Dette, Holger; Golosnoy, Vasyl; Kellermann, Janosch

The effect of intraday periodicity on realized volatility measures

Metrika

Provided in Cooperation with:
Springer Nature

Suggested Citation: Dette, Holger; Golosnoy, Vasyl; Kellermann, Janosch (2022) : The effect of intraday periodicity on realized volatility measures, Metrika, ISSN 1435-926X, Springer, Berlin, Heidelberg, Vol. 86, Iss. 3, pp. 315-342, https://doi.org/10.1007/s00184-022-00875-0

This Version is available at:
https://hdl.handle.net/10419/307917

Metrika (2023) 86:315–342
https://doi.org/10.1007/s00184-022-00875-0

The effect of intraday periodicity on realized volatility measures

Holger Dette[1] · **Vasyl Golosnoy**[2] · **Janosch Kellermann**[2]

Received: 17 August 2021 / Accepted: 5 July 2022 / Published online: 16 July 2022

Abstract
We focus on estimating daily integrated volatility (IV) by realized measures based on intraday returns following a discrete-time stochastic model with a pronounced intraday periodicity (IP). We demonstrate that neglecting the IP-impact on realized estimators may lead to invalid statistical inference concerning IV for a common finite number of intraday returns. For a given IP functional form, we analytically derive robust IP-correction factors for realized measures of IV as well as their asymptotic distributions. We show both in Monte Carlo simulations and empirically that the proposed bias corrections are the robust way to account for IP by computing realized estimators.

Keywords Bipower variation · Integrated volatility · Integrated quarticity · Intraday periodicity

1 Introduction

For the majority of financial markets, the pervasive intraday periodicity (IP), which often takes a U- or mirrored J-form during the daily trading time, is a well documented empirical feature of intraday absolute returns (Wood et al. 1985; Harris 1986). As it appears to be highly correlated with intraday variation of trading volume, Admati and Pfleiderer (1988) propose to explain daily IP-shape by strategic interaction of traders around market openings and closures, whereas the periodicity at weekly or

✉ Vasyl Golosnoy
vasyl.golosnoy@rub.de

Holger Dette
holger.dette@rub.de

Janosch Kellermann
janosch.kellermann@rub.de

1 Department of Mathematics, Ruhr-Universität Bochum, Bochum, Germany

2 Faculty of Management and Economics, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany

monthly horizons could be attributed to the impact of slowly varying macroeconomic fundamentals (Andersen and Bollerslev 1998b; Andersen et al. 2001, 2003).

The availability of high-frequency data allows for the construction of precise estimators of daily integrated volatility (IV) for risky asset returns. The realized volatility (RV) defined as a sum of squared intraday returns is known to be a consistent estimator of daily IV in absence of jumps. Other realized measures such as the bipower variation (BV) should be used for IV estimation in presence of finite-activity jumps during the day (cf. Aït-Sahalia and Jacod 2014). Barndorff-Nielsen and Shephard (2004) derive the asymptotic properties of these quantities for the number of intraday returns $M \to \infty$ under mild assumptions on the corresponding pricing process. In practice, however, the number of intraday returns available for computing realized estimators often remains limited due to insufficient liquidity and/or irregular trading activity. Even for highly liquid stocks a practitioner could prefer to rely on simple realized estimators based e.g. on 5 min returns primarily in order to escape from adverse effects of market microstructure noise (MMN) which are particularly pronounced at (ultra) high sampling frequencies (cf. Aït-Sahalia and Jacod 2014).

In this paper we analyze and quantify the impact of IP on the finite M properties of RV and BV estimators by providing corresponding formal statements given the IP functional form. To the best of our knowledge, this research agenda has not been explored yet, although there is a vast amount of literature devoted to modeling and estimating IP (cf. Engle et al. 1990; Andersen et al. 2019; Christensen et al. 2018). Thus, our investigation provides useful insights for exploring differences between the asymptotic theory and the practical finite sample performance of realized measures based on intraday data.

To model intraday returns, we presume a discrete time stochastic specification as in Andersen and Bollerslev (1997), where the variance of intraday returns is written as a product of the deterministic periodic and stochastic volatility (SV) components. The IP is assumed to be constant for all days, whereas the SV part is slowly changing over time. Our framework is motivated by the empirical evidence that the IP captures a vast part of intraday volatility variation whereas the SV impact is of a smaller order (cf. Christensen et al. 2018). We show that for the commonly available finite number of intraday returns M neglecting the impact of IP would lead to non-valid statistical inference concerning daily IV. For a given IP, we compute the first and the second moments of RV and BV, moreover, we establish the asymptotic bivariate distribution of these measures as $M \to \infty$. We also quantify the impact of IP on realized tri-power (TP) and quad-power (QP) estimators of daily integrated quarticity (IQ) required for statistical inference about IV.

Our major finding is that for the commonly available finite number of intraday returns the impact of IP should be explicitly addressed when making statistical inference concerning IV. While the RV estimator of IV is unaffected by IP, BV has a finite sample bias negligible only for large sample sizes which are not always available in practice. Moreover, by estimating IQ one should account – at least for small M values – for scaling factors, which depend on the functional form of the IP. We derive the explicit expressions for these IP-correction factors and provide the asymptotic distribution for their estimators.

Our theoretical results are illustrated in a Monte Carlo study, where we investigate the impact of IP on various realized measures. We find that the IP-correction procedures proposed in this paper are helpful against the adverse impact of IP on realized measures in finite samples. Moreover, our IP-corrected estimates are advantageous compared to immediate removing of the estimated IP due to their robustness with respect to IP misspecifications at days with unusual pattern of intraday volatility (cf. Gabrys et al. 2013; Kokoszka and Reimherr 2013). In such situations our approach is preferable in terms of relative bias and mean squared error (MSE) compared to the standard procedure with immediate scaling of intraperiod returns by the estimated IP profile as in Boudt et al. (2011), for example. In the empirical application we estimate the IP for the daily volatility of the Dow Jones Industrial Average Index with and without IP bias corrections.

The remaining part of the paper is organized as follows. In Sect. 2, we introduce the model for intraday returns and discuss the realized estimators of daily IV and IQ. The theoretical results are derived in Sect. 3 where we establish both finite sample and asymptotic stochastic properties of commonly applied realized estimators for a given IP form. Moreover, we provide expressions for IP correction factors and derive the asymptotic distributions of their estimators. Our approach is illustrated in Sect. 4 by means of a simulation study and in Sect. 5 by an empirical application. Sect. 6 concludes, whereas the proofs are placed in the Appendix.

2 Measuring daily volatility based on intraday information

Before we introduce our model in Sect. 2.2, we provide definitions of the objects which are of importance for our analysis. For this purpose we start from a general jump-diffusion model for log price increments in order to define the objects of our interest, namely the daily IV and daily IQ. Next, we present the corresponding realized estimators which are based on M intraday returns. These realized measures are consistent estimators of IV for $M \to \infty$, however, in practice we often face $M \leq 10^2$ primarily because ultra high frequency returns are contaminated by MMN. Thus, it is of importance to explore the finite sample stochastic properties of realized estimators. For this purpose we then consider a discrete time model for intraday returns with an explicit functional specification of intraday periodicity (IP) and study the impact of IP on the realized measures of IV and IQ for finite M.

2.1 Model for intraday returns and realized measures

Assume that log-prices of risky assets $p(t) = \ln P_t$ follow a continuous time process with (possible) additive jump components. We consider a day t as the period of interest with the daily return $r_t = p(t) - p(t-1)$ and focus on the integrated volatility (IV), which is defined for day t as

$$IV_t = \sigma_t^2 = \int_{t-1}^{t} \sigma^2(u)du,$$

where $\sigma(u)$ is a spot volatility. In order to make statistical inference about IV measures one also needs statements concerning the daily integrated quarticity (IQ) defined by

$$IQ_t = \int_{t-1}^{t} \sigma^4(u)du.$$

The availability of intraday returns allows to construct precise *realized* estimators (Andersen and Bollerslev 1998a) for the daily IV which are of immense practical importance for estimation and inferences concerning daily volatility. Assume that M equally spaced intraday returns are available for day t. We denoted them by $r_{t,m} = p(t-1+m/M) - p(t-1+(m-1)/M)$ for $m = 1, \ldots, M$. Then the daily return r_t is the sum of intraday returns with $r_t = \sum_{m=1}^{M} r_{t,m}$. The most popular IV estimator is the realized volatility (RV) measure given as

$$RV_t = \sum_{m=1}^{M} r_{t,m}^2.$$

Barndorff-Nielsen and Shephard (2004) show that RV_t is a consistent estimator of IV_t without jumps at day t, i.e. $RV_t \xrightarrow{p} IV_t$ as $M \to \infty$. Although the estimator RV_t possesses a set of appealing properties, it is not appropriate in the presence a non-zero jump component.

The bipower variation (BV) proposed by Barndorff-Nielsen and Shephard (2004)

$$BV_t = \frac{M}{M-1}\frac{\pi}{2}\sum_{m=2}^{M} |r_{t,m}||r_{t,m-1}| \tag{1}$$

is a jump-robust estimator of IV. It is consistent even in presence of jumps, i.e. $BV_t \xrightarrow{p} IV_t$ as $M \to \infty$. However, RV has a smaller variance than BV if there is no jumps, so the common practice is first to test for a jump during day t. Then, in case of a significantly large positive distance between RV and BV indicating jumps, one should apply BV; otherwise RV is to use.

Intraday returns are also suitable for the purpose of estimating the unknown daily IQ required for computing variances of RV and BV measures. The realized quarticity (RQ)

$$RQ_t = \frac{M}{3}\sum_{m=1}^{M} r_{t,m}^4,$$

is a consistent estimator of IQ in the case of no jumps with $RQ_t \xrightarrow{p} IQ_t$ as $M \to \infty$. However, as RQ_t measure is not robust (cf. Andersen et al. 2014), Barndorff-Nielsen

and Shephard (2004) suggest to use the realized tri-power (TP) and quad-power (QP) measures defined by

$$QP_t = \frac{M^2}{M-3} \cdot \frac{\pi^2}{4} \cdot \sum_{m=4}^{M} |r_{t,m-3}||r_{t,m-2}||r_{t,m-1}||r_{t,m}|, \tag{2}$$

$$TP_t = \frac{M^2}{M-2} \cdot \mu_{4/3}^{-3} \sum_{m=3}^{M} |r_{t,m-2}|^{\frac{4}{3}} |r_{t,m-1}|^{\frac{4}{3}} |r_{t,m}|^{\frac{4}{3}}, \quad \text{with} \quad \mu_{4/3} = 0.8309. \tag{3}$$

Although both RV and BV have appealing stochastic properties as $M \to \infty$, their practical implementation is often based on (say) 5 min intraday returns which makes $M = 78$ intraday observations for a usual 6.5-hours trading day, because of MMN which hinders the use of ultra high frequency data for construction of realized estimators (McAleer and Medeiros 2008). To overcome these problems, the recent research has been focused on making realized estimators more robust to these features. However, in many situations the common practice still remains to sample returns at a lower frequency, i.e. to consider 5–, 10–, or even 15-min intraday returns (Andersen et al. 2011). We follow this strand of literature and concentrate on profound understanding of stochastic properties of realized estimators for comparatively small values M. Although we focus in our analysis primarily on the classical RV, BV, QP, and TP measures, we also provide a discussion of recently proposed further realized measures in Sect. 5.2 and in Dette et al. (2022).

