=1}^{M}y^{*2}_{k,i+(j-1)M}\right)^{-2(2+\delta)}}\sqrt{E^*\left(\left|\sum_{i=1}^{M}y^*_{k,i+(j-1)M}u^*_{i+(j-1)M}\right|^{2(2+\delta)}\right)}\\ &\leq \mu_{2(2+\delta)}b^{\frac{1}{2}}_{M,4(2+\delta)}M^{1+\frac{\delta}{2}}\hat{\Gamma}^{-\frac{2+\delta}{2}}_{k(j)}\hat{V}^{\frac{2+\delta}{2}}_{(j)}\\ &= \mu_{2(2+\delta)}b^{\frac{1}{2}}_{M,4(2+\delta)}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{2+\delta}{2}},\end{aligned}$$

where the second inequatily used part a1) and a2) of Lemma C.8 and $\mu_{2(2+\delta)}=E\left|v\right|^{2(2+\delta)}$ with $v\sim N(0,1)$. Finally, given the definition of $R_{\beta,2+\delta}$ and the fact that $M\approx ch^{-\alpha}$, we can write

$$\begin{aligned}\sum_{j=1}^{1/Mh}E^*\left|z^*_{j,\beta}\right|^{2+\delta} &\leq \mu_{2(2+\delta)}b^{\frac{1}{2}}_{M,4(2+\delta)}M^{2+\delta}h^{1+\frac{\delta}{2}}\sum_{j=1}^{1/Mh}\left(\frac{\hat{\Gamma}_{l(j)}}{\hat{\Gamma}_{k(j)}}-\left(\frac{\hat{\Gamma}_{lk(j)}}{\hat{\Gamma}_{k(j)}}\right)^2\right)^{\frac{2+\delta}{2}}\\ &= \mu_{2(2+\delta)}b^{\frac{1}{2}}_{M,4(2+\delta)}\left(\frac{M-1}{M}\right)^{\frac{2+\delta}{2}}b_{M,2+\delta}c_{M-1,2+\delta}M^{1+\delta}h^{\frac{\delta}{2}}R_{\beta,2+\delta}\\ &= O_p\left(h^{\frac{\delta}{2}-\alpha(1+\delta)}b^{\frac{1}{2}}_{M,4(2+\delta)}\left(\frac{M-1}{M}\right)^{\frac{2+\delta}{2}}b_{M,2+\delta}c_{M-1,2+\delta}\right)\\ &= o_p(1).\end{aligned}$$

Since for any $\delta>0$, such that $\alpha\in[0,1/2)$ we have $\frac{\delta}{2}-\alpha(1+\delta)>0$, and $\mu_{2(2+\delta)}=E\left|v\right|^{2(2+\delta)}\leq\Delta<\infty$ where $v\sim N(0,1)$, moreover as $h\to 0$, $c_{M-1,2+\delta}\to 1$, $b_{M,4(2+\delta)}\to 1$, $b_{M,2+\delta}\to 1$ and by using Lemma C.7 we have $R_{\beta,2+\delta}=O_P(1)$.

Proof of Theorem 4.2 Let

$$H^*_{\beta,h,M} = \frac{\sqrt{h^{-1}}(\hat{\beta}^*_{lk} - \hat{\beta}_{lk})}{\sqrt{V^*_{\beta,h,M}}},$$

and note that

$$T^*_{\beta,h,M} = H^*_{\beta,h,M}\sqrt{\frac{V^*_{\beta,h,M}}{\hat{V}^*_{\beta,h,M}}},$$

where $\hat{V}^*_{\beta,h,M}$ is defined in the main text. Theorem 4.1 proved that $H^*_{\beta,h,M} \overset{d^*}{\to} N(0,1)$ in probability. Thus, it suffices to show that $\hat{V}^*_{\beta,h,M} - V^*_{\beta,h,M} \overset{P^*}{\to} 0$ in probability under Q_h and P. In particular, we show that (1) $Bias^*\left(\hat{V}^*_{\beta,h,M}\right) = 0$, and (2) $Var^*\left(\hat{V}^*_{\beta,h,M}\right) \overset{P}{\to} 0$. Results follows directly by using the bootstrap analogue of parts a1), a2) and a3) of Lemma C.5.

References

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

[2] 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.

