European University Institute

MAX WEBER PROGRAMME FOR POSTDOCTORAL STUDIES

EUI WORKING PAPERS

MWP 2013/21
Max Weber Programme

Quarticity Estimation on ohlc data

Janine Balter

European University Institute
Max Weber Programme

Quarticity Estimation on ohlc data

Janine Balter

EUI Working Paper **MWP** 2013/21

ISSN 1830-7728

Printed in Italy
European University Institute
Badia Fiesolana
I – 50014 San Domenico di Fiesole (FI)
Italy
www.eui.eu
cadmus.eui.eu

Abstract

Integrated quarticity, a measure of the volatility of volatility, plays a key role in analyzing the volatility of financial time series. As it is an important ingredient for the construction of accurate confidence intervals for integrated volatility, its accurate estimation is of high interest. Given that it includes fourth order returns, it is relatively hard to estimate. This article proposes a new, very efficient and jump-robust estimator of integrated quarticity -based on intraday open, high, low and close prices (ohlc data) - and compares its performance to that of the realized quarticity.

Keywords

Volatility, integrated quarticity, high-low prices, high-frequency data, jumps.

JEL-codes: C10, C13, C14, G10

Janine Balter
Max Weber Fellow, 2012-2013
Janine.Balter@EUI.eu

Introduction

In finance, volatility is understood as a measure of the variation in the price of a financial instrument over time. Thus, it is one measure of the risk inherent in that financial instrument. Consequently, volatility plays an indispensable role in asset and derivative pricing, portfolio optimization and in investment decision making (see Busse (1999), Bollerslev and Mikkelsen (1999) or Fleming et al. (2003), among others). As volatility is unobservable, its appropriate estimation is a key issue. The measurement of volatility had a breakthrough with the availability of high-frequency data. Based on intraday returns, Andersen and others (2011) constructed the realized variance, RV (the expression realized volatility is often used synonymously in the literature). Barndorff-Nielsen and Shephard (2002) were able to show that for general semimartingales, RV converges in probability to the quadratic variation of the process when the number of intraday observations N increases (the time between the observations decreases). The quadratic variation is composed of integrated volatility, IV $:= \int_0^1 \sigma_s^2 ds$, and the sum of squared price jumps. Therefore, in the case of no price jumps, RV is a consistent estimator of the continuous part of volatility, IV. Integrated volatility is of high importance in option pricing theory, where the price of an option usually depends on the integrated volatility of the underlying stock price (see, for example, Hull and White (1987) or more recently Corsi et al. (2013)). By means of certain financial instruments, for example variance swaps, it is also possible to speculate on volatility itself. Hence, there is an interest in constructing confidence intervals for integrated volatility. Barndorff-Nielsen and Shephard (2002) showed that in the absence of price jumps, the $\sqrt{\text{N}}$-scaled difference between RV and IV is mixed-normal distributed with a variance depending on the (unfeasible) integrated quarticity IQ $:= \int_0^1 \sigma_s^4 ds$. Consequently, IQ is a critical ingredient when it comes to the construction of confidence intervals for IV. However, the estimation of integrated quarticity is particularly hard as it involves the fourth power of returns, which are even more exposed to measurement errors than squared ones. In addition to the requirement of accurate estimates, the efficiency of the estimator plays another important role: an estimator of IQ which is two times more efficient than an alternative estimator reduces the length of the confidence interval by a factor of $\sqrt{2}$.

Besides this, IQ is also attractive for itself as it describes the fluctuations in volatility. The importance of (knowledge about) the fluctuation in volatility should not be underrated as it constitutes a part of the risk investors are confronted with. See, for example, Carr and Wu (2009), who stated "When investing in a security, an investor faces at least two sources of uncertainty, namely the uncertainty about the return as captured by the return variance, and the uncertainty about the return variance itself." Further, Christensen et al. (2011) noted that unexpected changes in prices can often be attributed to changes in volatility instead of being attributed to the occurrence of price jumps. Consequently, Barndorff-Nielsen and Veraart (2012) established a stochastic volatility model which allows for stochastic volatility of volatility. With this model, heavy fluctuations in volatility can be established.