2.2 Discrete time model for intraday returns

The IP in absolute intraday returns is one of the most important stylized facts characterizing high frequency data. In order to investigate the IP impact on realized measures for a fixed number of intraday returns M, next we introduce a discrete time model in (4) which is central for study. There is a substantial scope of recent literature concerning discrete-time modeling of intraday returns whereas the IP is assumed to be a multiplicative scaling component (Boudt et al. 2011; Engle and Sokalska 2012; Bekierman and Gribisch 2021). Following Andersen and Bollerslev (1997), we focus on a discrete stochastic model for intraday return without jumps, which is given as

$$\begin{aligned} r_{t,m} &= \sigma_{t,m} \cdot u_{t,m}, \qquad \text{with} \qquad u_{t,m} \sim \text{iid}\, \mathcal{N}(0,1), \\ \sigma_{t,m}^2 &= 1/M \cdot s_{t,m}^2 \cdot \gamma_{t,m}^2, \end{aligned} \tag{4}$$

where $s_{t,m} > 0$ is the deterministic IP component and $\gamma_{t,m}^2 > 0$ is the stochastic part. Note that there is no leverage effect in (4), as it is mostly of importance for daily returns but much less pronounced for high-frequency intraday returns, see e.g. Bollerslev et al. (2006).

For our theoretical derivations we presume that the stochastic part remains constant within day t (Andersen and Bollerslev 1998b; Hecq et al. 2012), i.e. $\gamma_{t,m} = \sigma_t$ for all

$m = 1, \ldots, M$, but may change from one day to another. This assumption is justified by the empirical evidence that IP commonly accounts for a vast part of intraday volatility variation (cf. Christensen et al. 2018). In Sect. 4 we relax this assumption in the Monte Carlo simulation study by considering the intraday SV which follows a diffusion process as e.g. in Goncalves and Meddahi (2009). Based on the results of our Monte Carlo simulations, we conclude that our major findings also hold in the SV setting.

In line with the current literature (Hecq et al. 2012), we set the IP as constant at different days so that we further skip the time index with $s_{t,m} = s_m$. Moreover, the periodic component is standardized such that it sums up to M over the day with $\sum_{m=1}^{M} s_m^2 = M$. Of course, a special case $s_m = 1$ for all $m = 1, \ldots, M$ corresponds to no IP. Putting all together, we separate the intraday periodic component s_m which is solely responsible for intraday heteroskedasticity, and interday stochastic component σ_t which could change from one day to another by writing

$$\sigma_{t,m}^2 = 1/M \cdot s_m^2 \cdot \sigma_t^2 \qquad \text{with} \qquad \sum_{m=1}^{M} s_m^2 = M. \tag{5}$$

Although the model in (4) and (5) is fairly simple, it allows a detailed analysis of the IP impact on popular realized measures of the objects of our interest, which are the IV for day t

$$IV_t = Var(r_t) = \sum_{m=1}^{M} Var(r_{t,m}) = \frac{1}{M} \sum_{m=1}^{M} s_m^2 \sigma_t^2 = \sigma_t^2, \tag{6}$$

as well as the IQ written by (Andersen et al. 2014)

$$IQ_t = M/3 \cdot \sum_{m=1}^{M} E(r_{t,m}^4) = \sigma_t^4/M \cdot \sum_{m=1}^{M} s_m^4. \tag{7}$$

Note that as in general it holds that $\sum_{m=1}^{M} s_m^4 \geq M$, the IQ is directly influenced by IP. We aim to investigate the impact of IP on realized measures of IV and IQ.

3 The impact of intraday periodicity on *RV* and *BV*

To gain results on the IP impact on realized measures, re-write the normalized IP $\{s_m\}_{m=1}^{M}$ as

$$s_m^2 = g\left(\frac{m}{M}\right) / g_M, \qquad \text{with} \qquad g_M = \frac{1}{M} \sum_{m=1}^{M} g\left(\frac{m}{M}\right), \tag{8}$$

where $g : [0,1] \mapsto \mathbb{R}^+$ with $\mathbb{R}^+ := (0, +\infty)$ is a given non-normalized function. This normalization is very common in IP-literature (cf. Andersen and Bollerslev 1997, p. 153). The functional form of $g(\cdot)$ could be quite flexible and is subject to very general regularity conditions specified in the following propositions.

3.1 Bias and variance of realized estimators

For the discrete model of intraday returns (4)–(5) and the IP from $\{g(m/M)\}_{m=1}^{M}$, we derive expectation, bias and variance of RV and BV estimators of daily IV in the next proposition.

Proposition 1 *Assume that the IP component is given by (8) for some function* $g : [0,1] \mapsto \mathbb{R}^+$.
(A) The estimator RV_t *for daily* IV *is unbiased so that* $E[RV_t] = IV_t$. *The estimator* BV_t *is biased, that is* $E[BV_t] = M/(M-1)\sigma_t^2(1-R_M) = M/(M-1)IV_t(1-R_M)$, *where the factor* $R_M \in (0,1)$ *is given by*

$$R_M = \left(g\left(\frac{1}{M}\right) + \sum_{m=2}^{M} g\left(\frac{m}{M}\right)^{1/2} \left[g\left(\frac{m}{M}\right)^{1/2} - g\left(\frac{m-1}{M}\right)^{1/2} \right] \right) \Big/ \sum_{m=1}^{M} g\left(\frac{m}{M}\right).$$

If $g(\cdot)$ *is continuously differentiable on interval* $[0,1]$, *it holds as* $M \to \infty$

$$M \cdot R_M = \left[\frac{1}{2} \int_0^1 g'(x)dx \Big/ \int_0^1 g(x)dx + g(0) \Big/ \int_0^1 g(x)dx \right] \cdot (1 + o(1)),$$

so that $\lim_{M\to\infty} R_M = 0$, *i.e.* BV_t *is an asymptotically unbiased estimator of* IV.
(B) The (co)variances of RV_t *and* BV_t *are given as*

$$\begin{aligned}
Var(RV_t) &= \frac{2\sigma_t^4}{M^2 g_M^2} \sum_{m=1}^{M} g^2\left(\frac{m}{M}\right), \\
Var(BV_t) &= \frac{\sigma_t^4}{(M-1)^2 g_M^2} \left\{ \left(\frac{\pi^2}{4} - 1\right) \sum_{m=2}^{M} g\left(\frac{m}{M}\right) g\left(\frac{m-1}{M}\right) \right. \\
&\quad \left. + (\pi - 2) \sum_{m=3}^{M} g\left(\frac{m}{M}\right)^{1/2} g\left(\frac{m-1}{M}\right) g\left(\frac{m-2}{M}\right)^{1/2} \right\}, \\
Cov(RV_t, BV_t) &= \frac{\sigma_t^4}{M(M-1) g_M^2} \left[\sum_{m=2}^{M} g\left(\frac{m-1}{M}\right)^{1/2} g\left(\frac{m}{M}\right)^{3/2} \right. \\
&\quad \left. + \sum_{m=1}^{M-1} g\left(\frac{m+1}{M}\right)^{1/2} g\left(\frac{m}{M}\right)^{3/2} \right].
\end{aligned}$$

For $g(\cdot)$ continuously differentiable on interval $[0, 1]$ *and for* $M \to \infty$ *it holds that* $Var(RV_t) = (1/M) \cdot 2\sigma_t^4 \cdot \xi\ (1 + o(1))$,

$$Var(BV_t) = (1/M) \cdot \sigma_t^4 \cdot \xi \left(\frac{\pi^2}{4} - 3 + \frac{\pi}{4} \right) (1 + o(1)), \tag{9}$$

and $Cov(RV_t, BV_t) = (1/M) \cdot 2\sigma_t^4 \cdot \xi(1 + o(1))$. *The asymptotic scaling factor* ξ *is defined by*

$$\xi = \int_0^1 g^2(x)dx \Big/ \left(\int_0^1 g(x)dx \right)^2. \tag{10}$$

It holds that $\xi \geq 1$ *with* $\xi = 1$ *if and only if* $g(\cdot)$ *is almost everywhere constant, i.e. there is no IP.*

The property $R_M \in (0, 1)$ follows from the proof of Proposition 1 in the Appendix. More precisely, by the Cauchy-Schwarz inequality, we have

$$\begin{aligned} 0 < 1 - R_M &= \frac{\sum_{m=2}^{M} \left[g\left(\frac{m}{M}\right) g\left(\frac{m-1}{M}\right) \right]^{1/2}}{\sum_{m=1}^{M} g\left(\frac{m}{M}\right)} \\ &\leq \frac{\left(\sum_{m=2}^{M} g\left(\frac{m}{M}\right) \right)^{1/2} \left(\sum_{m=2}^{M} g\left(\frac{m-1}{M}\right) \right)^{1/2}}{\sum_{m=1}^{M} g\left(\frac{m}{M}\right)} < 1. \end{aligned}$$

Thus, in the case of IP, RV is an unbiased estimator for IV but BV has a finite M bias which should be corrected in applications. Since the expectation of BV is given by

$$E[BV_t] = \frac{\sigma_t^2}{(M-1) \cdot g_M} \cdot \sum_{m=2}^{M} g\left(\frac{m}{M}\right)^{1/2} g\left(\frac{m-1}{M}\right)^{1/2} = \frac{\sigma_t^2}{M-1} \cdot \sum_{m=2}^{M} s_m s_{m-1},$$

we suggest the following bias-corrected measure

$$\widetilde{BV}_t = \frac{\pi}{2} \cdot M \cdot \left(\sum_{m=2}^{M} s_m s_{m-1} \right)^{-1} \cdot \sum_{m=2}^{M} |r_{t,m}||r_{t,m-1}|. \tag{11}$$

Hence, it holds that $\widetilde{BV}_t = (M-1) \left(\sum_{m=2}^{M} s_m s_{m-1} \right)^{-1} BV_t$ leading the correction factor ζ_M for BV:

$$\zeta_M = \zeta_{M,BV} = BV_t / \widetilde{BV}_t = \frac{1}{M-1} \cdot \sum_{m=2}^{M} s_m s_{m-1}, \tag{12}$$

which should be replaced by its empirical counterpart $\hat{\zeta}_M$ based on IP estimates $\hat{s}_m$ in practice. By the same principle, we define the IP factor ξ_M in RQ for finite M values as

$$\xi_M = \xi_{M,RQ} = \frac{1}{M} \cdot \sum_{m=1}^{M} s_m^4, \tag{13}$$

with $\lim_{M\to\infty} \xi_M = \xi$ as in (10) of Proposition 1. Note that in case of no IP it holds that $\xi_M = \xi_{M,RQ} = 1$. Of course, in applications we replace $\xi_{M,RQ}$ by its estimator $\hat{\xi}_M = \hat{\xi}_{M,RQ}$ which is discussed below in Sect. 3.3. Hence, we show analytically that although for BV it holds that $\lim_{M\to\infty} \zeta_M = 1$, its counterpart ξ for the realized measures of IQ is still present and could be (depending on the data) rather substantial. In the context of the IP-bias corrections for further realized measures, we investigate by means of numerical analysis several popular MMN-robust realized estimators such as min RV or med RV in the follow-up paper of Dette et al. (2022).

