Correcting Intraday Periodicity Bias in Realized Volatility Measures

Holger Dette,[a] Vasyl Golosnoy,[b] and Janosch Kellermann[c]

July 8, 2020

Abstract

Diurnal fluctuations in volatility are a well-documented stylized fact of intraday price data. We investigate how this intraday periodicity (IP) affects both finite sample as well as asymptotic properties of several popular realized estimators of daily integrated volatility which are based on functionals of M intraday returns. We demonstrate that most of the estimators considered in our study exhibit a finite-sample bias due to IP, which can however get negligible if the number of intraday returns diverges to infinity. We suggest appropriate correction factors for this bias based on estimates of the IP. The adequacy of the new corrections is evaluated by means of a Monte Carlo simulation study and an empirical example.

JEL: C14, C15, C58

Keywords: integrated volatility, realized measures, intraday periodicity, simulation-based methods

[a] Faculty of Mathematics, Ruhr-Universität Bochum, *E-mail:* holger.dette@rub.de

[b] Faculty of Management and Economics, Ruhr-Universität Bochum, Universitätsstr. 150, 44801 Bochum, Germany, Tel.: +49(0)234-32-22917, Fax: +49(0)234-32-14528, *E-mail:* vasyl.golosnoy@rub.de

[c] Faculty of Management and Economics, Ruhr-Universität Bochum, *E-mail:* janosch.kellermann@rub.de

1 Introduction

Measurement of daily integrated volatility (IV) of financial risky assets is a task with much empirical relevance. The class of realized volatility measures which are based on intraday high-frequency returns allows to estimate daily IV with a high precision. During the last two decades many different types and refinements of realized estimators have been suggested, mostly in order to resolve estimation problems associated with jumps and market microstructure noise. The asymptotic properties of these estimators are nowadays well understood if the number M of intraday returns satisfies $M \to \infty$. For example, Jacod and Protter (2014) provide the laws of large numbers and central limit theorems for a quite general class of such estimators.

In this paper we focus on another stylized fact of intraday returns, namely the intraday periodicity (IP), and investigate its impact on various realized volatility measures. The assessment of IP in realized volatility contexts was already investigated in the early papers of Andersen and Bollerslev (1997) and Taylor and Xu (1997), however, this topic has also gained substantial attention recently, see e.g. the contributions of Gabrys et al. (2013), Dette et al. (2016), Christensen et al. (2018), Andersen et al. (2019), Dumitru et al. (2019).

In particular, Dette et al. (2016) is the first paper with analytical statements for quantifying a (downward) IP bias in the well-known bipower variation (BV) measure of daily IV. They compute a closed form expression for this bias depending on the IP profile specification which could be of a quite arbitrary functional form. Although this IP-bias in BV is negligible asymptotically for $M \to \infty$, it could make a substantial impact for finite M leading to underestimation of IV which is highly undesired from the risk management perspective. Further, they derive expressions for IP-bias in tri-power (TP) and quad-power (QP) variation which are popular measures for integrated quarticity (IQ) required for statical inference on IV. In contrast to the BV case, the IP-bias in TP and QP remains also visible in the case $M \to \infty$. This happens because the presence of IP changes the 'true' value of IQ, i.e. when $IQ = 1$ in case of no IP, it would hold that $IQ = \xi > 1$ in case of IP, see Theorem 1 in Dette et al. (2016). All this evidence underscores the importance of IP-bias quantification and correction in realized measures.

Extending the results of Dette et al. (2016), in this paper we provide a systematic investigation of the IP bias correction methods for the most popular realized volatility estimators both analytically as well as by means of Monte Carlo simulations. Our approach is based on computing a scalar IP-correction factor as a function of the estimated IP profile for each type of realized volatility estimator. The advantage of such an approach is that our scalar correction factor is a kind of average over IP profile components which provides certain robustness properties. In particular, we derive the analytical expression for IP-bias in the popular $\min RVRV$ estimator, and show how to compute correction factors for the other considered realized estimators in simulations.

We analyze the performance of our approach in an extensive simulation study where we quantify IP-biases for various realized IV estimators depending on the amount of IP curvature and the number of intraday returns M. We also consider estimation of IP curvatures and compare our IP-correction procedure with those of Boudt et al. (2011) which is based on IP-filtered returns. The latter approach suggests an immediate correction of each intraday return by its corresponding IP component which makes it more sensitive to IP profile misspecifications compared to the method proposed in this paper. In general, we find that our IP-bias correction approach works well both in simulations and empirically. As the procedure proposed by us appears to be in some sense more robust than the approach of Boudt et al. (2011), we recommend it for practical applications.

The remaining part of this paper is organized as follows: in Section 2 we introduce the daily volatility IV and its popular realized estimators based on intraday returns. In Section 3 we consider the IP modeling, discuss and quantify the impact of IP on the estimators of IV, and introduce our methodology for IP bias correction. The performance of our correction method is investigated in Section 4 by means of a simulation study, whereas in Section 5 we provide an empirical illustration. Section 6 concludes the paper, the proofs are placed in the Appendix.

2 Daily integrated volatility (IV) and its realized estimators

2.1 Model setup

We assume that the log-price process of a risky financial asset is given as

$$p(t) = \mu(t)dt + \sigma(t)dW(t), \tag{1}$$

where $\mu(t)$ is the drift, $dW(t)$ is the increment of the Brownian motion and $\sigma(t)$ is a cadlag process under rather general assumptions (cf. Barndorff-Nielsen and Shephard, 2004), which can incorporate for example stochastic volatility and intraday periodicity (IP). We do not consider jumps in this setting as we are interested in isolating the finite sample bias caused by a deterministic IP, while jumps constitute another (different) source of bias in realized measures (Andersen et al., 2012).

Our quantity of interest is the daily integrated volatility (IV)

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

which needs to be estimated from discretized data, i.e. from high-frequency returns. In this paper we consider several popular estimators of IV_t, for which we quantify and correct the bias in these estimators caused by IP.

To set up the finite sample analysis, we consider M log returns sampled from an evenly spaced intraday time grid per each day t. Since our objective is the analysis of properties of IV-estimators for finite M values, for the rest of the paper we exploit the discrete time model

$$r_{m,t} = s_m \nu_{t,m} z_{t,m}, \qquad z_{t,m} \sim \text{ iid}\,\mathcal{N}(0,1), \qquad m = 1, ..., M,$$

which is a standard representation in many recent studies (cf. Bekierman and Gribisch, 2020).

The deterministic diurnal IP-components s_m summarized to the vector $\boldsymbol{s} := (s_1, ..., s_M)'$ are allowed to vary during the day, but are assumed to be stable for all days under consideration. We address the issue of time invariance of IP profile $\boldsymbol{s}$ later in this paper. The assumptions about $z_{t,m}$ result from the continuous time framework above in (1). The (stochastic) volatility component $\nu_{t,m}$ could be either constant or varying both inter- and intra-daily. As we focus on the effect of IP on IV estimation, we set $\nu_{t,m} = \sigma^2/M$ for the rest of the paper, which could be justified by analysis in Dette et al. (2016) who show that the impact of persistent intraday stochastic volatility on realized measures is usually of smaller order. In order to disentangle the diurnal part of volatility which causes IP, we impose the common identification restriction (Boudt et al., 2011)

$$\sum_{m=1}^{M} s_m^2 = M,$$

hence, in our setting it holds that $IV_t = \sigma^2$.

2.2 Realized estimators of IV

Now we introduce several popular realized estimators for IV which are commonly used in empirical research. Hereafter we skip day t index for the simplification of notation. Since our interest is to quantify the effect of IP, we also consider some estimators that are not robust to jumps or microstructure noise. Note that even though microstructure noise is mainly a problem when M is very large (e.g., in case of tick data or data sampled under ultra-high frequencies), it might also be relevant in finite M situations when risky assets are not liquid enough.