Recently, special attention has been paid to the estimation and performance of integrated quarticity. The well-known realized quarticity, which is constructed from the fourth power of intraday returns, is a natural estimator of integrated quarticity and is a straightforward extension of the concept of realized variance. Unfortunately, it fails under many realistic scenarios. For example, it is not consistent if the underlying price process exhibits jumps (see Barndorff-Nielsen and Shephard (2002)).

The first consistent estimators of IQ in the case of jumps of finite activity were provided by Barndorff-Nielsen and Shephard (2002). They extended the theory of multipower variation estimators to the fourth power. These estimators are based on the multiplication of a fixed number of adjacent fourth-power intraday returns. The idea is to mitigate the influence of a possible jump by multiplying the absolute value of the affected intraday return by the absolute value of adjacent, unaffected intraday returns. Clearly, this requires the assumption that not too many (adjacent) intervals contain a jump. Generally, lower order multipower-based statistics (by order, we refer to the number of adjacent returns which are multiplied) are more efficient but less jump-robust than those of a higher order. This leads to a trade-off between efficiency and jump robustness. Moreover, Veraart (2010) showed that in volatility estimation, multipower variations of orders up to ten generate a finite sample bias when the underlying process contains price jumps (see also Corsi et al. (2010) and Barndorff-Nielsen et al. (2006)). Regarding multipower-based estimators, we also face a trade-off between locality and jump robustness. The reason is that the estimator's value of the i-th intraday interval includes not only the information contained in the i-th interval but also the information from the adjacent intervals. Hence, the higher the order the less local the multipower-based estimator is. This can especially result in a bias when volatility fluctuates heavily.

Recently, Kolokolov and Renò (2012) analyzed the efficiency of truncated multipower variation estimators. These estimators are an extension of the standard multipower variations (see Corsi et al. (2010)). The idea is to increase the robustness to jumps (under competitive efficiency) by computing the estimators on the basis of intraday returns which are not influenced by jumps. An appropriate threshold for the returns was provided by Mancini (2009). Kolokolov and Renò (2012) figured out that (truncated) multipower variations with equal exponents (like tripower quarticity, where each of the adjacent returns is taken to the power 4/3) are not the most efficient ones within the class of possible estimators. Hence they determine the vectors of exponents which generate estimators of higher efficiency. Additional symmetrization leads to estimators which have a distinctly smaller mean square error compared to commonly used multipower variations.

Andersen et al. (2012) introduced two jump-robust estimators of integrated quarticity, which are also based on adjacent intraday returns. Instead of multiplying, they choose the minimum or median of two and three adjacent returns respectively in order to reduce the influence of jumps. In Andersen et al. (2011), they further suggested to use a certain filtering procedure when it comes to the estimation of integrated powers of volatility, $\int_0^1 \sigma_s^p ds$. This filtering procedure enhances the estimator's robustness to noise and to possible outliers in the data. It takes a fixed number m (a block) of adjacent intraday returns (a block of size smaller than 6 is suggested). Then, an unbiased local estimator of the p-th power variation is computed for each block. This is done by raising a fixed number j<m of returns of each block to p-th power (for example the j largest returns) and to scale them appropriately. Among these candidates of unbiased estimators, the best one is chosen (for example by taking the minimum or median of the estimates). Another appropriate scaling (where the scaling depends on the chosen composition) yields the final unbiased estimator.

Mancino and Sanfelici (2012) provided a framework for the estimation of both spot and integrated quarticity. It is an extension of Mancino and Sanfelici (2008), were estimators for the integrated volatility, based on the calculated Fourier coefficients of the volatility process, are established. By linking the Fourier series of the underlying price process (without jumps) with the one of the volatility process, the zero-th Fourier coefficient of the fourth power of volatility can be calculated. This serves as an estimator of integrated quarticity, which is characterized by not being sensitive to certain microstructure noise effects.