In the next proposition we provide the expectations of realized estimators RQ, TP and QP serving as the measures for IQ under the model defined in (4) and (5) with the function $g(\cdot)$.

Proposition 2 *Assume that the IP is given by (8), then the expectations of RQ_t, QP_t and TP_t are given as $E[RQ_t] = \frac{\sigma_t^4}{M \cdot g_M^2} \sum_{m=1}^{M} g(\frac{m}{M})^2$, i.e. RQ_t is unbiased, and*

$$E[QP_t] = \frac{\sigma_t^4}{(M-3) \cdot g_M^2} \sum_{m=4}^{M} \left[g\left(\frac{m-3}{M}\right) g\left(\frac{m-2}{M}\right) g\left(\frac{m-1}{M}\right) g\left(\frac{m}{M}\right) \right]^{\frac{1}{2}},$$

$$E[TP_t] = \frac{\sigma_t^4}{(M-2) \cdot g_M^2} \sum_{m=3}^{M} \left[g\left(\frac{m-2}{M}\right) g\left(\frac{m-1}{M}\right) g\left(\frac{m}{M}\right) \right]^{\frac{2}{3}}.$$

Moreover, $\lim_{M\to\infty} E[RQ_t] = E[TP_t] = E[QP_t] = \sigma_t^4 \cdot \xi = IQ_t$ if $g(\cdot)$ is square integrable, where the asymptotic scaling factor ξ which comprises the impact of IP is given by (10).

Proposition 2 suggests the following factorization of QP and TP measures of IQ in finite samples:

$$\widetilde{QP}_t = QP_t / \xi_{M,QP}, \qquad \xi_{M,QP} = \frac{1}{M-3} \cdot \sum_{m=4}^{M} s_m s_{m-1} s_{m-2} s_{m-3},$$

$$\widetilde{TP}_t = TP_t / \xi_{M,TP}, \qquad \xi_{M,TP} = \frac{1}{M-2} \cdot \sum_{m=3}^{M} (s_m s_{m-1} s_{m-2})^{4/3},$$

whereby $\xi_{M,QP}$ and $\xi_{M,TP}$ are the IP scaling factors for QP and TP, respectively. This factorization appears to be useful both in the simulations in Sect. 4 and in the empirical study in Sect. 5.

3.2 Asymptotic distribution in case of intraday periodicity

Next, we provide the corresponding bivariate limit distribution for RV and BV as $M \to \infty$ for our discrete time model of intraday returns with IP.

Theorem 1 *Consider model* (4) *and* (5) *without jumps and assume that the IP component is given by* (8) *with a continuously differentiable function* $g : [0, 1] \mapsto \mathbb{R}$. *Then, as* $M \to \infty$,

$$M^{1/2} \cdot IQ_t^{-1/2} \cdot \begin{pmatrix} RV_t - \sigma_t^2 \\ BV_t - \sigma_t^2 \end{pmatrix} \xrightarrow{L} \mathcal{N}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 & 2 \\ 2 & \frac{\pi^2}{4} + \pi - 3 \end{bmatrix} \right).$$

The integrated quarticity $IQ_t = \xi \cdot \sigma_t^4$ *can be consistently estimated by* RQ_t, QP_t, *or* TP_t.

Thus, a pronounced IP with $\xi > 1$ causes more variability of IV estimators compared to the case of no IP where $\xi = 1$. The asymptotic $(1 - \alpha)$-confidence interval for daily IV and IQ based on RV and RQ measures is given according to Theorem 1 as

$$CI_t(1-\alpha) = \left[RV_t + z_{\alpha/2} \left[\frac{2}{M} \right]^{1/2} \cdot RQ_t^{1/2}, \quad RV_t - z_{\alpha/2} \left[\frac{2}{M} \right]^{1/2} \cdot RQ_t^{1/2} \right],$$

where $z_{\alpha/2}$ is the $(\alpha/2)$-quantile of the standard normal distribution. The asymptotic confidence intervals based on BV, TP or QP measures are constructed similarly.

The results in Proposition 2 and Theorem 1 are useful for statistical inference on IV. As RQ is not robust, one could estimate IQ in presence of jumps by either TP or QP directly as in (2) or (3), i.e. without estimating ξ separately. However, as we show in the Monte Carlo simulation, the approximation $TP_t \approx IQ_t$ is only precise for fairly large M. For this reason, for finite M we recommend to use the IP-scaled estimators $\widetilde{QP}_t$ or $\widetilde{TP}_t$ as well as the estimated scaling factor $\hat{\xi}_{M,RQ}$ from Eq. (15) below for the construction of confidence intervals as

$$\widetilde{CI}_t(1-\alpha) = \left[\widetilde{BV}_t + z_{\alpha/2} \left[\frac{\pi^2/4 + \pi - 4}{M} \right]^{1/2} \cdot \hat{\xi}_{M,RQ}^{1/2} \cdot \widetilde{QP}_t^{1/2}, \right.$$
$$\left. \widetilde{BV}_t - z_{\alpha/2} \left[\frac{\pi^2/4 + \pi - 4}{M} \right]^{1/2} \cdot \hat{\xi}_{M,RQ}^{1/2} \cdot \widetilde{QP}_t^{1/2} \right],$$

because of $E[\hat{\xi}_{M,RQ} \cdot \widetilde{QP}_t] = E[\hat{\xi}_{M,RQ} \cdot \widetilde{TP}_t] = E[TP_t] = E[QP_t] = E[RQ_t]$.

3.3 Estimation of IP correction factors

For a given number of intraday returns M, one needs consistent estimators of the IP functions $\{\hat{s}_m^2\}_{m=1}^M$ in order to obtain the estimators of the quantities ξ_M and ζ_M defined

in (13) and (12), respectively. For this purpose we exploit the SD estimator (cf. Boudt et al. 2011) given as

$$\widehat{SD}^2_{m,T} = \frac{1}{T}\sum_{t=1}^{T} r^2_{t,m}\,, \quad \text{and} \quad \hat{s}_m = \frac{\widehat{SD}_{m,T}}{\left((1/M)\sum_{m=1}^{M}\widehat{SD}^2_{m,T}\right)^{1/2}}, \tag{14}$$

and use the statistics

$$\hat{\xi}_M = \hat{\xi}_{M,RQ} = (1/M)\cdot\sum_{m=1}^{M}\hat{s}^4_m\,, \qquad \hat{\zeta}_M = \hat{\zeta}_{M,BV} = (1/M)\cdot\sum_{m=2}^{M}\hat{s}_m\hat{s}_{m-1} \tag{15}$$

as the estimators of ξ_M and ζ_M, respectively. If the variance

$$\sigma^2 := \lim_{T\to\infty}\sigma^2_T := \lim_{T\to\infty}\frac{1}{T}\sum_{t=1}^{T}\sigma^2_t > 0 \tag{16}$$

exists, it follows from (5) that $\lim_{T\to\infty} E[\widehat{SD}^2_{m,T}] = \sigma^2 s^2_m/M$ which motivates the definition (14). As a consequence, we expect that the estimators in (14) and (15) are consistent for s_m, $m = 1,\ldots,M$, as well as for ξ_M and ζ_M, respectively. Note that as long as the ergodicity condition (16) is met, the assumption of constant intraday volatility as in Sect. 2 could be relaxed for the following analysis in Sect. 3.3, e.g. by allowing intraday stochastic volatility.

In order to make the intuitive arguments above more precise we investigate in the following the asymptotic distribution of the statistics $\hat{\zeta}_M$ and $\hat{\xi}_M$ for finite M and $T\to\infty$. Note that under common assumptions, such as mixing conditions (see, for example Dehling et al. 1986, among many others) or physical dependence conditions (cf. Wu 2005), the vector $\widehat{\mathbf{SD}}_{M,T} = (\widehat{SD}^2_{1,T},\ldots,\widehat{SD}^2_{M,T})^\top$ is asymptotically normal distributed if $T\to\infty$, that is

$$\sqrt{T}\left(\widehat{\mathbf{SD}}_{M,T} - \mathbf{SD}_{M,T}\right) \xrightarrow{L} \mathcal{N}(\mathbf{0}, \sigma^4\Sigma) \quad \text{with} \quad \mathbf{SD}_{M,T} = (SD^2_{1,T},\ldots,SD^2_{M,T})^\top\,, \tag{17}$$

where $SD^2_{m,T} = E(\widehat{SD}^2_{m,T}) = \sigma^2_T s^2_m/M$ for $m = 1,\ldots,M$, $\sigma^4\Sigma \in \mathbb{R}^{M\times M}$ denotes a covariance matrix reflecting the underlying dependence structure and σ^2, σ^2_T are defined in (16).

Now we provide the theoretical result concerning the estimators $\hat{\xi}_M$ and $\hat{\zeta}_M$ for finite number of intraday observations M and the estimation period $T\to\infty$, which are based on the SD-estimator of the IP components s_m, $m = 1,\ldots,M$, although similar type of result could be also obtained for other IP-estimators. We denote by $\langle x, y\rangle = x^\top y$ the common inner product on $\mathbb{R}^M$ with the corresponding norm $\|x\| = \langle x, x\rangle^{1/2}$ and by $\mathbf{1}_M$ an M-dimensional vector of unit entries.