The most simple realized volatility (RV) estimator is given by

$$RV = \sum_{m=1}^{M} r_m^2, \tag{2}$$

whereby Barndorff-Nielsen and Shephard (2002) provide the asymptotic theory of RV if $M \to \infty$. While RV is the efficient estimator for IV, it becomes inconsistent when the price process exhibits jumps during the day. As a remedy, Barndorff-Nielsen and Shephard (2004) suggest the class of multipower variations which are based on products of adjacent returns. The most commonly used estimator of this class is the bipower variation (BV) defined by

$$BV = \frac{M}{M-1}\frac{\pi}{2}\sum_{m=2}^{M}|r_m||r_{m-1}|. \tag{3}$$

While BV is a consistent IV-estimator in the case of jumps, there is a finite sample bias caused by the jumps especially when M is small. In an attempt to reduce this bias without much loss in efficiency, Andersen et al. (2012) suggest the minimum RV (minRV) and the median RV (medRV) estimators given by

$$\text{min}RV = \frac{\pi}{\pi-2}\frac{M}{M-1}\sum_{m=2}^{M} min(|r_m|,|r_{m-1}|)^2, \tag{4}$$

$$\text{med}RV = \frac{\pi}{6-4\sqrt{3}+\pi}\frac{M}{M-2}\sum_{m=3}^{M} med(|r_m|,|r_{m-1}|,|r_{m-2}|)^2. \tag{5}$$

The measures minRV and medRV can also been seen as special cases of the quantile-based RV measures proposed by Christensen et al. (2010). Further extensions of these estimators are discussed by Andersen et al. (2014), who consider general order statistics of functionals computed from blocks of returns.

Besides jumps, market microstructure noise is another source for possible bias and inconsistency of IV estimators which motivate further realized estimators of daily IV, such as the class of realized kernel (RK) estimators suggested by Barndorff-Nielsen et al. (2008). Given a suitable kernel function $\mathcal{K}(\cdot)$, the RK estimator is given by

$$RK = \sum_{h=-H}^{H}\mathcal{K}\left(\frac{h}{H+1}\right)\gamma_h, \qquad \gamma_h = \sum_{m=|h|+1}^{M} r_m r_{m-|h|}. \tag{6}$$

The choice of kernel function $\mathcal{K}(\cdot)$ and of the bandwidth H are discussed in detail in Barndorff-Nielsen et al. (2009). To obtain a measure which is robust against market microstructure noise, Jacod et al. (2009) suggest an alternative approach based on pre-averaging of intraday returns which is usually applied for intraday returns sampled at 30 seconds or even higher frequencies: Intraday returns are first averaged over a local window of size K and then an estimate is computed from these pre-averaged returns. The pre-averaged RV ($paRV$) measure is therefore given by

$$paRV = \frac{M}{M-K+2}\frac{1}{K\psi_2^K}\sum_{m=0}^{M-K+1}\bar{r}_m^2, \qquad \bar{r}_m = \sum_{j=1}^{K-1} g\left(\frac{m}{K}\right) r_{m+j}, \tag{7}$$

where $\psi_2^K = (1/K)\sum_{j=1}^{K-1} g^2(j/K)$ with the function $g(x) = \min(x, 1-x)$ and the usual choice $K = M^{1/2}$. Since neither RK nor $paRV$ estimators are robust to jumps, Christensen et al. (2014) further suggest a pre-averaged version of the bipower variation given by

$$paBV = \frac{M}{M-2K+2}\frac{\pi}{2K\psi_2^K}\sum_{m=0}^{M-2K+1}|\bar{r}_m||\bar{r}_{m+K}|, \qquad \bar{r}_m = \sum_{j=1}^{K-1} g\left(\frac{m}{K}\right) r_{m+j}, \tag{8}$$

which is also consistent for IV in the presence of jumps. To summarize, we have reviewed seven popular measures – RV, BV, $\min RV$, $\text{med}RV$, RK, $paRV$ and $paBV$ – as estimators of daily IV. Next we focus on the impact of IP profile $\boldsymbol{s} = (s_1, ..., s_M)'$ on these estimators for finite M number of intraday returns.

3 IP-bias in realized estimators of IV

The estimators of daily IV introduced in Section 2 are based on the sum of specified functionals of intraday returns and could therefore be expressed in a general form as

$$\widehat{IV} = \sum_{m=j+1}^{M-k} h(r_{m-j}, \ldots, r_m, \ldots, r_{m+k}). \tag{9}$$

For a proper estimator-specific correction factor $C(\boldsymbol{s})$ it should hold that

$$C(\boldsymbol{s}) \cdot \mathbb{E}(\widehat{IV}) = IV = \sigma^2, \tag{10}$$

whereby the scaling factor $C(\boldsymbol{s})$ depends both on the functional $h(\cdot)$ characterizing each estimator and on the specific IP profile $\boldsymbol{s} = (s_1, ..., s_M)'$. Hence, for a particular estimator $\widehat{IV}$ of the form (9) the presence of IP could cause a bias which should be completely quantified and eliminated by a multiplicative correction with the factor $C(\boldsymbol{s})$. This means, two distinct profiles with $\boldsymbol{s}_1 \neq \boldsymbol{s}_2$ which cannot be obtained from each other by a structure-preserving permutation could lead to the same correction factors $C(\boldsymbol{s}_1) = C(\boldsymbol{s}_2)$. This makes our IP-bias correction attractive from the empirical applicability perspective, as it is only a scalar $C(\boldsymbol{s})$ we need to compute for each estimator of IV. Even in case of iid intraday returns an analytical expression for $C(\boldsymbol{s})$ is possible only for some special forms of the functional $h(\cdot)$. Already Andersen et al. (2014) note that closed form expressions can be hardly obtained for more advanced functionals even in the iid case and point out the need to obtain them by means of Monte Carlo simulations. Of course, in presence of IP calculation of $C(\boldsymbol{s})$ becomes more challenging as then the iid assumption is violated (cf. Andersen et al., 2014).

In the next proposition we provide the finite sample IP-biases under quite general conditions on the IP form for some of the above-mentioned IV-estimators. We rewrite the normalized IP components s_m^2 for $m = 1, ..., M$ as

$$s_m^2 = g(\tfrac{m}{M}) \,/\, g_M, \qquad \text{with} \qquad g_M = \frac{1}{M} \sum_{m=1}^{M} g(\tfrac{m}{M}), \tag{11}$$

where $g : [0, 1] \mapsto \mathbb{R}$ is a given continuously differentiable function.

Proposition 1. *Assume that $r_m \sim \mathcal{N}(0, s_m^2\sigma^2/M)$ with $\sum_{m=1}^{M} s_m^2 = M$, and the IP component as in (11) for some function $g : [0, 1] \mapsto \mathbb{R}$. The estimator RV for IV is unbiased so that $\mathbb{E}(RV) = \sigma^2$. The estimators BV and $\min RV$ are biased for finite M such that*

$$\begin{aligned} \mathbb{E}(BV) &= \frac{\sigma^2}{M-1} \sum_{m=2}^{M} s_m s_{m-1} \qquad (12) \\ &= \sigma^2 \left(1 - \frac{1}{M-1}\Big(\frac{1}{2}\frac{\int_0^1 g'(x)dx}{\int_0^1 g(x)dx} + \frac{g(0)}{\int_0^1 g(x)dx}\Big) \cdot (1 + o(1))\right), \\ \mathbb{E}(\min RV) &= \frac{\sigma^2}{(\pi-2)(M-1)} \sum_{m=2}^{M} \left(\pi s_m^2 - 2 s_{m-1} s_m + 2\arctan\left(s_m/s_{m-1}\right)\left(s_{m-1}^2 - s_m^2\right)\right) \qquad (13) \\ &= \sigma^2 \left(1 + \frac{1}{M} - \frac{g(0)}{M\int_0^1 g(x)dx}\right) - \frac{\sigma^2}{2M}\frac{\int_0^1 g'(x)dx}{\int_0^1 g(x)dx} + O\left(\frac{1}{M^2}\right). \end{aligned}$$

Consequently, if $M \to \infty$ it holds that $\mathbb{E}(BV) \to \sigma^2$ and $\mathbb{E}(\min RV) \to \sigma^2$.