Independent from this work, Jacod and Rosenbaum (2012) provided a framework for the estimation of integrals of functions (of class C^3) of the volatility matrix. The underlying considered process is quite general, being an Itô semimartingale plus possible price jumps of finite activity. The idea is to approximate the integral over the desired function of volatility by Riemann sums and to replace the spot volatility by an appropriate estimate. This approach is not new. The originality lies in the fact that the authors could provide two estimators (one makes use of overlapping intervals to stabilize the estimates in the case of outliers) which allow for an (easily made feasible) unbiased central limit theorem with convergence rate $\sqrt{1/\mathrm{N}}$. Moreover, the authors showed that their estimators are (under certain assumptions on the model) efficient in the sense that the corresponding variances reach the lower bounds of the Hajek convolution theorem (see Clément et al. (2013)).

The present work provides an alternative estimator of integrated quarticity which is based on the ideas on volatility estimation given in Klößner (2009). Its computation is based on the intraday open, high, low and close (ohlc) values of a log price process and thus takes into account the full range of prices. Consequently, we might expect estimators of higher efficiency compared to estimators which exploit just the intraday open and close values. The use of daily ohlc prices dates from Parkinson (1980) and Garman and Klass (1980). In fact, Parkinson (1980) found that volatility estimators based on daily ohlc data are over eight times more efficient than simple squared daily returns. Using intraday ohlc values in order to estimate volatility was the natural next step. Thus, Christensen and Podolskij (2007) used the intraday price range (the difference between the highest and smallest value within each intraday interval) in order to estimate integrated volatility and integrated quarticity (without providing a central limit theorem for the latter). Unfortunately, some problems arise with the use of intraday ohlc data, as the correct determination of the high and low values requires continuous observations. As already pointed out by Garman and Klass (1980), under discrete sampling, the observed high (low) values will be lower (higher) than the ones reached under continuous prices. In simulation studies this problem can be mitigated by choosing the discretization fine enough. Nevertheless, the estimates remain subject to bias in empirical studies, when – in order to account for microstructure noise effects – data are considered at comparably low frequencies. However, the bias-correction factor of our estimator is not influenced by discrete sampling. The reason is that it does not depend on the observations (or number of intraday observations) as it is derived from the common density of the high, low and close values of a standard Brownian motion and thus is independent from N. In our Monte-Carlo study, it turns out that the new ohlc-based estimator performs poorly in the case of daily bounce-backs. However, in the case of stochastic volatility and large daily jumps, it is a highly precise estimator of integrated quarticity and produces narrow confidence intervals for IV. Its robustness to adjacent jumps is interesting in particular, as estimators which are functions of adjacent non-truncated returns are known to react very sensitive to adjacent jumps.

The rest of this article is organized as follows. Section 1 presents the new ohlc-based estimator of integrated quarticity and some alternative estimators provided by the literature. The properties of the ohlc-based estimator and some of its drawbacks are discussed. The performance of the estimator is examined by Monte-Carlo simulations in section 2. Section 3 concludes. The detailed proofs can be found in the Web Appendix.

1 Theory

Let the log price process $\mathrm{p} := (\mathrm{p}_t)_{t\in[0,1]}$ be defined on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,1]}, P)$ so that p is adapted to the filtration $(\mathcal{F}_t)_{t\in[0,1]}$. We assume that p follows an Itô semimartingale described by

$$\mathrm{p}_t = \mathrm{p}_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + \mathrm{J}_t, \tag{A1}$$

where $(\mu_t)_{t\in[0,1]}$ is a drift process, which is locally bounded and predictable, $W := (W_t)_{t\in[0,1]}$ is a standard Brownian motion, σ_t is the never vanishing càdlàg spot volatility and $\mathrm{J} := (\mathrm{J}_t)_{t\in[0,1]}$ is a pure jump process with jumps of finite activity.

In some cases, we will need some further mild structural assumptions on the volatility process:

$$\sigma_t = \sigma_0 + \int_0^t \tilde{\mu}_s ds + \int_0^t \tilde{\sigma}_s dW_s + \int_0^t \tilde{v}_s dB_s, \tag{A2}$$

where $(\tilde{\mu}_t)_{t\in[0,1]}, (\tilde{\sigma}_t)_{t\in[0,1]}, (\tilde{v}_t)_{t\in[0,1]}$ are càdlàg, $(\tilde{\mu}_t)_{t\in[0,1]}$ is locally bounded as well as predictable and $B := (B_t)_{t\in[0,1]}$ is a standard Brownian motion independent of W. Obviously, the Brownian semimartingale $\mathrm{p}_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$ captures the continuous part of the price process, while jumps $\Delta\,\mathrm{p}_t := \mathrm{p}_t - \mathrm{p}_{t-}$ are given by J_t.