Proposition 3 *Assume that* (16) *and* (17) *hold, then, as* $T \to \infty$, *that*

$$\sqrt{T}\,(\hat{\xi}_M - \xi_M) \xrightarrow{L} \mathcal{N}(0, \tau_1^2) \tag{18}$$

$$\sqrt{T}\,(\hat{\zeta}_M - \zeta_M) \xrightarrow{L} \mathcal{N}(0, \tau_2^2)\,, \tag{19}$$

where the asymptotic variances are given by

$$\tau_1^2 = 4\left(\mathbf{S}_M^\top \Sigma \mathbf{S}_M - \frac{2}{M}\|\mathbf{S}_M\|^2 \mathbf{1}_M^\top \Sigma \mathbf{S}_M + \frac{\mathbf{1}_M^\top \Sigma \mathbf{1}_M}{M^2}\|\mathbf{S}_M\|^4\right), \tag{20}$$

$$\tau_2^2 = \frac{1}{4}\left(\mathbf{R}_M^\top \Sigma \mathbf{R}_M - \frac{2}{M}\langle \mathbf{R}_M, \mathbf{S}_M\rangle \mathbf{1}_M^\top \Sigma \mathbf{R}_M + \frac{\mathbf{1}_M^\top \Sigma \mathbf{1}_M}{M^2}\langle \mathbf{R}_M, \mathbf{S}_M\rangle^2\right), \tag{21}$$

and the vectors $\mathbf{S}_M$ *and* $\mathbf{R}_M$ *are defined by*

$$\mathbf{S}_M := \left(s_1^2, \ldots, s_M^2\right)^\top, \qquad \mathbf{R}_M := \left(\frac{s_2}{s_1}, \frac{s_1 + s_3}{s_2}, \ldots, \frac{s_{M-2} + s_M}{s_{M-1}}, \frac{s_{M-1}}{s_M}\right)^\top,$$

respectively. In particular, if $\Sigma = \operatorname{diag}(\vartheta_1^2, \ldots, \vartheta_M^2)$ *is a diagonal matrix, we have*

$$\tau_1^2 = 4\left\{\sum_{m=1}^M \vartheta_m^2 s_m^4 - \frac{2}{M}\sum_{m=1}^M \vartheta_m^2 s_m^2 \sum_{m=1}^M s_m^4 + \frac{1}{M^2}\left(\sum_{m=1}^M s_m^4\right)^2 \sum_{m=1}^M \vartheta_m^2\right\}, \tag{22}$$

$$\begin{aligned}\tau_2^2 = \frac{1}{4}\Bigg\{&\sum_{m=1}^M \frac{\vartheta_m^2}{s_m^2}(s_{m-1} + s_{m+1})^2 - \frac{2}{M}\sum_{m=1}^M s_m(s_{m-1} + s_{m+1}) \sum_{m=1}^M \vartheta_m^2 \frac{s_{m-1} + s_{m+1}}{s_m} \\ &+ \frac{1}{M^2}\sum_{m=1}^M \vartheta_m^2 \left(\sum_{m=1}^M s_m(s_{m-1} + s_{m+1})\right)^2\Bigg\}\,,\end{aligned} \tag{23}$$

where we put $s_0 = s_{M+1} = 0$.

When estimating IP elements s_m, we require for consistency that the number of days $T \to \infty$, whereas the number of intraday observations M is here fixed. Hence, we derive the distributions of the correction factors $\hat{\zeta}_M$ and $\hat{\xi}_M$ for fixed M and $T \to \infty$ in Proposition 3. Differently, inference about the *asymptotic* correction factor ξ in Theorem 1 would require that $M \to \infty$ because then $\hat{\xi}_M$ is a consistent estimator of ξ. These results allow to make statistical inferences and conduct tests for IP correction factors ξ_M and ζ_M with $T \to \infty$. Of course, an extension of our theoretical findings for IP estimated from a finite sample T is also of interest. However, this is also a quite challenging task, for which Christensen et al. (2018) present with some theoretical considerations in their Proposition 3.1. We investigate this issue in the Monte Carlo simulations in Sect. 4.2.

In Proposition 3 we provide the explicit asymptotic results for IP scaling factors given the SD estimator. In general, our results for the SD estimator could be extended

for any consistent estimator of IP. In this paper both in our simulations and empirical studies we use a more robust WSD estimator of Boudt et al. (2011) which is described in the Appendix. However, it is much more difficult to get analytical results such as in Proposition 3 for this more complicated WSD estimator which is based on order statistics. For this reason we recommend to make statistical inferences for the WSD approach by using bootstrap procedures, as e.g. in Goncalves and Meddahi (2009) or in Dette et al. (2022).

4 Simulation study

We illustrate our theoretical findings by means of an extensive Monte Carlo simulation study which is structured as follows. First, we introduce the U-shaped IP functional form and discuss the parameter choice for the model in (4)–(5). We generate intraday returns for estimation of IP shape and construction of various realized measures. In Sect. 4.1 we study the impact of IP on various realized estimators. In Sect. 4.2 we investigate our IP-corrections in terms of MSE whereby we compare our approach with those of Boudt et al. (2011).

We consider $M = 26$, $M = 78$, or $M = 390$ intraday returns, which roughly correspond to sampling at 15 min, 5 min, or 1 min for 6.5-hour trading days, respectively. For constant intraday volatility, we fix $IV = \sigma^2 = 1$ and $IQ = \xi \cdot \sigma^4 = \xi$. We generate M intraday returns for each of $T = 10^4$ days with $r_{t,m} \sim \mathcal{N}(0, \gamma_{t,m}^2 s_m^2)$ where γ^2 and s_m^2 are the respective SV and IP components. As a baseline, we set $\gamma_{t,m}^2 = 1/M$ which corresponds to $IV = 1$ and $IQ = \xi$.

Later we also consider a SV model where we assume that $\gamma_{t,m}^2$ is governed by the process

$$\Delta\gamma_{t,m}^2 = 0.035(0.636 - \gamma_{t,m}^2)\Delta t + 0.144\gamma_{t,m}^2(\Delta t)^{1/2}u_{t,m}, \tag{24}$$

where $\Delta t = 1/M$ and $u_{t,m} \sim \text{iid}\mathcal{N}(0, 1)$, with $\gamma_{1,1}^2 = 1$. This model is a discretized version of the GARCH (1,1) diffusion used by Andersen and Bollerslev (1998a), Goncalves and Meddahi (2009) with parameters implying an autoregressive persistence; in the time series context it is related to state-space models for realized volatilities (cf. Golosnoy et al. 2021). In presence of intraday SV there is no exact analytical expression for the impact of IP on realized measures. However, since the SV in (24) follows a highly persistent process, the empirical contribution of $\gamma_{t,m}^2$ to intraday heteroskedasticity is of a smaller order compared to the impact of IP (see e.g., Christensen et al. 2018; Bekierman and Gribisch 2021). Moreover, from the empirical perspective, one could precisely estimate the SV components $\gamma_{t,m}$ ex post, see Bekierman and Gribisch (2016). Then one could calculate the product components $\hat{s}_{t,m} = \hat{s}_m\hat{\gamma}_{t,m}$ and compute the IP correction factors for this day t based on the obtained estimates $\hat{s}_{t,m}$.

The IP components follow a quadratic convex U-shaped given by

$$g(m/M) = \left[c_1 + c_2(m - M/2)^2\right]^2, \qquad c_1, c_2 > 0, \tag{25}$$

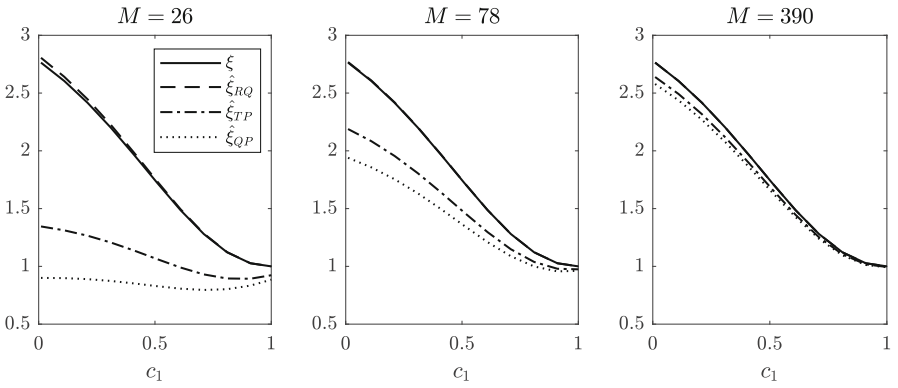

Fig. 1 The asymptotic factor ξ in (10) and averages of finite sample estimates $\hat{\xi}_{M,TP}, \hat{\xi}_{M,QP}, \hat{\xi}_{M,RQ}$ for $M = 26, 78, 390$ depending on parameter c_1 of the intraday profile (25)

so that the standardized IP values are $s_m^2 = g(m/M)/g_M$. The choice of the functional form is (25) is motivated by our empirical findings, see Fig. 5. Note that our theoretical results are applicable for any functional form of IP including asymmetric mirrored J-shape specifications, as e.g. a more flexible asymmetric U-shaped IP specification as in Hasbrouck (1999) and Andersen et al. (2012). Because of $\sum_{m=1}^{M} s_m^2 = M$, it holds for (25) that $c_2 = 12\,(1 - c_1)\,/\,(M^2 + 2)$. We select $c_1 \in \{0.01, 0.11, \ldots, 0.91, 1\}$ with the most IP curvature for $c_1 \to 0^+$ whereas $c_1 = 1$ is the no-IP case. Next we focus on the value $c_1 = 0.71$ corresponding to the evidence from U.S. stock market. Note that the IP form could be very distinct at 'special' days characterized by specific announcements and/or unexpected events where the IP curvature could be much more (or less) pronounced.

The asymptotic scaling factor ξ defined in (10) is plotted as a function of c_1 in Fig. 1 where we observe that ξ is substantially larger than one even for c_1 close to one. As the IP is unknown in practice, we construct estimators $\hat{g}(\cdot)$ and $\hat{s}^2(\cdot)$ by applying the WSD estimator of Boudt et al. (2011) which does not require a-priori specification of the IP functional form; the implementation details for the WSD estimator are provided in the Appendix. Then we compute the average estimates of the scaling factors $\hat{\xi}_{M,RQ}$, $\hat{\xi}_{M,TP}, \hat{\xi}_{M,QP}$ which are plotted in Fig. 1. We observe that the factor $\hat{\xi}_{M,RQ}$ appears to be very close to ξ_M even for a fairly small value $M = 26$.

4.1 The impact of IP on realized measures

As RV measure is not affected by IP, we study the IP impact on other measures, such as BV for IV, and TP, QP for IQ. After generating IID intraday returns as specified above, we calculate BV_t, TP_t, and QP_t as well as the IP-corrected estimators denoted by $\widetilde{BV}_t$, etc. for each day $t = 1, \ldots, T$. Additionally, we compute the jump-robust medRV and minRV measures of Andersen et al. (2012). Then we build time averages, e.g. $\overline{BV} = (1/T) \cdot \sum_{t=1}^{T} BV_t$, for all measures. These averages of BV, $\widetilde{BV}$, medRV, and minRV are shown in Fig. 2; whereas of TP, QP, $\widetilde{TP}$, and $\widetilde{QP}$ in Fig. 3 for different M and c_1 values. All measures should be equal to one for no IP with $c_1 = 1$.