The results for RV and BV are derived by Dette et al. (2016) in their Proposition 1, whereas the proof for the expectation of minRV in (13) is provided in the Appendix. Based on these results we obtain the closed form expressions for the IP-correction factors for BV and minRV as

$$C(\boldsymbol{s}) = IV/\mathbb{E}(\widehat{IV}),$$

e.g., for the BV measure it is

$$C(\boldsymbol{s}) = M\left(\sum_{m=2}^{M} s_m s_{m-1}\right)^{-1}.$$

The focus of our paper is to quantify IP bias in realized measures of IV for a finite number of intraday returns M. Dette et al. (2016) show that for $M \to \infty$ the IP bias of the BV measure converges to zero. As we show in the proof of Proposition 1, the same holds for minRV. In the simulation study of this paper, we demonstrate that this seems also to be the case for the other estimators where no analytical expression for the IP-bias is available. The values of M where the bias becomes negligible, however, depend on the particular type of the realized estimator and on the form of the IP.

3.1 Estimation of IP

In practice the IP components $\boldsymbol{s} = (s_1, ..., s_M)'$ are unknown and should be estimated from the data. We denote the vector of IP estimates by $\widehat{\boldsymbol{s}} = (\hat{s}_1, ..., \hat{s}_M)'$ and briefly discuss the issue of IP estimation here. Basically, there are both parametric and non-parametric IP estimators. A non-parametric estimator based on the inter-day standard deviation of returns was suggested by Taylor and Xu (1997), however, this estimator appears to be not robust with respect to jumps. Robust non-parametric procedures for IP estimation are discussed by Boudt et al. (2011), in particular, they suggest the weighted standard deviation (WSD) estimator which offers a favorable trade-off between efficiency and robustness and is therefore the preferred estimator in our study. A parametric approach for IP estimation is proposed by Andersen and Bollerslev (1997), who use a flexible Fourier transform to obtain the IP estimates. The latter approach is more efficient than the non-parametric alternatives but is not robust with respect to jumps.

For any consistent estimators of $\boldsymbol{s}$ it holds that $C(\widehat{\boldsymbol{s}}) \xrightarrow{p} C(\boldsymbol{s})$ for the estimation window $T \to \infty$; this result follows from the continuous mapping theorem (see e.g. Hamilton, 1994, p. 482-483) as long as the function $C(\cdot)$ is sufficiently smooth. Additionally, we assume the stochastic independence of $C(\widehat{\boldsymbol{s}})$ which is estimated from historical data, and of $\widehat{IV}$ computed for the current day, as it is also commonly done in the literature (cf. Boudt et al., 2011, Bekierman and Gribisch, 2020).

An important question for IP-estimation is the length T of the estimation window to use, e.g. one year or ten years of daily data. There is an ongoing literature discussion whether IP remains constant over time, see e.g. Andersen et al. (2001), Gabrys et al. (2013); the evidence in Figure 1 in our empirical study supports the conjecture that IP could be time-varying. Hence, smaller estimation windows based on more recent data might be preferable from this perspective. We investigate the effect of the estimation window length T in our simulation study in Section 4.

3.2 Correction factors for IP bias

The analytical results as in Proposition 1 indicate that one could expect a bias due to IP in the popular IV estimators discussed above with the exception of RV measure. Our strategy for IP bias correction is based on the representation in (10) for the given IP profile $\boldsymbol{s}$ and various estimators $\widehat{IV}$. Then given the basic setting $IV = \sigma^2 = 1$, we compute the scalar correction factors $C(\boldsymbol{s})$ such that $C(\boldsymbol{s}) \cdot \mathbb{E}(\widehat{IV}) = 1$ or, alternatively, $C(\boldsymbol{s})^{-1} = \mathbb{E}(\widehat{IV})$.

For the estimated IP profile $\widehat{\boldsymbol{s}}$ we obtain the explicit expressions for the correction factors for BV- and $\min RV$-estimators, which are given by

$$C_{BV}(\widehat{\boldsymbol{s}}) = \left(\frac{1}{M}\sum_{m=2}^{M} \hat{s}_m \hat{s}_{m-1}\right)^{-1}, \tag{14}$$

$$C_{\min RV}(\widehat{\boldsymbol{s}}) = \left[\frac{1}{(\pi-2)(M-1)}\sum_{m=2}^{M}\left(\pi \hat{s}_m^2 - 2\hat{s}_{m-1}\hat{s}_m + 2\arctan\left(\frac{\hat{s}_m}{\hat{s}_{m-1}}\right)(\hat{s}_{m-1}^2 - \hat{s}_m^2)\right)\right]^{-1}. \tag{15}$$

Unfortunately, the expectations $\mathbb{E}(\widehat{IV})$ in case of IP are not known for the other realized estimators of IV, so that there is no analytical expressions available for their correction factors. In particular, finding a closed form expression is hardly possible for both realized kernel RK and pre-averaging estimators, as it is technically challenging to take into account either different kernels or weighting functions. For these reasons we compute the bias correction factors $C(\boldsymbol{s})$ within a Monte Carlo simulation study as it is outlined in Section 4 by a numerical approximation of $\mathbb{E}(\widehat{IV})$ for the given form of IP.

3.3 Using IP-filtered intraday returns

An alternative approach to IP bias correction is to use the estimated profile $\hat{\boldsymbol{s}}$ in order to remove immediately the IP components from the intraday returns, so that IP-filtered returns are given by $r_m^* = r_m/\hat{s}_m$, $m = 1, ..., M$. The filtered returns are then plugged into the IV estimators discussed above instead of the original returns. Throughout this paper we denote by $\widehat{IV}^*$ an estimator of IV which is calculated based on IP-filtered intraday returns. This rescaling method is considered by Andersen and Bollerslev (1997) and Boudt et al. (2011), but has not been analysed up to now in the context of IP bias correction. Here we discuss the rescaling approach in more detail.

First, due to Jensen's inequality, it holds for each squared return rescaled by the estimate $\hat{s}_m^2$ that

$$\mathbb{E}\left((r_m^*)^2\right) = \mathbb{E}\left(\frac{r_m^2}{\hat{s}_m^2}\right) \geq \frac{\mathbb{E}(r_m^2)}{\mathbb{E}(\hat{s}_m^2)} = \frac{\sigma^2}{M},$$

so that, e.g. for the RV^* estimator of IV, it holds that $\mathbb{E}(RV^*) \geq IV$. Hence, the rescaling of returns would introduce some upward bias in case of the imprecisely estimated IP-profile component s_m. Remarkable, this upward bias would be present even in the IP-unbiased measure RV which is rather undesired from the empirical perspective. We quantify this filtering bias in our simulation study in Section 4.

Another important point is related to the fact that the estimation of the scalar correction factor $C(\boldsymbol{s})$ could be done more precisely than estimation of single profile components s_m, $m = 1, ..., M$. As we can observe in Proposition 1, our bias correction factors are based on averages with $(1/M)$, so that the central limit theorem arguments would apply in the case $M \to \infty$. On the other hand, by the estimation of the components of the IP profile $\boldsymbol{s} = (s_1, ..., s_M)'$ which are required for filtering as in Boudt et al. (2011), one could only rely on the asymptotics with respect to the number of days T used for the profile estimation. However, this could be a problem when the IP profile changes over time and the estimation can only be performed from a small sample, see e.g. Figure 1 in our empirical study.

Finally, note that immediate rescaling of returns is much more sensitive to (unexpected) changes in IP compared to the correction proposed in this paper, as it is shown by Dette et al. (2016) in a similar setup for BV measure. In particular, for the days of (unexpected) announcements the IP profile could be very different from those IP estimates which one confronts in 'ordinary' days. On the contrary, as long as the U-shape (or J-shape) is there, the IP-correction factors $C(\boldsymbol{s})$ based on averages would be numerically much less affected than single IP components s_m. This robustness property is another argument for using our approach for IP-bias corrections.