We further assume that the intraday open, high, low and close values of p are given on an equidistant grid $\pi := \{0 < \frac{1}{\mathrm{N}} < ... < \frac{\mathrm{N}-1}{\mathrm{N}} < 1\}$ for every trading day:

$$\mathrm{p}_{\frac{i}{\mathrm{N}}} \quad \text{(intraday log prices)}, \qquad (\mathrm{p}^*)_{i,\mathrm{N}} := \sup_{\frac{i-1}{\mathrm{N}} \le t \le \frac{i}{\mathrm{N}}} \mathrm{p}_t \quad \text{(intraday highs)}$$

$$(\mathrm{p}_*)_{i,\mathrm{N}} := \inf_{\frac{i-1}{\mathrm{N}} \le t \le \frac{i}{\mathrm{N}}} \mathrm{p}_t \quad \text{(intraday lows)},$$

where N is the number of intraday observations and $i = 1, ..., \mathrm{N}$.

First estimators of integrated quarticity $\mathrm{IQ} := \int_0^1 \sigma_\tau^4 d\tau$ were provided by Barndorff-Nielsen and Shephard (2002, 2004, 2006). The so-called realized quarticity is a consistent estimator of IQ in the case of no jumps. It is

$$\mathrm{RQ} = \frac{\mathrm{N}}{3} \sum_{i=1}^{\mathrm{N}} (\mathrm{p}_{\frac{i}{\mathrm{N}}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}})^4 \xrightarrow{\mathrm{P}} \mathrm{IQ}$$

for $\mathrm{N} \to \infty$ and P denotes convergence in probability. It is a natural extension of the concept of realized variance ($\mathrm{RV} := \sum_{i=1}^{\mathrm{N}} (\mathrm{p}_{\frac{i}{\mathrm{N}}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}})^2$) which is a consistent estimator of the quadratic variation of a price process. Unfortunately, RQ does not remain consistent when jumps occur.

First estimators of IQ which are asymptotically robust to jumps were developed by Barndorff-Nielsen and Shephard (2004) and Andersen et al. (2007). They are a natural extension of multipower variation to the fourth power and are formally defined as

$$\mathrm{MPV}(\mathrm{m},4) := \mu_{\frac{4}{\mathrm{m}}}^{-\mathrm{m}} \frac{\mathrm{N}^2}{\mathrm{N}-\mathrm{m}+1} \sum_{i=\mathrm{m}}^{\mathrm{N}} |\, \mathrm{p}_{\frac{i}{\mathrm{N}}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}} |^{\frac{4}{\mathrm{m}}} \cdot ... \cdot |\, \mathrm{p}_{\frac{i-(\mathrm{m}-1)}{\mathrm{N}}} - \mathrm{p}_{\frac{i-\mathrm{m}}{\mathrm{N}}} |^{\frac{4}{\mathrm{m}}} \tag{1}$$

with

$$\mu_a = \mathbb{E}(|Z|^a), \quad Z \sim \mathbb{N}(0,1), \quad a > 0 \quad \text{and} \quad \mathrm{m} > 2.$$

$\mathrm{MPV}(\mathrm{m},4)$ converges to IQ in probability for $\mathrm{N} \to \infty$. Some commonly used versions are the realized tripower quarticity $\mathrm{MPV}(3,4)$, the realized quadpower quarticity $\mathrm{MPV}(4,4)$ and the realized quintpower quarticity, $\mathrm{MPV}(5,4)$. Central limit theorems for estimators of the form (1) are given under the jump alternative for $\mathrm{m} > 4$. It holds that

$$\sqrt{\mathrm{N}}\,(\mathrm{MPV}(\mathrm{m},4) - \mathrm{IQ}) \xrightarrow{d_S} \mathbb{MN}(0, \mu_{\frac{4}{\mathrm{m}}}^{-\mathrm{m}} \int_0^1 \sigma_s^8 ds), \tag{2}$$

where $\mu_{\frac{4}{\mathrm{m}}}^{-\mathrm{m}}$ is the efficiency factor, d_S denotes stable convergence in law and $\mathbb{MN}$ stands for the mixed normal distribution.