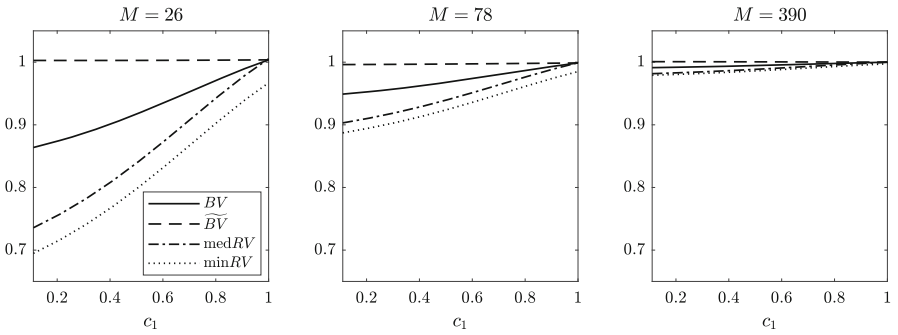

Fig. 2 Uncorrected BV, medRV, minRV as well as corrected $\widetilde{BV}$ as a function of c_1

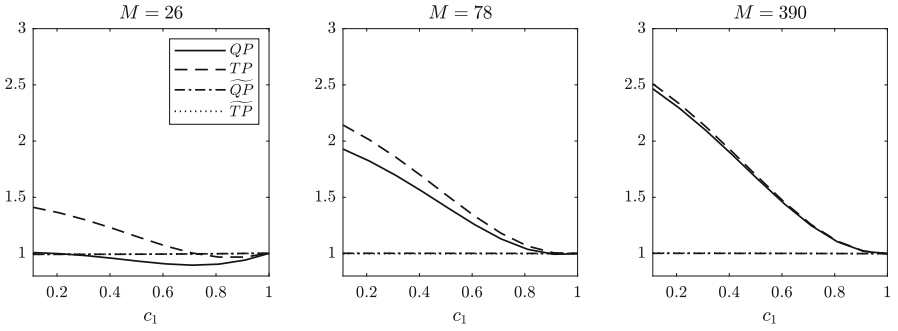

Fig. 3 Original QP and TP as well as scaled $\widetilde{QP}$ and $\widetilde{TP}$ as a function of c_1

In Fig. 2 the IP-bias in BV is quite pronounced for $M = 26$ and $M = 78$ for large and medium curvatures and is still visible even for $M = 390$. Remarkably, the biases in medRV and minRV are even stronger than in BV. As expected, the bias-corrected mean of $\widetilde{BV}$ is close to the true IV for all M, so the suggested correction functions properly.

The averages of the original TP, QP and scaled $\widetilde{TP}$, $\widetilde{QP}$ are reported in Fig. 3. The original measures are downward biased for finite $M = 26, 78$, compared to almost unbiased measures for $M = 390$ which is close to the asymptotic value ξ, see Fig. 1. Remarkably, in case of 5 minute returns with $M = 78$ the bias is still quite substantial for empirical relevant values of IP curvature parameter $c_1 \in [0.6, 0.8]$. As expected, QP is more biased than TP for finite M due to longer lags involved in its computing. Our scaled $\widetilde{TP}$ and $\widetilde{QP}$ measures are equal to the value $\sigma^4 = 1$ for all considered values of c_1 and M which is an appealing property. Summarizing, the IP has a substantial influence on BV, minRV, medRV, so that IP-corrections are needed to get valid statistical inference on IV.

In order to shed light on the finite sample validity of the result in Proposition 3, we provide the exemplary QQ-plots in Fig. 4 for the statistics $\hat{\zeta}_M$ and $\hat{\xi}_M$ for the case of $M = 78$ and $c_1 = 0.6$ which are based on the WSD estimator of IP (cf. Boudt et al. 2011). Being standardized properly, both statistics seem to approach normality with the increase of estimation period T, in particular, even $T = 250$ the QQ plots show a decent fit.

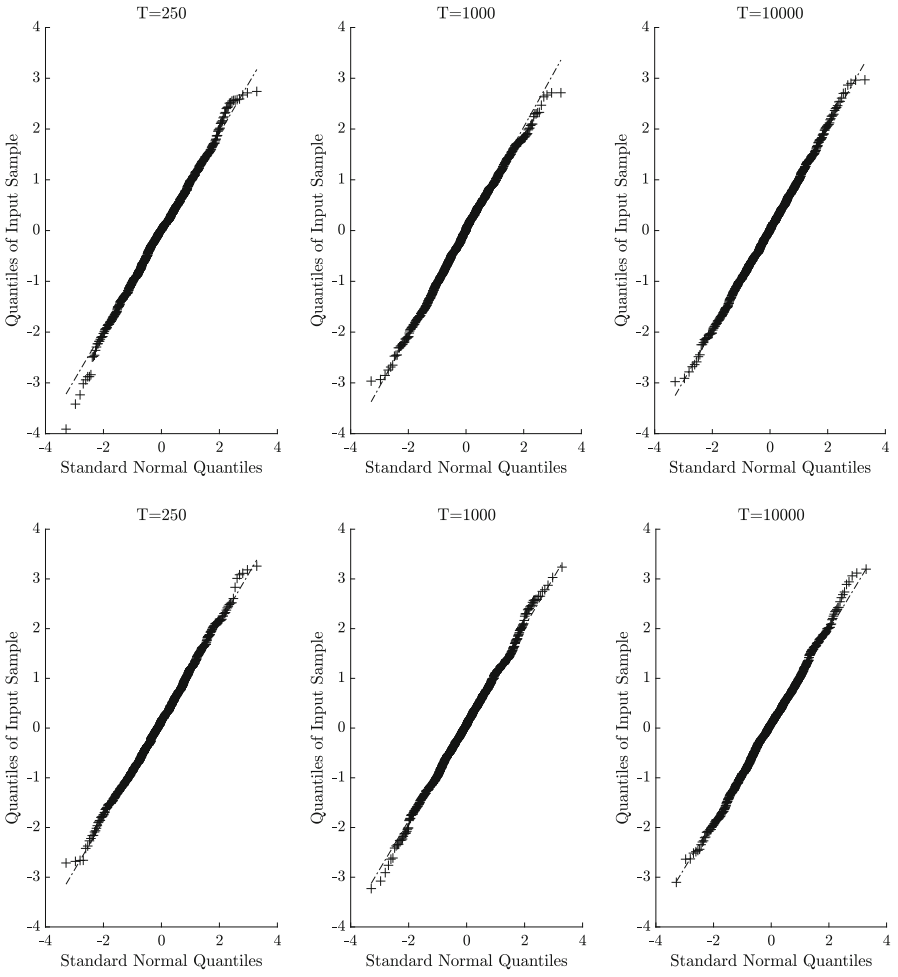

Fig. 4 QQ-plots for the statistics $\hat{\zeta}_M$ (top panel) and $\hat{\xi}_M$ (bottom panel) for various values of T, $M = 78$ and $c_1 = 0.6$

4.2 The comparison of IP-bias corrected estimators

In Proposition 1 we show analytically that the original BV is downward biased so that the true level of risk measured by IV is underestimated. This is a rather undesired scenario from the risk management point of view making an MSE comparison of biased and unbiased measures not reasonable because the MSE is symmetric for upward and downward biases. For this reason, we provide an MSE comparison only for IP-bias corrected estimators.

In particular, we contrast the relative bias and MSE of our corrected estimator $\widetilde{BV}_t$ in (11) with those of Boudt et al. (2011) where the estimated IP component $\hat{s}_m$ is immediately removed from intraday returns by computing $r^*_{t,m} = r_{t,m}/\hat{s}_m$. Then the BCL-estimator BV^*_t is given as

$$BV_t^* = \frac{M}{M-1} \cdot \frac{\pi}{2} \cdot \sum_{m=2}^{M} |r_{t,m}^*||r_{t,m-1}^*|, \qquad \text{with} \qquad E[BV_t^*] = IV_t. \tag{26}$$

This immediate removing of IP appears to be a common approach in the current literature (cf. Golosnoy et al. 2012; Bekierman and Gribisch 2021; Christensen et al. 2018), whereby the IP estimates $\hat{s}_m$ are based on historical data.

To contrast the relative biases and MSEs of BV_t^* and $\widetilde{BV}_t$, we generate intraday returns for $T = 250$ pre-sample days with the IP parameter value $c_1^h = 0.5$ as in (4) and (25) and use them to get IP estimates $\hat{s}_m$. We denote by c_1^h the 'historical' value of c_1 assumed to be constant during the pre-sample period. Then, we focus on the next day's IP which functional form is described by the 'current value' of c_1 denoted by c_1^c. That means, for this next (single) day of our interest, intraday returns are generated with either unchanged current IP parameter $c_1^c = 0.5 = c_1^h$ or changed parameter $c_1^c \neq c_1^h$. The latter case could occur at some 'special' days characterized by announcements, unexpected events etc. Note that the values $c_1^c = 0.1$ (extremely pronounced IP) and $c_1^c = 0.9$ (almost no IP) are not empirically relevant but considered for the illustration purposes only. We calculate BV^* and $\widetilde{BV}$ for this day of interest using the historical IP estimates $\hat{s}_m$, so there is an IP misspecification when $c_1^c \neq c_1^h$. We repeat the procedure (generating $T = 250$ pre-sample days and an additional 'day of interest') 10^4 times and put the computed relative biases and MSEs in Table 1, where we show results for constant intraday volatility in Block A and intraday SV is Block B.

Hence, we study both the effect of estimation risk in case of unchanged IP and the effect of a change in the IP form. The latter is of much practical importance, as there are many empirical confirmations for time variability of IP even after excluding days with important macroeconomic announcements, see Andersen et al. (2001), Hecq et al. (2012), or Andersen et al. (2019).

We observe in Table 1 that in case of no IP change with $c_1^h = c_1^c$, the realized measure BV^* of Boudt et al. (2011) is preferable in terms of relative bias and MSE. These findings are in line with the recent theoretical results of Ghysels et al. (2021) who show under similar assumptions that the most efficient quarticity estimators (in terms of MSE) can be obtained by an immediate adjustment of intraday returns for the IP as e.g. by Boudt et al. (2011) which is also confirmed by our evidence. However, even a comparatively small change in IP, e.g. from $c_1^h = 0.5$ to $c_1^c = 0.7$, leads to a substantial increase in the relative bias and MSE of BV^*. This evidence remains also for intraday SV in Block B. Hence, our IP-correction approach is more robust compared to the procedure of Boudt et al. (2011) which is rather sensible to even small changes in the IP form.

5 Empirical study

In our application we work with intraday returns for the Dow Jones Industrial Average Index with the focus on measuring daily IV. Our dataset consists of intraday observations from January 1996 to December 2010 with non-regular trading time days skipped. Days with non-regular trading times are those where the trading time have been substantially shorter than the common 6.5 hours because of these or that reasons.

Table 1 Relative bias and MSE of BV^* and $\widetilde{BV}$. The IP parameter c_1 equals to c_1^h during estimation period of the IP components and to c_1^c when realized measures are calculated

Relative bias in %

Block A: constant intraday volatility						Block B: intraday SV			
		$M = 78$		$M = 390$		$M = 78$		$M = 390$	
c_1^h	c_1^c	BV^*	$\widetilde{BV}$	BV^*	$\widetilde{BV}$	BV^*	$\widetilde{BV}$	BV^*	$\widetilde{BV}$
0.5	0.5	0.12	0.25	0.06	0.07	−0.08	0.05	0.02	0.03
0.5	0.1	−37.77	−0.58	−36.44	0.34	−39.69	−1.66	−38.10	−0.15
0.5	0.3	−23.24	0.22	−22.13	0.52	−25.48	−0.83	−24.19	0.01
0.5	0.7	39.60	2.44	40.13	1.05	35.94	1.53	36.32	0.46
0.5	0.9	86.69	3.64	86.88	1.35	81.93	2.83	81.77	0.71

Mean squared error (MSE)

Block A: constant intraday volatility						Block B: intraday SV			
		$M = 78$		$M = 390$		$M = 78$		$M = 390$	
c_1^h	c_1^c	BV^*	$\widetilde{BV}$	BV^*	$\widetilde{BV}$	BV^*	$\widetilde{BV}$	BV^*	$\widetilde{BV}$
0.5	0.5	0.0332	0.0567	0.0067	0.0115	0.0390	0.0703	0.0083	0.0190
0.5	0.1	0.1767	0.0819	0.1498	0.0170	0.2134	0.1053	0.1823	0.0308
0.5	0.3	0.0873	0.0705	0.0635	0.0145	0.1046	0.0896	0.0775	0.0256
0.5	0.7	0.1936	0.0439	0.1444	0.0087	0.2390	0.0524	0.1766	0.0127
0.5	0.9	0.8054	0.0369	0.6946	0.0070	0.9995	0.0432	0.8499	0.0089

We consider 15 min, 10 min, 5 min and 2 min intraday returns; 5 min returns is the most popular choice in practice. To avoid the opening bias effects, we skip the first daily observation which is the common practice to escape from the impact of noisy overnight quotes. The final sample consists of 3329 days with $M = 24, 37, 76$ or 193 observations for 15, 10, 5 or 2 min frequency, respectively.