4 Monte Carlo simulation study

The aims of the Monte Carlo simulation study is to quantify biases induced by IP on selected realized volatility measures as well as to evaluate the performance of different approaches for IP bias correction. For this purpose we isolate the effect of IP on the realized measures by the design of our computational procedure. We fix daily IV at $\sigma_t^2 = 1$. The IP components s_m for $m = 1, ..., M$ are given by

$$s_m^2 = f_m / \sum_{m=1}^{M} f_m, \qquad \text{with} \qquad f_m = \left[c_1 + c_2(m - M/2)^2\right]^2, \qquad c_1, c_2 > 0, \tag{16}$$

in order to mimic the empirically observed U-shaped IP-pattern. The parameter c_1 modulates the amount of curvature, with the most IP curvature for $c_1 \to 0^+$, whereas the value $c_1 = 1$ corresponds to the case of no IP at all. We choose $c_1 \in \{0.3, 0.5, 1\}$, whereas the value of c_2 is obtained from the equality $\sum_{m=1}^{M} s_m^2 = M$ for each given c_1.

We generate intraday returns as $r_m \sim \mathcal{N}(0, s_m^2/M)$ for T days, and estimate the IP profile $\boldsymbol{s} = (s_1, ..., s_M)'$ using the robust WSD estimator of Boudt et al. (2011) with the estimates denoted as $\widehat{\boldsymbol{s}}_T$. We do not allow for intraday stochastic volatility due to the evidence that its impact on realized measures is negligible when the stochastic volatility process is sufficiently persistent, see for example Andersen et al. (2014), Dette et al. (2016).

In order to calculate the bias correction factors, we generate intraday returns for additional B days and compute estimates $\widehat{IV}_t$ of realizations of the RV, BV, RK, medRV, minRV, $paRV$ and $paBV$ measures for each day $t = 1, ..., B$. For the realized kernel, we use the Parzen kernel and the bandwidth selection procedure as described in Barndorff-Nielsen et al. (2009). For the pre-averaging measures $paRV$ and $paBV$ we set $g(x) = \min(x, 1 - x)$. Then we obtain the Monte Carlo correction factor for each particular realized estimator as

$$\widehat{C}(\widehat{\boldsymbol{s}}_T) = \left((1/B) \sum_{t=1}^{B} \widehat{IV}_t \right)^{-1}.$$

We investigate the choice of 15-, 10-, 5- and 1-minute returns corresponding to $M \in \{26, 39, 78, 390\}$ intraday returns for an 6.5-hours trading day. Although tick price observations are often available, using 10- or 5-minute returns still remains a popular choice in many applications, especially when risky assets under consideration are not very liquid (cf. Bekierman and Gribisch, 2020, Buccheri and Corsi, 2020).

4.1 IP bias in IV measures

Now we quantify the impact of IP bias on realized volatility measures. We estimate IP based on $T = 10^4$ days and repeat the procedure $B = 10^4$ times in order to get a very precise simulated estimates of the IP bias, so that we can neglect estimation error in IP-profile $\boldsymbol{s}$. Taking the reciprocal of this bias leads to the IP-bias correction factors.

In Table 1, Block A, we show the relative bias in % of all considered estimators for different M and c_1 values. As expected, the RV measure is not biased in the presence of IP, whereas IP bias in RK is negligible. For all other measures we observe a downward IP bias, which is rather undesired as then IV would be underestimated in risk management applications. In line with the theory, we experience that the bias decreases when the curvature of IP is less pronounced. Also, given a fixed curvature parameter c_1, the bias decreases as M increases. The measure BV has the lowest bias which almost disappears for $M = 390$. The bias of minRV is slightly higher but comparable to BV whereas the bias of medRV is substantially larger. This suggests that there might be a trade-off between robustness to jumps and IP robustness, as the medRV estimator is much less affected by jumps, see e.g. Andersen et al. (2012). Mostly biased are the pre-averaged estimators $paRV$ and $paBV$ where even for $M = 390$ and $c_1 = 0.5$ the downward IP-biases are about 10% and 20%, respectively.

To assess the impact of IP on the estimation efficiency, we report the mean squared error (MSE) of all estimators in Table 1, Block B, where we observe that more pronounced IP leads to MSE increases for all estimators. For a better visual assessment, we also report the ratios of MSEs with $c_1 \in \{0.3, 0.5\}$ and without IP with $c_1 = 1$ in Table 1, Block C. For RV, the ratios are the highest, which means that the efficiency of RV suffers mostly in presence of IP. For other estimators the ratios are smaller, however, they increase as M increases. Hence, the IP effect on the MSE does not disappear as M gets larger, which is consistent with the theoretical findings of Dette et al. (2016) who show that a stronger IP inflates the integrated quarticity IQ. Therefore, the impact of IP should be carefully taken into account when conducting statistical inference about IV. However, as in our paper we focus on the bias in the point estimators of IV, a detailed investigation of the IP impact on confidence intervals is left for future research.

4.2 IP-correction factors

Next we consider IP estimates $\widehat{\boldsymbol{s}}_T$ instead of the true IP profile $\boldsymbol{s}$ in order to investigate the feasibility of our Monte Carlo simulation procedures for IP bias corrections and to compare it to the approach of Boudt et al. (2011) which is based on IP-filtered intraday returns. We use the WSD estimator of Boudt et al. (2011) with the estimation windows T of 63, 250, and 1000 days for the IP components, which corresponds to roughly to one quarter, one year and four years of daily data.

First we provide the Monte Carlo simulated correction factors in Table 2. As the benchmark, we use the reciprocal of the exact factors from Table 1 which are placed in the lines with $T = 10^4$. For each choice of the IP curvature parameter c_1 and number of intraday returns M, we compute correction factors for three different estimation windows T and report the averages of the estimated correction factors over $B = 10^4$ simulation runs. In general, we can observe a rather quick convergence to the true correction factor with increasing T, as already for $T = 250$ we can observe a small bias usually in the range much less than 1%, whereby the bias mostly disappears for $T = 1000$. However, the choice $T = 63$ leads to more substantial differences with the 'true' factors computed based on $T = 10^4$. As the correction factors for $paRV$ and $paBV$ are still larger than one even for $M = 390$, we provide additional simulation results for $M \in \{2340, 23400\}$ corresponding to 10-second and 1-second returns in Table 3. We can see that even for 1-second returns, the correction factors for $paBV$ are still larger than one, which suggests that IP correction should also be conducted for these measures when using a very high sampling frequency.

Then we consider BV and $\min RV$ measures and compare in Table 4 their average correction factors obtained from the simulation based method with those resulting from using the analytical expressions in Eqns. (14) and (15) from Proposition 1. First we note that the analytical factors coincide with the Monte Carlo simulated factors for $T = 10^4$ from Table 2. Second, we also observe a slight upward bias even for an estimation window of $T = 250$ days which almost disappears for $T = 1000$. These findings suggest that given the IP estimate $\widehat{\boldsymbol{s}}$, one could use the computational approach in cases when an explicit expression for the IP bias is not readily available for a particular IV estimator.

Finally, in Table 5 we report the relative IP biases of the IV estimators when filtered returns r^*_m are used as in Boudt et al. (2011). First we note that the relative bias decreases as the estimation window length T for $\widehat{\boldsymbol{s}}_T$ increases. For the small estimation windows, we observe some upward biases. Already for $T = 250$ this bias is never larger than 1%, with the exception of $paBV^*$ estimator. As also RV measure which is unaffected by IP exhibits a small upward bias when returns are filtered, we recommend not to apply the filtering for RV estimator. In general, also this method of bias correction works satisfactory for estimation windows about $T = 250$ when the IP profile is correctly specified.