A technique to increase the robustness of multipower variation to jumps was suggested by Corsi et al. (2010). They truncate the intraday returns according to a certain threshold function θ_t. The necessary requirements on θ_t were introduced by Mancini (2009). Precisely, they estimate integrated quarticity by

$$\begin{aligned}\mathrm{TMPV}(\mathbf{r},4) := \mathrm{c}_{\mathbf{r},4} \frac{\mathrm{N}^2}{\mathrm{N}-(\mathrm{m}-1)-\mathrm{N_J}} \sum_{i=\mathrm{m}}^{\mathrm{N}} |\, \mathrm{p}_{\frac{i}{\mathrm{N}}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}} |^{\frac{4}{\mathrm{r}_1}} \cdot ... \cdot |\, \mathrm{p}_{\frac{i-(\mathrm{m}-1)}{\mathrm{N}}} - \mathrm{p}_{\frac{i-\mathrm{m}}{\mathrm{N}}} |^{\frac{4}{\mathrm{r_m}}} \cdot \\ \mathrm{I}_{\{|\mathrm{p}_{\frac{i}{\mathrm{N}}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}}| \le \theta_{\frac{i-1}{\mathrm{N}}}\}} \cdot ... \cdot \mathrm{I}_{\{|\mathrm{p}_{\frac{i-(\mathrm{m}-1)}{\mathrm{N}}} - \mathrm{p}_{\frac{i-\mathrm{m}}{\mathrm{N}}}| \le \theta_{\frac{i-\mathrm{m}}{\mathrm{N}}}\}},\end{aligned} \tag{3}$$

where $\mathrm{N_J}$ is the number of intraday returns which are truncated (vanish) due to the indicator functions, the vector $\mathbf{r} := (\mathrm{r}_1, ..., \mathrm{r_m})'$ has positive, real entries with $\sum_{i=1}^{\mathrm{m}} \mathrm{r}_i = 4$ and $\mathrm{c}_{\mathbf{r},4} := (\mu_{\mathrm{r}_1} \cdot ... \cdot \mu_{\mathrm{r_m}})^{-1}$ is a constant which makes TMPV unbiased. In Corsi et al. (2010) it is shown that under the two conditions $\theta_t \to 0$ and $\mathrm{N}^{-1} \log(\mathrm{N})/\theta_t \to 0$ (for $\mathrm{N} \to \infty$)

$$\sqrt{\mathrm{N}} \left(\mathrm{TMPV}(\mathbf{r},4) - \int_0^1 \sigma_s^4 ds \right) \xrightarrow{d_S} \mathbb{MN}(0, V_{\mathbf{r}} \int_0^1 \sigma_s^8 ds)$$

with a certain efficiency factor $V_{\mathbf{r}}$ (see Corsi et al. (2010), formula (2.13)). In Kolokolov and Renò (2012), it is proved that the common practice of choosing equal exponents, meaning that $\mathrm{r}_1 = ... = \mathrm{r_m}$, does not result in the most efficient estimator (the one with the smallest possible $V_{\mathbf{r}}$) within the class of estimators of the form (3), but the scheme

$\mathbf{r}^* = (3.5455, 0.2182, 0.2362)'$,[1] which yields $V_{\mathbf{r}} = 9.70$ (instead of, for example, 13.65, which is the efficiency factor of MPV(3,4)). Consequently, Kolokolov and Renò (2012) suggest using the following symmetrized estimator

$$\mathrm{STMPV} := \frac{\mathrm{TMPV}\left(\mathbf{r}^*, 4\right) + \mathrm{TMPV}\left(\mathbf{sr}^*, 4\right)}{2}$$

with $\mathbf{sr}^* := (0.2362, 0.2182, 3.5455)'$. Choosing $\theta_t = 5\sigma_t^{\mathrm{N}}$ as the threshold function and estimating the intraday volatility σ_t^{N} as in Corsi et al. (2010), the gained mean square error is up to 30% lower when using STMPV instead of the corresponding estimator with $\mathrm{r}_1 = \ldots = \mathrm{r}_\mathrm{m}$.