[3] Barndorff-Nielsen, O. and N. Shephard, (2004a). "Econometric analysis of realised covariation: high frequency based covariance, regression and correlation in financial economics," *Econometrica*, 72, 885-925.

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

[5] 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.

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

[7] Davidson, R., and E. Flachaire, (2001). "The Wild Bootstrap, Tamed at Last," *Working Paper Darp58, STICERD, London School of Economics.*

[8] Dovonon, P., Gonçalves, S. and N. Meddahi, (2013). "Bootstrapping realized multivariate volatility measures," *Journal of Econometrics* 172,49-65.

[9] Gonçalves, S. and N. Meddahi, (2009). "Bootstrapping realized volatility," *Econometrica*, 77(1), 283.306.

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

[11] 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.*

[12] Hounyo, U. , Gonçalves, S., and N. Meddahi, (2013). "Bootstrapping pre-averaged realized volatility under market microstructure noise," manuscript.

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

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

[15] Meddahi, N., (2002). "A theoretical comparison between integrated and realized volatility," *Journal of Applied Econometrics* 17, 475-508.

[16] Mykland, P.A. and L. Zhang, (2009). "Inference for continous semimartingales observed at high frequency," *Econometrica* 77, 1403-1455.

[17] Mykland, P.A. and L. Zhang, (2011). "The double Gaussian approximation for high frequency data," *Scandinavian Journal of Statistics* 38, 215-236.

[18] Mykland, P.A., and L. Zhang, (2012). "The Econometrics of High Frequency Data," Statistical Methods for Stochastic Differential Equations, M. Kessler, A. Lindner, and M. Sørensen, eds. p. 109-190.

[19] Mykland, P.A. N. Shephard, and K. Shepphard, (2012). "Efficient and feasible inference for the components of financial variation using blocked multipower variation," Working paper, Oxford University.

[20] 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.

Research Papers 2013

2013-13: Tom Engsted, Stig V. Møller and Magnus Sander: Bond return predictability in expansions and recessions

2013-14: Charlotte Christiansen, Jonas Nygaard Eriksen and Stig V. Møller: Forecasting US Recessions: The Role of Sentiments

2013-15: Ole E. Barndorff-Nielsen, Mikko S. Pakkanen and Jürgen Schmiegel: Assessing Relative Volatility/Intermittency/Energy Dissipation

2013-16: Peter Exterkate, Patrick J.F. Groenen, Christiaan Heij and Dick van Dijk: Nonlinear Forecasting With Many Predictors Using Kernel Ridge Regression

2013-17: Daniela Osterrieder: Interest Rates with Long Memory: A Generalized Affine Term-Structure Model

2013-18: Kirstin Hubrich and Timo Teräsvirta: Thresholds and Smooth Transitions in Vector Autoregressive Models

2013-19: Asger Lunde and Kasper V. Olesen: Modeling and Forecasting the Volatility of Energy Forward Returns - Evidence from the Nordic Power Exchange

2013-20: Anders Bredahl Kock: Oracle inequalities for high-dimensional panel data models

2013-21: Malene Kallestrup-Lamb, Anders Bredahl Kock and Johannes Tang Kristensen: Lassoing the Determinants of Retirement

2013-22: Johannes Tang Kristensen: Diffusion Indexes with Sparse Loadings

2013-23: Asger Lunde and Anne Floor Brix: Estimating Stochastic Volatility Models using Prediction-based Estimating Functions

2013-24: Nima Nonejad: A Mixture Innovation Heterogeneous Autoregressive Model for Structural Breaks and Long Memory

2013-25: Nima Nonejad: Time-Consistency Problem and the Behavior of US Inflation from 1970 to 2008

2013-26: Nima Nonejad: Long Memory and Structural Breaks in Realized Volatility: An Irreversible Markov Switching Approach

2013-27: Nima Nonejad: Particle Markov Chain Monte Carlo Techniques of Unobserved Compdonent Time Series Models Using Ox

2013-28: Ulrich Hounyo, Sílvia Goncalves and Nour Meddahi: Bootstrapping pre-averaged realized volatility under market microstructure noise

2013-29: Jiti Gao, Shin Kanaya, Degui Li and Dag Tjøstheim: Uniform Consistency for Nonparametric Estimators in Null Recurrent Time Series

2013-30: Ulrich Hounyo: Bootstrapping realized volatility and realized beta under a local Gaussianity assumption