We estimate the intraday IP s_m, and calculate the IP correction factors ζ and ξ. Then we present descriptive statistics for both uncorrected and IP-corrected realized measures. Finally, we discuss the impact of IP bias in some further realized measures which are proposed in the literature.

5.1 Estimation of intraday pattern and descriptive statistics

We estimate the IP with the non-parametric WSD estimator of Boudt et al. (2011) and show them in Fig. 5 with components $\hat{s}_m$ normalized such that $\sum_{m=1}^{M} \hat{s}_m^2 = M$. The IP pattern has a convex U-shape for all considered sampling frequencies. It is high during morning and afternoon hours and low during the lunch break. In numerical terms, it is about twice as high during the peak in the morning compared to the trough in the middle of the day. Hence, in our empirical application for Dow Jones index the IP estimate approximately corresponds to $c_1 \approx 0.7$ in our simulation study. However, the IP curvature could be more pronounced for individual stocks. For example, in

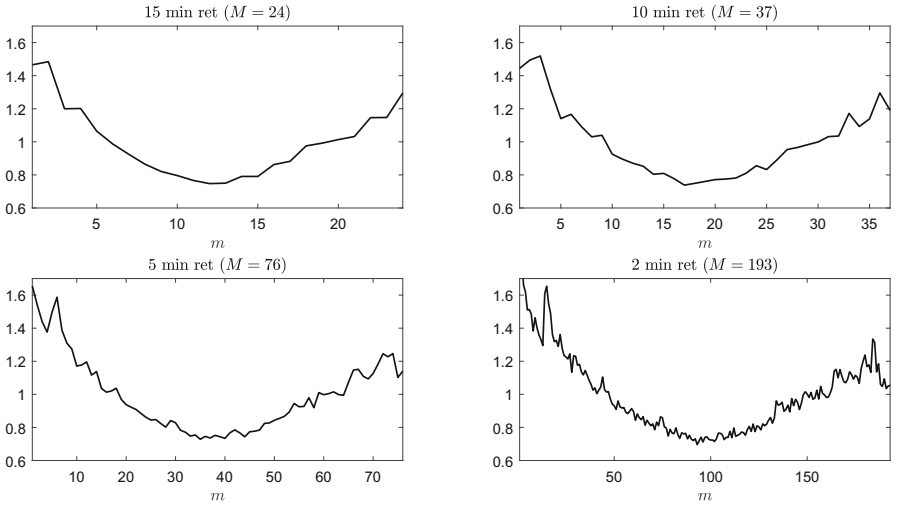

Fig. 5 Full sample estimated intraday patterns of the Dow Jones index

Table 2 Estimated empirical IP factors with bootstrapped 95% confidence intervals in parentheses

Correction factor	15 min, $M = 24$	10 min, $M = 37$	5 min, $M = 76$	2 min, $M = 193$
$\hat{\zeta}$	1.0619	1.0375	1.0190	1.0085
	(1.0607,1.0633)	(1.0368, 1.0383)	(1.0187, 1.0195)	(1.0084, 1.0088)
$\hat{\xi}_{M,TP}$	1.1325	1.0610	1.0036	0.9715
	(1.1287, 1.1364)	(1.0587, 1.0639)	(1.0017, 1.0056)	(0.9701, 0.9733)
$\hat{\xi}_{M,QP}$	1.2283	1.1225	1.0294	0.9829
	(1.2236, 1.2332)	(1.1199, 1.1258)	(1.0275, 1.0315)	(0.9815, 0.9845)
$\hat{\xi}_{M,RQ}$	1.0445	1.0452	1.0501	1.0584
	(1.0418, 1.0475)	(1.0431, 1.0478)	(1.0483, 1.0524)	(1.0566, 1.0607)

Christensen et al. (2018) the IP curvature parameter takes values around $c_1 = 0.60$ for some assets from the Dow Jones index.

Given these IP estimates, we calculate the finite M scaling factors for IP bias correction, all reported in Table 2 with 95% confidence intervals obtained by the bootstrap procedure with $10^3 - 1$ replications outlined in Dette et al. (2022). The estimated factor $\hat{\xi}_{M,RQ}$ is larger than one and numerically similar for all sampling frequencies which corresponds to the evidence of similar patterns in Fig. 5. We also estimate finite sample corrections for BV, TP, and QP. The correction estimate $\hat{\zeta}_M$ for BV gets closer to one with increasing sampling frequencies. The same holds for other finite M scaling factors so that BV is biased downward. For example, we get $\hat{\xi}_{M,QP} \approx 1.23$ for 15 min frequency but only $\hat{\xi}_{M,QP} \approx 1.03$ for 5 min frequency returns. This evidence indicates that the IP bias problem should not be very acute for the U.S. stock market during the considered period of time.

Next, we provide descriptive statistics of realized estimators. In Table 3 we report the full sample averages of both uncorrected and IP-corrected realized measures of

Table 3 Empirical averages of both uncorrected and IP-corrected realized measures

Uncorrected estimators					Corrected estimators				
Averages of	15 min	10 min	5 min	2 min	Averages of	15 min	10 min	5 min	2 min
$RV \times 10^{-3}$	0.0944	0.0971	0.1041	0.0997	$RV \times 10^{-3}$	0.0944	0.0971	0.1041	0.0997
$BV \times 10^{-3}$	0.0840	0.0899	0.0973	0.0955	$\widetilde{BV} \times 10^{-3}$	0.0892	0.0932	0.0992	0.0963
$TP \times 10^{-6}$	0.0433	0.0465	0.1081	0.2381	$\widetilde{TP} \times 10^{-6}$	0.0490	0.0493	0.1085	0.2314
$QP \times 10^{-6}$	0.0316	0.0442	0.0959	0.1919	$\widetilde{QP} \times 10^{-6}$	0.0388	0.0497	0.0987	0.1886
$(RV - BV)/RV$	0.1100	0.0748	0.0749	0.0425	$(RV\text{-}\widetilde{BV})/RV$	0.0549	0.0401	0.0472	0.0344

daily IV and IQ. The average RV is stable for all sampling frequencies, whereas the average BV is biased downwards compared to RV even after the IP correction, however, the corrected BV gets numerically closer to RV. The average values for TP and QP are almost the same for 15 and 10 min frequencies, but are much higher for 5 min and 2 min returns. Note that there are reported difficulties in estimating IQ based on frequencies of 5 min or higher (Andersen et al. 2014). Another important quantity is the average relative component $(RV - BV)/RV$ which, as expected, gets substantially smaller with the IP-bias corrected $\widetilde{BV}$ both for all sampling frequencies.

5.2 Discussion and further extensions

Our study is primarily focused on the popular realized measures which are widely used in practice. Recently, there have been much developments in estimation of IV and IQ based on high frequency observations. In particular, we would like to mention the robust threshold power variations (Corsi et al. 2010), the minRV and medRV estimators (Andersen et al. 2012), and the nearest neighborhood generalizations (Andersen et al. 2014). Moreover, in order to mitigate MMN-problems which arise by using ultra high frequency returns, several noise robust methods have been developed, as e.g. the pre-averaged BV (Podolskij and Vetter 2009), pre-averaged threshold RV (cf. Aït-Sahalia and Jacod 2014), or pre-averaged threshold BV (Christensen et al. 2014). The IP bias problem is also present for some of these more advanced realized measures as it could be shown by a simple Monte Carlo simulation exercise. To illustrate this point, we plot the IP-biases for minRV and medRV measures of Andersen et al. (2012) in Fig. 2 and observe that they are even larger than for the original BV estimator. The same argument applies to the minRQ and medRQ measures of IQ which are elaborated by Andersen et al. (2014). Of course, the IP-correction factors should be newly derived for each particular measure, however, this challenging task is beyond the scope of our paper and is left for future research, as e.g. the task of deriving IP correction factors for realized portfolio weights (Golosnoy et al. 2019, 2020, or Golosnoy and Gribisch 2022). The impact of IP on these additional estimators is investigated in the follow-up paper of Dette et al. (2022).

6 Summary

Availability of intraday high frequency returns on risky assets allows construction of precise realized estimators such as realized volatility (RV) and bipower variation

(BV) which are commonly applied for estimation of daily integrated volatility (IV), or tri-power (TP) and quad-power (QP) variations which are serving as measures for daily integrated quarticity (IQ).

In this paper we investigate the impact of intraday periodicity (IP) on the finite sample properties of these realized measures. For our analysis we assume a discrete time model for intraday returns on risky assets and postulate a multiplicative deterministic IP component which is often of U-shape empirically. For a number of intraday returns $M \to \infty$ the impact of IP is asymptotically negligible, however, we show that the IP-impact should be taken into account for a practically relevant situation with finite M. In particular, we prove that finite sample corrections of BV as well as of TP and QP measures are necessary to obtain valid statistical inferences concerning daily IV. We derive analytically the factors for IP-correction and analyse their stochastic properties. Our results are illustrated by means of a Monte Carlo simulation study for both constant and stochastic intraday volatility models. Finally, we evaluate IP correction factors empirically for daily IV of the Dow Jones Industrial Average Index.

Acknowledgements this work was partly supported by the Collaborative Research Center 'Statistical modeling of nonlinear dynamic processes' (SFB823, projects A1,C1) of German Research Foundation (DFG).

Funding Open Access funding enabled and organized by Projekt DEAL.

Declarations

Conflict of interest statement On behalf of all authors, the corresponding author states that there is no conflict of interest.