5 Empirical illustration

Now we illustrate our IP bias correction approach using the empirical data from U.S. stock market. We consider log intraday returns of the S&P 500 index and four highly liquid stocks, namely Amazon, American Express Company, Microsoft and Exxon Mobile. We sample returns at 15-, 10- and 5-minute frequencies from January 1998 to December 2012 with 3773 days in total. For each day, we compute the RV, BV, RK, medRV, minRV, $paRV$ and $paBV$ estimators of IV and apply either our IP corrections or the correction based on filtered returns as in Boudt et al. (2011). To implement our IP-corrections approach, we rely on the analytical expressions for BV and minRV and the simulation based procedure for minRV, $paRV$ and $paBV$ to obtain the correction factors.

The IP components are estimated with the WSD estimator of Boudt et al. (2011) with a moving window of $T = 250$ days, which is a reasonable choice according to the results of our Monte Carlo simulation study. For this reason the first 250 days are reserved for computing the initial estimate of the IP profile $\widehat{\boldsymbol{s}}_T(t)$ for day $t = 251$. Based on these estimated profiles we compute for each new day – either analytically or with Monte Carlo simulations – the IP correction factors for different estimators of IV. In Figure 1 we visualize the time evolution of the analytical correction factors for BV and simulated correction factors for medRV for different sampling frequencies.

In Figure 1 one could observe substantial fluctuations in IP correction factors both for BV and medRV measures over the considered period. In general, the estimates become smaller in magnitude as the sampling frequency increases, which is in line with both theoretical and simulation results. The observed patterns are quite similar for all sampling frequencies and both measures. Remarkably, the bias correction factors tend to increase with time, i.e. the issue of IP bias impact has gained in importance recently.

Now we provide a comparison of our IP-bias correction methodology with those of Boudt et al. (2011). To contrast these two different correction approaches empirically, we compute the mean ratios of corrected and uncorrected IV-estimators for all sampling frequencies. We report the mean

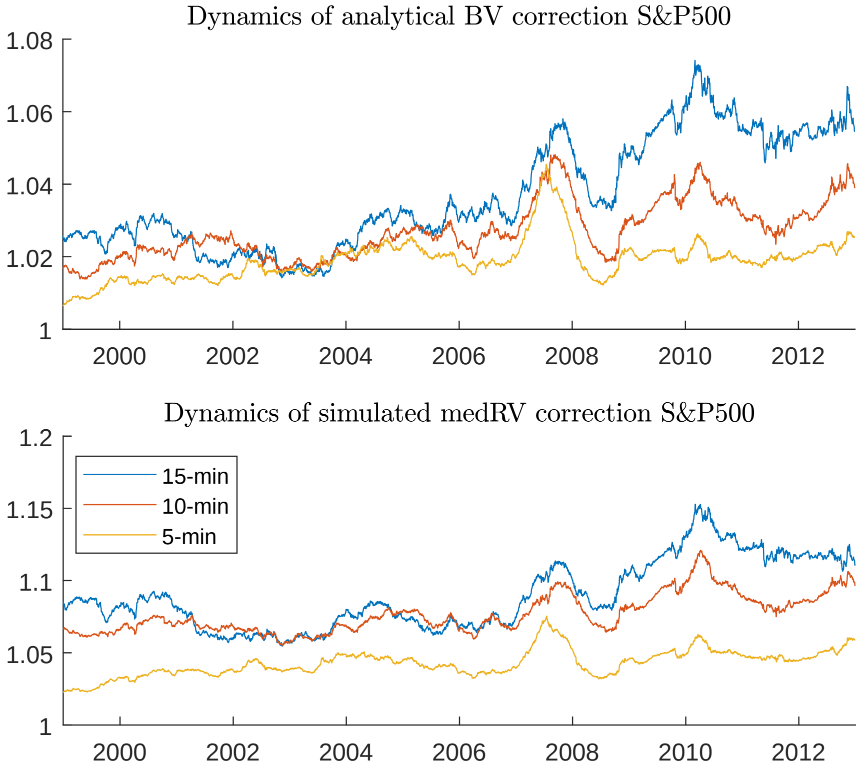

Figure 1: Time evolution of IP-correction factors based on returns sampled at 15-, 10-, 5-minutes: the analytical correction for BV (above) and Monte Carlo simulated correction for medRV (below).

ratios for our approach and for the approach of Boudt et al. (2011) in Tables 6 and 7, respectively. These ratios could be interpreted as a kind of full sample IP biases.

As RV and RK are not corrected in our approach, we refrain from putting ones for them in Table 6. For all corrected measures the average ratio goes down with increasing sampling frequency, in line with the theoretical results of Dette et al. (2016). The bias due to IP seem to the highest for Exxon Mobile and American Express. The estimator BV is the least affected estimator while $paBV$ has the largest average correction factor, similarly to our simulation results. As all correction factors are larger than one, the effect of IP should be accounted for when using these measures in empirical applications.

In Table 7 we provide the results for the approach of Boudt et al. (2011) by reporting the mean ratios for the measures computed from IP-filtered returns with those obtained from non-filtered returns. Here we also report the results for the RV and RK estimators as those measures differ when obtained from filtered or non-filtered data. First we note that the ratios are larger than one for RV and

RK, so the filtering of returns introduces an upward bias – as we discuss in Section 3.3 – of up to 5 %, which decreases with increasing sampling frequency. Therefore all average ratios are all slightly larger than those in Table 6 which could be due to the aforementioned bias. Another explanation could be a certain non-robustness of pre-filtering at some non-typical days, e.g. those with special announcements or unexpected events, as it is illustrated by Dette et al. (2016). Except for this, both methods provide similar results as we also observe in the simulation study. For practical purposes, these results suggest that our approach might be favorable both due to its robustness and because it does not cause an upward bias in measures like RV and RK which do not need IP corrections.

6 Conclusions

For measuring daily integrated volatility IV, there are various realized estimators based on intraday high-frequency information which are proposed in the literature. In this paper we focus on the impact of the intraday periodicity (IP) in absolute intraday returns on several popular realized volatility estimators. In particular, we show both analytically and in Monte Carlo simulations that realized estimators based on adjacent returns exhibit downward bias for finite number of intraday returns M which however disappears for $M \to \infty$. We also propose a scalar factor for IP bias correction for each of the considered realized estimators. Our correction approach is compared to the filtering approach of Boudt et al. (2011) where intraday returns are immediately rescaled by the corresponding IP-components. As our procedure shows certain robustness properties we recommend it for practical purposes of IP bias corrections.

Appendix

For the proof of Proposition 1, we need the following lemma:

Lemma 1. *For independent random variables $X \sim \mathcal{N}(0, \sigma_1^2)$ and $Y \sim \mathcal{N}(0, \sigma_2^2)$ we have*

$$\mathbb{E}[(\min\{|X|, |Y|\})^2] = \frac{1}{\pi} \cdot \left(\pi\sigma_1^2 - 2\sigma_1\sigma_2 + 2\arctan\left(\frac{\sigma_1}{\sigma_2}\right)(\sigma_2^2 - \sigma_1^2) \right). \tag{17}$$

Proof of Lemma 1. Without loss of generality we assume $\sigma_2 = 1$, then we have

$$\mathbb{E}[(\min\{|X|,|Y|\})^2] = A_1 + A_2, \qquad \text{where} \tag{18}$$

$$\begin{aligned} A_1 &= \frac{1}{2\pi\sigma_1} \int_{\mathbb{R}^2} I\{x^2 \le y^2\} x^2 \; e^{-\frac{x^2}{2\sigma_1^2} - \frac{y^2}{2}} dxdy \\ A_2 &= \frac{1}{2\pi\sigma_1} \int_{\mathbb{R}^2} I\{x^2 > y^2\} y^2 \; e^{-\frac{x^2}{2\sigma_1^2} - \frac{y^2}{2}} dxdy. \end{aligned}$$

The integrals are now evaluated separately using "polar coordinates"

$$x = \sigma_1 r \cos\varphi \;, \quad y = r \sin\varphi.$$

This gives for the first term

$$A_1 = \frac{1}{2\pi\sigma_1} \int_{[0,\infty)\times[0,2\pi)} I\{\sigma_1^2 \cos^2\varphi \le \sin^2\varphi\} \sigma_1^3 r^3 \cos^2\varphi e^{-r^2/2} d(r,\varphi)$$