Another method was suggested by Andersen et al. (2011). Their estimators of IQ are a natural extension of the estimators of integrated volatility proposed in Andersen et al. (2012):

$$\begin{aligned}
\mathrm{MinRQ} &= \mathrm{N}\,\mu_{\min}^{-1}\frac{\mathrm{N}}{\mathrm{N}-1}\sum_{i=2}^{\mathrm{N}}\min(|\,\mathrm{p}_{\frac{i}{\mathrm{N}}}-\mathrm{p}_{\frac{i-1}{\mathrm{N}}}\,|,|\,\mathrm{p}_{\frac{i-1}{\mathrm{N}}}-\mathrm{p}_{\frac{i-2}{\mathrm{N}}}\,|)^4,\\
\mathrm{MedRQ} &= \mathrm{N}\,\mu_{\mathrm{med}}^{-1}\frac{\mathrm{N}}{\mathrm{N}-2}\sum_{i=3}^{\mathrm{N}}\mathrm{med}(|\,\mathrm{p}_{\frac{i}{\mathrm{N}}}-\mathrm{p}_{\frac{i-1}{\mathrm{N}}}\,|,|\,\mathrm{p}_{\frac{i-1}{\mathrm{N}}}-\mathrm{p}_{\frac{i-2}{\mathrm{N}}}\,|,|\,\mathrm{p}_{\frac{i-2}{\mathrm{N}}}-\mathrm{p}_{\frac{i-3}{\mathrm{N}}}\,|)^4
\end{aligned}$$

with $\mu_{\min} = \mathbb{E}\left[\min(|Z_1|^4, |Z_2|^4)\right]$, $\mu_{\mathrm{med}} = \mathbb{E}\left[\mathrm{med}(|Z_1|^4, |Z_2|^4, |Z_3|^4)\right]$, $Z_j \sim \mathbb{N}(0,1)$ and Z_j independent from each other for $j \in \{1,2,3\}$. Both estimators are consistent for IQ in the presence of jumps of finite activity and allow for a central limit theorem:

$$\begin{aligned}
\sqrt{\mathrm{N}}\ (\mathrm{MinRQ} - \mathrm{IQ}) &\overset{d_S}{\to} \mathbb{MN}(0,\ 18.54\int_0^{\mathrm{T}}\sigma_s^8 ds),\\
\sqrt{\mathrm{N}}\ (\mathrm{MedRQ} - \mathrm{IQ}) &\overset{d_S}{\to} \mathbb{MN}(0,\ 14.16\int_0^{\mathrm{T}}\sigma_s^8 ds).
\end{aligned}$$

MinRQ and MedRQ are more local than MPV(m, 4) with m > 2 and m > 3, respectively, as their calculation include fewer adjacent returns.

Now we will introduce a new estimator of IQ, which is based on the intraday open, high, low and close (ohlc) values of a log price process. Thus it incorporates the full range of prices. At the beginning, we consider six estimators of the same structure. Later, we will combine these in an optimal way in order to obtain the estimator of minimal (under certain constraints) variance. The considered six estimators are formally given by

$$\mathrm{N}\sum_{i=1}^{\mathrm{N}}\mathrm{f}\left((\mathrm{p}^*)_{i,\mathrm{N}}-\mathrm{p}_{\frac{i-1}{\mathrm{N}}},(\mathrm{p}_*)_{i,\mathrm{N}}-\mathrm{p}_{\frac{i-1}{\mathrm{N}}},\mathrm{p}_{\frac{i}{\mathrm{N}}}-\mathrm{p}_{\frac{i-1}{\mathrm{N}}}\right), \tag{4}$$

[1] At least for m=3. The consideration of more than three adjacent returns leads to even higher efficiency but comes along with a loss of locality and a higher bias in the presence of jumps due to unequal weights. However, this effect is mitigated by the threshold.

where f is some appropriate homogeneous function of order four in its three arguments $x_{i,\mathrm{N}} := (\mathrm{p}^*)_{i,\mathrm{N}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}}$, $y_{i,\mathrm{N}} := (\mathrm{p}_*)_{i,\mathrm{N}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}}$ and $z_{i,\mathrm{N}} := \mathrm{p}_{\frac{i}{\mathrm{N}}} - \mathrm{p}_{\frac{i-1}{\mathrm{N}}}$.