Appendix

Proofs

Proof of Proposition 1 Note that the statement for RV_t is obvious and consider only $\overline{BV}_t = \frac{\pi}{2}\sum_{m=2}^{M}|r_{t,m}||r_{t,m-1}|$ to simplify calculations and noting that $E[BV_t] = \frac{M}{M-1}E[\overline{BV}_t]$. We use the fact that $E[|X|] = \sqrt{2/\pi}\,\sigma$ if X is a centered normal distributed random variable with variance σ^2, which implies

$$E[\overline{BV}_t] = \frac{\sigma_t^2}{M}\sum_{m=2}^{M} s_m s_{m-1} = \sigma_t^2 \frac{\sum_{m=2}^{M}\left[g\left(\frac{m}{M}\right)g\left(\frac{m-1}{M}\right)\right]^{1/2}}{\sum_{m=1}^{M} g\left(\frac{m}{M}\right)} = \sigma_t^2(1 - R_M),$$

where the term R_M is given by

$$R_M = \frac{g\left(\frac{1}{M}\right) + \sum_{m=2}^{M} g\left(\frac{m}{M}\right)^{1/2}\left(g\left(\frac{m}{M}\right)^{1/2} - g\left(\frac{m-1}{M}\right)^{1/2}\right)}{\sum_{m=1}^{M} g\left(\frac{m}{M}\right)}.$$

We assume that g is continuously differentiable, therefore g and g' are continuous and Riemann integrable on the interval $[0, 1]$, and also Lebesgue-integrable. Hence, the squares of g and g' are Riemann integrable as well because the squares are also continuous. Then we obtain by the mean value theorem

$$R_M = \frac{\frac{1}{M}\sum_{m=2}^{M} g\left(\frac{m}{M}\right)^{1/2}\left(g^{1/2}\right)'(\psi_m) + g\left(\frac{1}{M}\right)}{\sum_{m=2}^{M} g\left(\frac{m}{M}\right)},$$

where $\psi_m \in (\frac{m-1}{M}, \frac{m}{M}), m = 2, \ldots, M$. Consequently, using the fact that $(g^{1/2}(x))' = \frac{1}{2}\frac{g'(x)}{g^{1/2}(x)}$ and an approximation of the sums by a Riemann integral it follows that

$$M \cdot R_M = \left(\frac{1}{2}\frac{\int_0^1 g'(x)dx}{\int_0^1 g(x)dx} + \frac{g(0)}{\int_0^1 g(x)dx}\right) \cdot (1 + o(1)),$$

which proves the statement (A) of Proposition 1.

For the statement (B) in order to compute $Var(BV_t)$ we first look at $E(BV_t)^2$, setting $r_{t,m} = r_m$ for notation simplification. We have

$$E[(\overline{BV}_t)^2] = \frac{\pi^2}{4} E\left[\sum_{m_1=2}^{M}\sum_{m_2=2}^{M} |r_{m_1}||r_{m_1-1}|]|r_{m_2}||r_{m_2-1}|\right]$$

$$(E[\overline{BV}_t])^2 = \left[\frac{\sigma_t^2}{M}\sum_{m=2}^{M} s_m s_{m-1}\right]^2 = \frac{\sigma_t^4}{M^2}\left[\sum_{m=2}^{M} s_m s_{m-1}\right]^2.$$

As the random variables r_m are independent, the double sum contains three types of non-vanishing expectations

$$E(|r_{m_1}^2||r_{m_1-1}^2|) = \frac{\sigma_t^4}{M^2} s_m^2 s_{m-1}^2,$$

$$E(|r_{m_1}||r_{m_1-1}^2||r_{m_1-2}|) = \sqrt{\frac{2\sigma_t^2}{\pi M} s_{m_1}^2}\,\frac{\sigma_t^2}{M} s_{m_1-1}^2 \sqrt{\frac{2\sigma_t^2}{\pi M} s_{m_1-2}^2} = \frac{2\sigma_t^4}{\pi M^2} s_{m_1} s_{m_1-1}^2 s_{m_1-2},$$

$$E(|r_{m_1}||r_{m_1-1}||r_{m_2}||r_{m_2-1}|) = \frac{4\sigma_t^4}{\pi^2 M^2} s_{m_1} s_{m_1-1} s_{m_2} s_{m_2-1}.$$

This yields

$$Var(\overline{BV}_t) = \frac{\pi^2}{4}\frac{\sigma_t^4}{M^2}\left\{\sum_{m_1=2}^{M} s_m^2 s_{m-1}^2 + \frac{4}{\pi}\sum_{m_1=3}^{M} s_{m_1} s_{m_1-1}^2 s_{m_1-2} + \frac{4}{\pi^2}\underbrace{\sum_{m_1=2}^{M}\sum_{m_2=2}^{M}}_{|m_1-m_2|>1} s_{m_1} s_{m_1-1} s_{m_2} s_{m_2-1}\right\} - \frac{\sigma_t^4}{M^2}\left[\sum_{m=2}^{M} s_m s_{m-1}\right]^2.$$

Observing the definition of s_m^2 in (8), we get

$$Var(\overline{BV}_t) = \frac{\pi^2}{4}\frac{\sigma_t^4}{M^2 g_M^2}\left\{(1-\frac{4}{\pi^2})\sum_{m_1=2}^{M} g\left(\frac{m_1}{M}\right) g\left(\frac{m_1-1}{M}\right) + \frac{4}{\pi}\left(1-\frac{2}{\pi}\right)\sum_{m_1=3}^{M}\left\{g\left(\frac{m_1}{M}\right)\right\}^{1/2} g\left(\frac{m_1-1}{M}\right)\left\{g\left(\frac{m_1-2}{M}\right)\right\}^{1/2}\right\},$$

which proves the first assertion regarding the variance of $\overline{BV}_t$. The approximation (9) finally follows by interpreting the sum as approximations of a Riemann integral. The expressions for the variance $Var(RV_t)$ and the covariance $Cov(RV_t, \overline{BV}_t)$ are obtained by similar arguments. However, we omit them for the sake of brevity. Therefore, the statement (B) of Proposition 1 follows. □

Proof of Proposition 2 Recall the definition of QP_t in (2) and (8), then

$$\begin{aligned} E[QP_t] &= M\frac{\pi^2}{4}\left\{\sum_{m=4}^{M} E\left[|r_{m-3}|\right] E\left[|r_{m-2}|\right] E\left[|r_{m-1}|\right] E\left[|r_m|\right]\right\} \\ &= \frac{\sigma_t^4}{M}\sum_{m=4}^{M} s_{m-4} s_{m-2} s_{m-1} s_m \\ &= \frac{\sigma_t^4}{M}(g_M)^{-2}\sum_{m=4}^{M}\left[g\left(\frac{m-3}{M}\right) g\left(\frac{m-2}{M}\right) g\left(\frac{m-1}{M}\right) g\left(\frac{m}{M}\right)\right]^{1/2}. \end{aligned}$$

Then we get $\lim_{M\to\infty} E[QP_t] = \sigma_t^4 \xi / M\,(1+o(1))$ which follows again by a Riemann integral. The statement for RQ is obtained by analogy.

Next, consider TP_t defined in (3) and observe that $E[|Z|^{4/3}]/\sigma^{4/3} = \mu_{4/3} = 2^{2/3}\pi^{-1/2}\Gamma(7/6)$ if Z is a centered normal distributed random variable with variance σ^2. This yields observing (8)

$$E[TP_t] = M\mu_{\frac{4}{3}}^{-3} \sum_{m=3}^{M} \left\{ E\left[|r_{m-2}|^{\frac{4}{3}}\right] E\left[|r_{m-1}|^{\frac{4}{3}}\right] E\left[|r_m|^{\frac{4}{3}}\right] \right\} = \frac{\sigma_t^4}{M} \sum_{m=3}^{M} s_{m-2}^{\frac{4}{3}} s_{m-1}^{\frac{4}{3}} s_m^{\frac{4}{3}}$$
$$= \frac{\sigma_t^4}{M} (g_M)^{-2} \sum_{m=3}^{M} \left[g\left(\frac{m-2}{M}\right) g\left(\frac{m-1}{M}\right) g\left(\frac{m}{M}\right) \right]^{2/3}.$$

The statement $\lim_{M\to\infty} E[TP_t] = \sigma_t^4 \xi / M \ (1 + o(1))$ completes the proof of Proposition 2. □

Proof of Theorem 1 It follows from Proposition 2 that

$$V(BV_t) = \frac{\sigma_t^4}{M}\left(\frac{\pi^2}{4} + \pi - 3\right)\xi \cdot (1 + o(1)),$$
$$V(RV_t) = 2\frac{\sigma_t^4}{M}\xi \cdot (1 + o(1)),$$
$$Cov(RV_t, BV) = 2\frac{\sigma_t^4}{M}\xi \cdot (1 + o(1)),$$

where $\xi = \int_0^1 g^2(x)dx \Big/ \left(\int_0^1 g(x)dx\right)^2 \geq 1$ by the Cauchy–Schwarz inequality. Therefore, Theorem 1 is a consequence of a straightforward application of a central limit theorem for triangular arrays of m-dependent random variables (Romano and Wolf 2000) and the Cramer–Wold device. □

Proof of Proposition 3 Recall the definition of $\hat{s}_m$ in (14), then we obtain the representation

$$\widehat{\mathbf{S}}_M = \left(\hat{s}_1^2, \ldots, \hat{s}_M^2\right)^\top = g_1(\widehat{\mathbf{SD}}_{M,T}),$$

with $\widehat{\mathbf{SD}}_{M,T} = (\widehat{SD}_{1,T}^2, \ldots, \widehat{SD}_{M,T}^2)^\top$ defined by (14), and the function $g_1 : \mathbb{R}^M \to \mathbb{R}^M$ is given by

$$g_1(x) = (g_{11}(x), \ldots, g_{1M}(x))^\top = \frac{M}{\sum_{m=1}^M x_m}(x_1, \ldots, x_M)^\top. \tag{27}$$

Let

$$g_1'(x) = \frac{M}{\sum_{m=1}^M x_m}\left\{ \boldsymbol{I}_M - \frac{1}{\sum_{m=1}^M x_m}(x_1, \ldots, x_M)^\top \mathbf{1}_M^\top \right\}$$

denote the derivative of g_1 whereby $\boldsymbol{I}_M$ is the $M \times M$ identity matrix, then, by (16), it follows that

$$\lim_{T\to\infty} g_1'(\mathbf{SD}_{M,T}) = \frac{M}{\sigma^2}\boldsymbol{I}_M - \frac{1}{\sigma^2}\boldsymbol{S}_M \mathbf{1}_M^\top, \tag{28}$$

where $\boldsymbol{S}_M = (s_1^2, \ldots, s_M^2)^\top$ and we have used the fact that $\sum_{m=1}^M SD_{m,T}^2 = \sigma_T^2 \to \sigma^2$. Now a componentwise Taylor expansion gives

$$\sqrt{T}\left(g_1(\widehat{\mathbf{SD}}_{M,T}) - g_1(\mathbf{SD}_{M,T})\right) = g_1'(\mathbf{SD}_{M,T})\left(\widehat{\mathbf{SD}}_{M,T} - \mathbf{SD}_{M,T}\right) + \boldsymbol{\Delta}_{M,T} \,. \quad (29)$$