Observing that

$$\big\{\varphi \in [0,2\pi) \;\big|\; \sigma_1 |\cos\varphi| \le |\sin\varphi|\big\} = [\arctan\sigma_1, \pi - \arctan\sigma_1] \cup [\pi + \arctan\sigma_1, 2\pi - \arctan\sigma_1]$$

yields

$$\begin{aligned} A_1 &= \frac{1}{2\pi\sigma_1} \int_0^\infty \int_{\arctan\sigma_1}^{\pi-\arctan\sigma_1} \sigma_1^3 r^3 \cos^2\varphi \; e^{-r^2/2} dr d\varphi + \frac{1}{2\pi\sigma_1} \int_0^\infty \int_{\pi+\arctan\sigma_1}^{2\pi-\arctan\sigma_1} \sigma_1^3 r^3 \cos^2\varphi e^{-r^2/2} dr d\varphi \\ &= \sigma_1^2 \Big(1 - \frac{2\sigma_1}{\pi(1+\sigma_1^2)} - 2\frac{\arctan\sigma_1}{\pi}\Big). \end{aligned}$$

Now a similar argument for the term A_2 gives

$$A_2 = \frac{2}{\pi}\Big(- \frac{\sigma_1}{1+\sigma_1^2} + \arctan\sigma_1 \Big) \;,$$

and it follows from (18)

$$\mathbb{E}[(\min\{|X|,|Y|\})^2 = \frac{\pi\sigma_1^2 - 2\sigma_1 + 2\arctan\sigma_1 \cdot (1-\sigma_1^2)}{\pi} \;,$$

which is the assertion (17) in the case $\sigma_2 = 1$. □

Proof of Proposition 1.

As $r_m \sim \mathcal{N}(0, \sigma^2 \frac{s_m^2}{M})$ and the random variables $r_1, \dots, r_m$ are independent, a direct application of Lemma 1 yields

$$\mathbb{E}[(\min\{|r_m|, |r_{m-1}|\})^2 = \frac{\sigma^2}{M} \left\{ \frac{\pi s_m^2 - 2s_{m-1}s_m + 2\arctan\left(\frac{s_m}{s_{m-1}}\right)(s_{m-1}^2 - s_m^2)}{\pi} \right\}.$$

Using the definition of minRV, we then immediately obtain

$$\mathbb{E}[\min RV] = \frac{\sigma^2}{\pi - 2} \frac{1}{M-1} \sum_{m=2}^{M} \left\{ \pi s_m^2 - 2s_m s_{m-1} + 2\arctan\left(\frac{s_m}{s_{m-1}}\right) \cdot (s_{m-1}^2 - s_m^2) \right\}.$$

To show that minRV is asymptotically unbiased, we first rewrite $\mathbb{E}(\min RV)$ as

$$\begin{aligned}
&\mathbb{E}(\min RV) \\
&= \frac{\sigma^2}{\pi - 2} \frac{1}{M-1} \sum_{m=2}^{M} \left\{ \pi s_m^2 - 2s_{m-1}s_m + 2\arctan\left(\frac{s_m}{s_{m-1}}\right)(s_{m-1}^2 - s_m^2) \right\} \\
&= \frac{\sigma^2}{M-1} \sum_{m=2}^{M} s_m^2 + \frac{\sigma^2}{\pi - 2} \frac{2}{M-1} \sum_{m=2}^{M} (s_m - s_{m-1}) \left\{ s_m - \arctan\left(\frac{s_m}{s_{m-1}}\right)(s_m + s_{m-1}) \right\} \\
&= \frac{\sigma^2 M}{M-1} - \frac{\sigma^2}{M-1} s_1^2 + \frac{2\sigma^2}{(\pi - 2)(M-1)} \sum_{m=2}^{M} (s_m - s_{m-1}) \left\{ s_m - \arctan\left(\frac{s_m}{s_{m-1}}\right)(s_m + s_{m-1}) \right\} \\
&= \sigma^2 \left(1 + \frac{1 - s_1^2}{M-1}\right) + \frac{\sigma^2}{\pi - 2} \frac{2}{M-1} \sum_{m=2}^{M} (s_m - s_{m-1}) \left\{ s_m - \arctan\left(\frac{s_m}{s_{m-1}}\right)(s_m + s_{m-1}) \right\}.
\end{aligned}$$

Now we use the notation

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

where g is a continuously differentiable function per assumption. Then with $s_m = \sqrt{g(\frac{m}{M}) \,/\, g_M}$ and $g_M = \int_0^1 g(x)dx + O(1/M)$ we get

$$\begin{aligned}
&(s_m - s_{m-1}) \left\{ s_m - \arctan\left(\frac{s_m}{s_{m-1}}\right)(s_m + s_{m-1}) \right\} \\
&\quad = \frac{1}{\int_0^1 g(x)dx} \frac{g'(\frac{m}{M})}{2\sqrt{g(\frac{m}{M})}} \frac{1}{M} \left\{ \sqrt{g(\tfrac{m}{M})} - 2\sqrt{g(\tfrac{m}{M})} \arctan 1 \right\} \left(1 + O\left(\frac{1}{M}\right)\right) \\
&\quad = -\frac{g'(\frac{m}{M})}{4M \int_0^1 g(x)dx} (\pi - 2) \left(1 + O\left(\frac{1}{M}\right)\right)
\end{aligned}$$

uniformly with respect to $m \in \{1, \dots, M\}$, and, consequently,

$$\begin{aligned}\mathbb{E}(\min RV) &= \sigma^2\left(1+\tfrac{1-s_1^2}{M}\right) - \frac{\sigma^2}{\pi-2}\frac{1}{2M^2}\sum_{m=2}^{M}\frac{g'(\frac{m}{M})(\pi-2)}{\int_0^1 g(x)dx} + O\left(\frac{1}{M^2}\right)\\ &= \sigma^2\left(1+\frac{1}{M} - \frac{g(0)}{M\int_0^1 g(x)dx}\right) - \frac{\sigma^2}{2M}\frac{\int_0^1 g'(x)dx}{\int_0^1 g(x)dx} + O\left(\frac{1}{M^2}\right) \overset{M\to\infty}{\longrightarrow} \sigma^2,\end{aligned}$$

so that minRV is an asymptotically unbiased estimator of σ^2. □

Acknowledgements

This work was partly supported by the Collaborative Research Center 'Statistical modeling of non-linear dynamic processes' (SFB823, projects A1,C1) of German Research Foundation (DFG).

References

Andersen, T. and Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, 4:115–158.

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

Andersen, T., Dobrev, D., and Schaumburg, E. (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, 138:125–180.

Andersen, T., Dobrev, D., and Schaumburg, E. (2014). A robust neighborhood truncation approach to estimation of integrated quarticity. *Econometric Theory*, 30:3–59.

Andersen, T., Thyrsgaard, M., and Todorov, V. (2019). Time varying periodicity in intraday volatility. *Journal of American Statistical Association*, 528:1695–1707.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N. (2008). Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. *Econometrica*, 76:1481–1536.

Barndorff-Nielsen, O., Hansen, P., Lunde, A., and Shephard, N. (2009). Realized kernels in practice: Trades and quotes. *Econometrics Journal*, 12:C1–C32.

Barndorff-Nielsen, O. and Shephard, N. (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.

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

Bekierman, J. and Gribisch, B. (2020). A mixed frequency stochastic volatility model for intraday stock market returns. *Journal of Financial Econometrics*, DOI:10.1093/jjfinec/nbz021.

Boudt, K., Croux, C., and Laurent, S. (2011). Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance*, 18:353–367.

Buccheri, G. and Corsi, F. (2020). HARK the SHARK: Realized volatility modeling with measurement errors and nonlinear dependencies. *Journal of Financial Econometrics*, DOI:10.1093/jjfinec/nbz025.

Christensen, K., Hounyo, U., and Podolskij, M. (2018). Is the diurnal pattern sufficient to explain the intraday variation in volatility: A nonparametric assessment. *Journal of Econometrics*, 205:336–362.