These appropriate functions f combine the three arguments in a way that makes the estimators robust to jumps. Precisely, we exploit the specific reaction of the so-called stick lengths of the intraday ohlc value to jumps. The stick lengths are defined by $-y_{i,\mathrm{N}}$ and $x_{i,\mathrm{N}} - z_{i,\mathrm{N}} = (\mathrm{p}^*)_{i,\mathrm{N}} - \mathrm{p}_{\frac{i}{\mathrm{N}}}$ in the case that the i-th intraday return is positive (the closing price of the interval $[\frac{i-1}{\mathrm{N}}, \frac{i}{\mathrm{N}}]$ is higher than its opening price). Conversely, in the case that the i-th intraday return is negative (the closing price of the interval is lower than its opening price), the stick lengths are given by $x_{i,\mathrm{N}}$ and $z_{i,\mathrm{N}} - y_{i,\mathrm{N}} = \mathrm{p}_{\frac{i}{\mathrm{N}}} - (\mathrm{p}_*)_{i,\mathrm{N}}$. Let us illustrate the reaction of the stick lengths to negative and positive jumps. Suppose that a negative jump occurred in the i-th intraday interval. The corresponding return will be negative (as the jump should dominate the price movement) and the stick length $z_{i,\mathrm{N}} - y_{i,\mathrm{N}}$ will be comparably small since the continuous part of the price movement cannot revert the downward directed movement generated by the negative jump. In contrast, the second stick length, $x_{i,\mathrm{N}}$, will not be affected at all. Similarly, in the case of a positive jump, the difference between the highest and the closing price $(x_{i,\mathrm{N}} - z_{i,\mathrm{N}})$ will be small and the second stick length, $-y_{i,\mathrm{N}}$, will not be affected. Of course, this consideration requires that the jump is large enough to dominate the intraday return.

Based on these thoughts we will consider functions f, which account for positive and negative intraday increments separately and which combine the stick lengths in a way that f is not affected by jumps (asymptotically). Concretely, we consider the following six functions

$$\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_1}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) := d_1 \left((x_{i,\mathrm{N}} - z_{i,\mathrm{N}})^4 + y_{i,\mathrm{N}}^4\right) \mathrm{I}_{\{z_{i,\mathrm{N}}>0\}},$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_2}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) := d_2 \left((x_{i,\mathrm{N}} - z_{i,\mathrm{N}})^3(-y_{i,\mathrm{N}}) + (x_{i,\mathrm{N}} - z_{i,\mathrm{N}})(-y_{i,\mathrm{N}})^3\right) \mathrm{I}_{\{z_{i,\mathrm{N}}>0\}},$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_3}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) := d_3 \,(x_{i,\mathrm{N}} - z_{i,\mathrm{N}})^2 y_{i,\mathrm{N}}^2 \,\mathrm{I}_{\{z_{i,\mathrm{N}}>0\}},$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_{n_1}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) := d_1 \left(x_{i,\mathrm{N}}^4 + (z_{i,\mathrm{N}} - y_{i,\mathrm{N}})^4\right) \mathrm{I}_{\{z_{i,\mathrm{N}}<0\}},$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_{n_2}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) := d_2 \left(x_{i,\mathrm{N}}^3(z_{i,\mathrm{N}} - y_{i,\mathrm{N}}) + x_{i,\mathrm{N}}(z_{i,\mathrm{N}} - y_{i,\mathrm{N}})^3\right) \mathrm{I}_{\{z_{i,\mathrm{N}}<0\}},$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_{n_3}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) := d_3 \, x_{i,\mathrm{N}}^2(z_{i,\mathrm{N}} - y_{i,\mathrm{N}})^2 \,\mathrm{I}_{\{z_{i,\mathrm{N}}<0\}}\,.$$