The components of the vector $\boldsymbol{\Delta}_{M,T} = (\Delta_{1,T}, \ldots, \Delta_{M,T})^\top$ are given by

$$\Delta_{m,T} = \sqrt{T}\left(\widehat{\mathbf{SD}}_{M,T}) - \mathbf{SD}_{M,T}\right)^\top \nabla^2 g_{1m} \left((1-\psi_m)\widehat{\mathbf{SD}}_{M,T} + \psi_m \mathbf{SD}_{M,T}\right)\left(\widehat{\mathbf{SD}}_{M,T} - \mathbf{SD}_{M,T}\right) \,,$$

where $\nabla^2 g_{1m}$ denotes the Hessian matrix of the m-th component g_{1m} of the function g_1 and $\psi_m \in (0,1)$ is an intermediate point, $m = 1, \ldots, M$. By (17) and (16) we have

$$\lim_{T\to\infty} \mathbf{SD}_{M,T} = \frac{\sigma^2}{M}\mathbf{S}_M \quad \text{and} \quad \lim_{T\to\infty} \widehat{\mathbf{SD}}_{M,T} = \frac{\sigma^2}{M}\mathbf{S}_M \quad \text{in probability} \,,$$

and $\Delta_{m,T} = O_P(1/\sqrt{T})$. Note that g_1' and $\nabla^2 g_{1m}$ are continuous at the point $\frac{\sigma^2}{M}\mathbf{S}_M$, $m = 1, \ldots, M$. Consequently, it follows from (28), (29) and (17) that

$$\begin{aligned} &\sqrt{T}\left(\left(\hat{s}_1^2, \ldots, \hat{s}_M^2\right)^\top - \left(s_1^2, \ldots, s_M^2\right)^\top\right) \\ &= \sqrt{T}\left(g_1(\widehat{\mathbf{SD}}_{M,T}) - g_1(\mathbf{SD}_{M,T})\right) \xrightarrow{L} \mathcal{N}(\mathbf{0}, \Sigma_1) \,, \end{aligned} \quad (30)$$

where the $M \times M$ matrix Σ_1 is given by

$$\Sigma_1 = M^2\Sigma - M\left(\mathbf{S}_M \mathbf{1}_M^\top \Sigma + \Sigma \mathbf{1}_M \mathbf{S}_M^\top\right) + \mathbf{1}_M^\top \Sigma \mathbf{1}_M \mathbf{S}_M \mathbf{S}_M^\top \,.$$

Now note that $\hat{\xi}_M = g_2(\hat{\mathbf{S}}_M)$, where the function $g_2 : \mathbb{R}^M \to \mathbb{R}$ is defined by

$$g_2(x) = \frac{1}{M}\sum_{m=1}^M x_m^2 \,.$$

Observing that $\nabla g_2(x) = \frac{2}{M}(x_1 \ldots, x_M)$ assertion (18) follows by the Delta-method, where a straightforward calculation yields the limiting variance in (20). Similarly, (19) is obtained observing that $\hat{\zeta}_M = g_3(\hat{\mathbf{S}}_M)$, where the function $g_3 : \mathbb{R}^M \to \mathbb{R}$ is defined by

$$\begin{aligned} g_3(x) &= \frac{1}{M}\sum_{m=2}^M \sqrt{x_m x_{m-1}} \,, \\ \nabla g_3(x) &= \frac{1}{2M}\left(\frac{\sqrt{x_2}}{\sqrt{x_1}}, \frac{\sqrt{x_1}+\sqrt{x_3}}{\sqrt{x_2}}, \ \ldots \ , \frac{\sqrt{x_{M-2}}+\sqrt{x_M}}{\sqrt{x_{M-1}}}, \frac{\sqrt{x_{M-1}}}{\sqrt{x_M}}\right)^\top \,. \end{aligned}$$

Finally, if Σ is a diagonal matrix as specified in Proposition 3 the formulas (22) and (23) follow by a direct calculation which completes the proof of Proposition 3.

Moreover, the arguments provided in this proof show that the result of Proposition 3 would remain correct for any estimator $\left(\hat{s}_1^2, \ldots, \hat{s}_M^2\right)^\top$, for which an asymptotic normality in the sense (30) can be established. However, the limiting variances τ_1^2 and τ_2^2 may change as they depend on the specific form of the estimator. □

Implementation of the *WSD* estimator

In the following we outline the implementation of the WSD estimator as in Boudt et al. (2011). Let $m \in \{1, \ldots, M\}$ be a fixed intraday interval and compute the returns $x_{t,m} = r_{t,m}/\hat{\sigma}_t$ where $\hat{\sigma}_t^2$ is the BV measure. Next, consider the T order statistics $x_{(1),m} \leq x_{(2),m} \leq \cdots \leq x_{(T),m}$. Let $h_T = \lfloor T/2 \rfloor + 1$. We first calculate $\mathrm{SH}_m = 0.741 \cdot \min(x_{(h_T),m} - x_{(1),m}, \ldots, x_{(T),m} - x_{(T-h_T+1),m})$ and then

$$\hat{s}_m^{\mathrm{SH}} = \frac{\mathrm{SH}_m}{\sqrt{\frac{1}{M}\sum_{j=1}^M \mathrm{SH}_j^2}}.$$

The WSD estimates are then given by

$$\hat{s}_m = \frac{\mathrm{WSD}_m}{\sqrt{\frac{1}{M}\sum_{j=1}^M \mathrm{WSD}_j^2}} \quad \text{with} \quad WSD_j = \sqrt{1.081 \cdot \frac{\sum_{t=1}^T w_{t,j} x_{t,j}^2}{\sum_{t=1}^T w_{t,j}}}.$$

The weights are given by $w_{t,j} = \theta(x_{t,j}/\hat{s}_m^{\mathrm{SH}})$ and the weight function is $\theta(z) = \mathbb{1}(z^2 \leq 6.635)$ where $\mathbb{1}(\cdot)$ denotes the indicator function.

References

Admati A, Pfleiderer P (1988) A theory of intraday patterns: volume and price variability. Rev Financ Stud 1:3–40

Aït-Sahalia Y, Jacod J (2014) High-Frequency Financial Econometrics. Princeton University Press, New Jersey

Andersen T, Bollerslev T (1997) Intraday periodicity and volatility persistence in financial markets. J Empir Financ 4:115–158

Andersen T, Bollerslev T (1998) Answering the skeptics: yes, standard volatility models do provide accurate forecasts. Int Econ Rev 39:885–905

Andersen T, Bollerslev T (1998) Deutsche Mark-Dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. J Financ 53:219–265

Andersen T, Bollerslev T, Das A (2001) Variance-ratio statistics and high-frequency data: testing for changes in intraday volatility patterns. J Financ 56:305–327

Andersen T, Bollerslev T, Diebold F, Vega C (2003) Micro effects of macro announcements: real-time price discovery in foreign exchange. American Econ Rev 93:38–62

Andersen T, Bollerslev T, Huang X (2011) A reduced form framework for modeling volatility of speculative prices based on realized variation measures. J Econom 160:176–189

Andersen T, Dobrev D, Schaumburg E (2012) Jump-robust volatility estimation using nearest neighbor truncation. J Econometr 138:125–180
Andersen T, Dobrev D, Schaumburg E (2014) A robust neighborhood truncation approach to estimation of integrated quarticity. Econometr Theor 30:3–59
Andersen T, Thyrsgaard M, Todorov V (2019) Time varying periodicity in intraday volatility. J Am Stat Assoc 114:1695–1707
Barndorff-Nielsen O, Shephard N (2004) Power and bipower variation with stochastic volatility and jumps. J Financ Econometr 2:1–37
Bekierman J, Gribisch B (2016) Estimating stochastic volatility models using realized measures. Stud Nonlinear Dyn Econom 20:279–300
Bekierman J, Gribisch B (2021) A mixed frequency stochastic volatility model for intraday stock market returns. J Financ Econometr 19(3):496–530
Bollerslev T, Litvinova J, Tauchen G (2006) Leverage and volatility feedback effects in high-frequency data. J Financ Econometr 4:353–384
Boudt K, Croux C, Laurent S (2011) Robust estimation of intraweek periodicity in volatility and jump detection. J Empir Financ 18:353–367
Christensen K, Hounyo U, Podolskij M (2018) Is the diurnal pattern sufficient to explain the intraday variation in volatility: A nonparametric assessment. J Econometr 205:336–362
Christensen K, Oomen R, Podolskij M (2014) Fact or friction: Jumps at ultra high frequency. J Financ Econ 114:576–599
Corsi F, Pirino D, Reno R (2010) Threshold bipower variation and the impact of jumps on volatility forecasting. J Econometr 114:576–599
Dehling H, Denker M, Philipp W (1986) Central limit theorems for mixing sequences of random variables under minimal conditions. Ann Probab 14(4):1359–1370
Dette H, Golosnoy V, Kellermann J (2022) Correcting intraday periodicity bias in realized volatility measures. Econometr Stat 23:36–52
Engle R, Ito T, Lin W-L (1990) Meteor showers or heat waves? Heteroskedastic intra-daily volatility in the foreign exchange market. Econometrica 58:525–542
Engle R, Sokalska M (2012) Forecasting intraday volatility in the US equity market. Multiplicative component GARCH. J Financ Econometr 10:54–83
Gabrys R, Hörmann S, Kokoszka P (2013) Monitoring the intraday volatility pattern. J Time Series Econometr 5:87–116
Ghysels E, Mykland P, Renault E (2021) In-sample asymptotics and across-sample efficiency gains for high frequency data statistics. Econometr Theor. https://doi.org/10.1017/S0266466621000359
Golosnoy V, Gribisch B (2022) Modeling and forecasting realized portfolio weights. J Bank Finance 138:106404
Golosnoy V, Gribisch B, Seifert MI (2019) Exponential smoothing of realized portfolio weights. J Empir Financ 53:222–237
Golosnoy V, Köhler S, Schmid W, Seifert MI (2021) Tests for validity of linear state space representations. Appl Stoch Model Bus Ind 37:1060–1079
Golosnoy V, Okhrin I, Schmid W (2012) Statistical surveillance of volatility forecasting models. J Financ Econometr 10:513–545
Golosnoy V, Schmid W, Seifert MI, Lazariv T (2020) Statistical inferences for realized portfolio weights. Econometr Stat 14:49–62
Goncalves S, Meddahi N (2009) Bootstrapping realized volatility. Econometrica 77:283–306
Harris L (1986) A transaction data study of weekly and intradaily patterns in stock returns. J Financ Econ 16:99–117
Hasbrouck J (1999) The dynamics of discrete bid and ask quotes. J Financ 54:2109–2142
Hecq A, Laurent S, Palm F (2012) Common intraday periodicity. J Financ Econometr 10:325–353
Kokoszka P, Reimherr M (2013) Predictability of shapes of intraday price curves. Econometr J 16:285–308
McAleer M, Medeiros M (2008) Realized volatility: a review. Econometr Rev 27:10–45
Podolskij M, Vetter M (2009) Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Bernoulli 15:634–658
Romano J, Wolf M (2000) A more general central limit theorem for m-dependent random variables with unbounded m. Stat Prob Letters 47:115–124
Wood R, McInish T, Ord J (1985) An investigation of transaction data for NYSE stocks. J Financ 25:723–739

Wu WB (2005) Nonlinear system theory: Another look at dependence. Proc Natl Acad Sci USA 102(40):14150–14154