Christensen, K., Oomen, R., and Podolskij, M. (2010). Realised quantile-based estimation of the integrated variance. *Journal of Econometrics*, 159:74–98.

Christensen, K., Oomen, R., and Podolskij, M. (2014). Fact or friction: Jumps at ultra high frequency. *Journal of Financial Economics*, 114:576–599.

Dette, H., Golosnoy, V., and Kellermann, J. (2016). The effect of intraday periodicity on realized volatility measures. *SFB 823 Working Paper.*

Dumitru, A.-M., Hizmeri, R., and Izzeldin, M. (2019). Forecasting the realized variance in the presence of intraday periodicity. *Working Paper.*

Gabrys, R., Hörmann, S., and Kokoszka, P. (2013). Monitoring the intraday volatility pattern. *Journal of Time Series Econometrics*, 5:87–116.

Hamilton, J. (1994). *Time Series Analysis.* Princeton University Press, New Jersey.

Jacod, J., Li, Y., Mykland, P., Podolskij, M., and Vetter, M. (2009). Microstructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and Their Applications*, 119(7):2249–2276.

Jacod, J. and Protter, P. (2014). *Discretization of Processes.* Springer, Heidelberg.

Taylor, S. and Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, 4:317–340.

	c_1	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
Block A: relative bias in %								
$M = 26$	0.3	-0.00	-11.59	-0.03	-22.32	-12.45	-36.64	-64.51
	0.5	-0.00	-8.55	-0.09	-16.47	-9.06	-28.13	-46.93
	1	-0.00	-0.00	0.04	-0.00	-0.00	0.01	6.23
$M = 39$	0.3	0.00	-7.84	-0.01	-15.31	-8.23	-32.06	-58.43
	0.5	-0.00	-5.80	0.04	-11.32	-6.04	-24.44	-43.06
	1	-0.00	-0.00	-0.02	-0.00	-0.00	-0.02	3.68
$M = 78$	0.3	0.00	-3.97	0.11	-7.86	-4.08	-24.18	-45.77
	0.5	0.00	-2.94	0.10	-5.82	-3.00	-18.30	-33.97
	1	-0.00	-0.00	0.08	0.00	0.00	-0.03	1.56
$M = 390$	0.3	-0.00	-0.80	0.07	-1.60	-0.81	-13.47	-26.12
	0.5	-0.00	-0.60	0.06	-1.19	-0.60	-10.09	-19.46
	1	0.00	-0.00	0.05	0.00	0.00	-0.00	0.29
Block B: MSE								
	c_1	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
$M = 26$	0.3	0.1730	0.1726	0.4380	0.1777	0.2472	0.3081	0.4615
	0.5	0.1354	0.1420	0.3591	0.1453	0.2054	0.2586	0.3154
	1	0.0769	0.1025	0.2470	0.1192	0.1502	0.2672	0.4101
$M = 39$	0.3	0.1146	0.1246	0.3326	0.1305	0.1804	0.2691	0.3947
	0.5	0.0898	0.1011	0.2707	0.1061	0.1470	0.2244	0.2748
	1	0.0513	0.0679	0.1769	0.0783	0.0993	0.2107	0.3002
$M = 78$	0.3	0.0571	0.0679	0.2510	0.0735	0.0989	0.2005	0.2777
	0.5	0.0448	0.0542	0.2022	0.0590	0.0792	0.1639	0.1996
	1	0.0256	0.0337	0.1288	0.0385	0.0492	0.1322	0.1749
$M = 390$	0.3	0.0114	0.0146	0.1337	0.0164	0.0213	0.1061	0.1364
	0.5	0.0090	0.0115	0.1063	0.0129	0.0168	0.0847	0.1026
	1	0.0051	0.0067	0.0642	0.0076	0.0098	0.0571	0.0708
Block C: Ratio of MSE compared to no IP case								
	c_1	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
$M = 26$	0.3	2.2493	1.6834	1.7729	1.4906	1.6463	1.1532	1.1252
	0.5	1.7600	1.3856	1.4535	1.2183	1.3677	0.9677	0.7690
$M = 39$	0.3	2.2353	1.8361	1.8795	1.6676	1.8168	1.2773	1.3147
	0.5	1.7520	1.4897	1.5299	1.3560	1.4809	1.0649	0.9153
$M = 78$	0.3	2.2267	2.0146	1.9479	1.9070	2.0088	1.5158	1.5880
	0.5	1.7478	1.6097	1.5690	1.5303	1.6078	1.2394	1.1414
$M = 390$	0.3	2.2233	2.1780	2.0840	2.1520	2.1773	1.8571	1.9276
	0.5	1.7458	1.7163	1.6564	1.6983	1.7161	1.4829	1.4494

Table 1: The impact of IP on realized measures of IV.

$M = 26$						
	T	BV	medRV	minRV	$paRV$	$paBV$
$c_1 = 0.3$	63	1.1652	1.3673	1.1918	1.6557	3.0242
	250	1.1339	1.2911	1.1473	1.5752	2.8148
	1000	1.1312	1.2882	1.1429	1.5755	2.8149
	10^4	1.1311	1.2874	1.1421	1.5782	2.8178
$c_1 = 0.5$	63	1.1245	1.2645	1.1449	1.4555	2.0202
	250	1.0960	1.2004	1.1041	1.3897	1.8820
	1000	1.0947	1.1982	1.1013	1.3899	1.8838
	10^4	1.0935	1.1971	1.0996	1.3915	1.8843

$M = 39$						
	T	BV	medRV	minRV	$paRV$	$paBV$
$c_1 = 0.3$	63	1.1142	1.2186	1.1342	1.5015	2.4466
	250	1.0879	1.1850	1.0942	1.4703	2.3997
	1000	1.0855	1.1814	1.0904	1.4709	2.4022
	10^4	1.0851	1.1808	1.0897	1.4719	2.4055
$c_1 = 0.5$	63	1.0890	1.1619	1.1049	1.3509	1.7848
	250	1.0635	1.1310	1.0681	1.3223	1.7554
	1000	1.0617	1.1281	1.0652	1.3229	1.7556
	10^4	1.0615	1.1277	1.0642	1.3235	1.7561

$M = 78$						
	T	BV	medRV	minRV	$paRV$	$paBV$
$c_1 = 0.3$	63	1.0665	1.1124	1.0776	1.3456	1.9199
	250	1.0436	1.0886	1.0465	1.3183	1.8432
	1000	1.0415	1.0858	1.0430	1.3187	1.8434
	10^4	1.0414	1.0853	1.0425	1.3190	1.8441
$c_1 = 0.5$	63	1.0516	1.0853	1.0606	1.2480	1.5629
	250	1.0327	1.0654	1.0352	1.2238	1.5141
	1000	1.0311	1.0629	1.0322	1.2239	1.5142
	10^4	1.0304	1.0618	1.0310	1.2240	1.5144

$M = 390$						
	T	BV	medRV	minRV	$paRV$	$paBV$
$c_1 = 0.3$	63	1.0222	1.0350	1.0294	1.1792	1.3771
	250	1.0102	1.0195	1.0120	1.1549	1.3527
	1000	1.0086	1.0171	1.0091	1.1554	1.3532
	10^4	1.0081	1.0163	1.0081	1.1557	1.3535
$c_1 = 0.5$	63	1.0190	1.0292	1.0258	1.1323	1.2612
	250	1.0084	1.0156	1.0102	1.1120	1.2415
	1000	1.0064	1.0127	1.0067	1.1121	1.2416
	10^4	1.0060	1.0121	1.0060	1.1122	1.2417

Table 2: Mean of simulated correction factors with estimated IP

$M = 2340$			
	T	$paRV$	$paBV$
$c_1 = 0.3$	63	1.0608	1.1265
	250	1.0590	1.1269
	1000	1.0589	1.1259
	10^4	1.0591	1.1260
$c_1 = 0.5$	63	1.0435	1.0887
	250	1.0419	1.0887
	1000	1.0418	1.0881
	10^4	1.0419	1.0883

$M = 23400$			
	T	$paRV$	$paBV$
$c_1 = 0.3$	63	1.0167	1.0355
	250	1.0165	1.0349
	1000	1.0167	1.0353
	10^4	1.0167	1.0354
$c_1 = 0.5$	63	1.0118	1.0256
	250	1.0115	1.0250
	1000	1.0117	1.0254
	10^4	1.0118	1.0255

Table 3: Mean of simulated correction factors with estimated IP

	$M = 26$				$M = 39$		
	T	BV	minRV		T	BV	minRV
$c_1 = 0.3$	63	1.1652	1.1918	$c_1 = 0.3$	63	1.1142	1.1342
	250	1.1339	1.1469		250	1.0875	1.0940
	1000	1.1317	1.1433		1000	1.0857	1.0908
	analyt.	1.1311	1.1421		analyt.	1.0851	1.0897
$c_1 = 0.5$	63	1.1245	1.1449	$c_1 = 0.5$	63	1.0890	1.1049
	250	1.0959	1.1039		250	1.0640	1.0685
	1000	1.0941	1.1007		1000	1.0621	1.0652
	analyt.	1.0935	1.0996		analyt.	1.0615	1.0642

	$M = 78$				$M = 390$		
	T	BV	minRV		T	BV	minRV
$c_1 = 0.3$	63	1.0665	1.0776	$c_1 = 0.3$	63	1.0222	1.0294
	250	1.0438	1.0466		250	1.0104	1.0122
	1000	1.0419	1.0435		1000	1.0087	1.0091
	analyt.	1.0414	1.0425		analyt.	1.0081	1.0081
$c_1 = 0.5$	63	1.0516	1.0636	$c_1 = 0.5$	63	1.0190	1.0258
	250	1.0327	1.0351		250	1.0083	1.0100
	1000	1.0310	1.0321		1000	1.0066	1.0070
	analyt.	1.0304	1.0310		analyt.	1.0060	1.0060

Table 4: Comparison of simulation correction factors with analytic expressions for estimated IP

	T	RV^*	BV^*	RK^*	medRV^*	minRV^*	$paRV^*$	$paBV^*$
$M = 26$								
$c_1 = 0.3$	63	3.57	2.64	3.61	2.22	1.94	3.68	9.82
	250	0.90	0.66	0.85	0.59	0.49	0.89	7.12
	1000	0.16	0.11	0.19	0.09	0.06	0.19	6.43
$c_1 = 0.5$	63	3.67	2.74	3.64	2.30	2.01	3.72	9.86
	250	0.91	0.67	0.84	0.58	0.50	0.85	7.05
	1000	0.22	0.17	0.19	0.15	0.14	0.17	6.45
$M = 39$								
	T	RV^*	BV^*	RK^*	medRV^*	minRV^*	$paRV^*$	$paBV^*$
$c_1 = 0.3$	63	3.43	2.45	3.48	2.01	1.73	3.46	7.05
	250	0.79	0.58	0.83	0.46	0.42	0.83	4.45
	1000	0.18	0.14	0.23	0.10	0.11	0.24	3.93
$c_1 = 0.5$	63	3.64	2.69	3.61	2.28	1.98	3.63	7.24
	250	1.04	0.81	1.03	0.71	0.63	1.07	4.80
	1000	0.22	0.18	0.21	0.12	0.14	0.19	3.86
$M = 78$								
	T	RV^*	BV^*	RK^*	medRV^*	minRV^*	$paRV^*$	$paBV^*$
$c_1 = 0.3$	63	3.50	2.56	3.61	2.13	1.87	3.50	4.94
	250	0.84	0.62	0.96	0.51	0.44	0.83	2.42
	1000	0.21	0.14	0.28	0.11	0.09	0.18	1.75
$c_1 = 0.5$	63	3.68	2.72	3.81	2.29	2.01	3.70	5.21
	250	0.93	0.70	0.96	0.59	0.53	0.86	2.45
	1000	0.20	0.14	0.35	0.11	0.10	0.23	1.86
$M = 390$								
	T	RV^*	BV^*	RK^*	medRV^*	minRV^*	$paRV^*$	$paBV^*$
$c_1 = 0.3$	63	3.73	2.78	3.79	2.34	2.08	3.73	3.96
	250	0.90	0.67	0.94	0.56	0.49	0.90	1.17
	1000	0.22	0.16	0.30	0.14	0.11	0.24	0.52
$c_1 = 0.5$	63	3.73	2.77	3.78	2.34	2.07	3.74	3.95
	250	0.91	0.68	0.99	0.57	0.50	0.92	1.19
	1000	0.23	0.17	0.29	0.15	0.12	0.24	0.53

Table 5: Relative bias in % of IV estimators calculated from IP filtered returns $r^*_{t,m}$.

S&P 500

Min	BV	medRV	minRV	$paRV$	$paBV$
15	1.0370	1.0897	1.0440	1.1320	1.2615
10	1.0270	1.0767	1.0326	1.1511	1.2635
5	1.0190	1.0428	1.0250	1.0953	1.1864

Amazon

Min	BV	medRV	minRV	$paRV$	$paBV$
15	1.0881	1.1826	1.0980	1.2699	1.4153
10	1.0657	1.1446	1.0725	1.2723	1.4139
5	1.0373	1.0815	1.0412	1.1952	1.3092

American Express

Min	BV	medRV	minRV	$paRV$	$paBV$
15	1.1240	1.2246	1.1435	1.3143	1.4383
10	1.1067	1.1956	1.1234	1.3239	1.4363
5	1.0710	1.1276	1.0818	1.2435	1.3383

Microsoft

Min	BV	medRV	minRV	$paRV$	$paBV$
15	1.0776	1.1601	1.0858	1.2403	1.3554
10	1.0597	1.1284	1.0660	1.2418	1.3492
5	1.0360	1.0740	1.0404	1.1689	1.2616

Exxon Mobile

Min	BV	medRV	minRV	$paRV$	$paBV$
15	1.1078	1.1975	1.1249	1.2747	1.3768
10	1.0939	1.1716	1.1090	1.2798	1.3675
5	1.0617	1.1091	1.0715	1.2040	1.2757

Table 6: Ratio of means of estimators with correction factors and uncorrected estimators.

S&P 500

Min	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
15	1.0244	1.0487	1.0110	1.0830	1.0475	1.1452	1.3010
10	1.0253	1.0414	1.0177	1.0612	1.0372	1.1277	1.2686
5	1.0314	1.0342	1.0235	1.0427	1.0275	1.0976	1.1980

Amazon

Min	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
15	1.0540	1.1135	1.0555	1.1866	1.1117	1.3400	1.5184
10	1.0497	1.0942	1.0445	1.1420	1.0904	1.2925	1.4615
5	1.0416	1.0639	1.0400	1.0877	1.0622	1.1946	1.3291

American Express

Min	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
15	1.0425	1.1518	1.0448	1.2317	1.1643	1.3323	1.5002
10	1.0333	1.1351	1.0336	1.1939	1.1472	1.3069	1.4582
5	1.0258	1.0963	1.0274	1.1370	1.1013	1.2350	1.3481

Microsoft

Min	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
15	1.0324	1.0952	1.0287	1.1508	1.0924	1.2489	1.4023
10	1.0307	1.0743	1.0280	1.1160	1.0733	1.2146	1.3584
5	1.0329	1.0502	1.0323	1.0685	1.0450	1.1540	1.2688

Exxon Mobile

Min	RV	BV	RK	medRV	minRV	$paRV$	$paBV$
15	1.0431	1.1378	1.0035	1.2034	1.1474	1.2730	1.4248
10	1.0323	1.1274	1.0224	1.1710	1.1421	1.2466	1.3772
5	1.0281	1.0897	1.0260	1.1210	1.0945	1.1878	1.2852

Table 7: Ratio of means of estimators computed from IP-filtered and non-filtered returns.