Inserting these functions in (4), we obtain six estimators of IQ:

$$\widehat{\mathrm{IQ}}_{p_1} := \mathrm{N} \sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{p_1}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}})\,, \qquad \widehat{\mathrm{IQ}}_{p_2} := \mathrm{N} \sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{p_2}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}})\,,$$
$$\widehat{\mathrm{IQ}}_{p_3} := \mathrm{N} \sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{p_3}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) \qquad (5)$$

and

$$\widehat{\mathrm{IQ}}_{n_1} := \mathrm{N} \sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{n_1}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}})\,, \qquad \widehat{\mathrm{IQ}}_{n_2} := \mathrm{N} \sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{n_2}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}})\,,$$

$$\widehat{\mathrm{IQ}}_{n_3} := \mathrm{N}\sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{n_3}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right). \tag{6}$$

Note that the constants d_j, $j \in \{1,2,3\}$, which are given by[2]

$$d_1 = \frac{16}{3}, \qquad d_2 = \frac{32}{96\ln(2) - 54 - 9\zeta(3)}, \qquad d_3 = \frac{32}{3 - 2\zeta(3)},$$

make the estimators asymptotically unbiased. To calculate the desired values d_j exactly, we have to derive the common density of the high, low and close values of a standard Brownian motion, see formula (12). With the help of this density, we can calculate the reciprocals of the expected values of the estimators, which gives us the values of d_j, $j \in \{1,2,3\}$. Hence, the constants are independent from N and thus do not suffer from a finite sampling error, which can be especially severe when using ohlc values. Consequently, the estimators $\widehat{\mathrm{IQ}}_{p_j}$ and $\widehat{\mathrm{IQ}}_{n_j}$ are unbiased for standard Brownian motion. In the Web Appendix A, Lemma 1, **(I)** we prove the asymptotic unbiasedness of the estimators for processes of the form (A1).

The estimators in (5) and (6) are robust to jumps of finite activity, as is stated in the following Proposition

Proposition 1 *Let* $\mathrm{J} := (\mathrm{J}_t)_{t\in[0,1]}$ *be a jump process of finite activity,* $\mathrm{p}_t = \mathrm{p}_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s$*,* σ_t *is the never vanishing càdlàg (spot) volatility and* $\widehat{U}$ *denotes one of the six estimators* $\widehat{\mathrm{IQ}}_{p_1},\ldots,\widehat{\mathrm{IQ}}_{n_3}$*. Then it holds that*

$$\mathrm{N}\sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{U}}(x_{i,\mathrm{N}}^{(\mathrm{J})}, y_{i,\mathrm{N}}^{(\mathrm{J})}, z_{i,\mathrm{N}}^{(\mathrm{J})}) - \mathrm{N}\sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{U}}(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}) \xrightarrow{P} 0.$$

with $x_{i,\mathrm{N}}^{(\mathrm{J})} := ((\mathrm{p}+\mathrm{J})^*)_{i,\mathrm{N}} - (\mathrm{p}+\mathrm{J})_{\frac{i-1}{\mathrm{N}}}$*,* $y_{i,\mathrm{N}}^{(\mathrm{J})} := ((\mathrm{p}+\mathrm{J})_*)_{i,\mathrm{N}} - (\mathrm{p}+\mathrm{J})_{\frac{i-1}{\mathrm{N}}}$ *and* $z_{i,\mathrm{N}}^{(\mathrm{J})} := (\mathrm{p}+\mathrm{J})_{\frac{i}{\mathrm{N}}} - (\mathrm{p}+\mathrm{J})_{\frac{i-1}{\mathrm{N}}}$*.*

The proof is straightforward and will therefore not be shown.

We can prove consistency of the following combined estimators

$$\widehat{\mathrm{IQ}}_1 := \frac{1}{2}\left(\widehat{\mathrm{IQ}}_{p_1} + \widehat{\mathrm{IQ}}_{n_1}\right), \quad \widehat{\mathrm{IQ}}_2 := \frac{1}{2}\left(\widehat{\mathrm{IQ}}_{p_2} + \widehat{\mathrm{IQ}}_{n_2}\right), \quad \widehat{\mathrm{IQ}}_3 := \frac{1}{2}\left(\widehat{\mathrm{IQ}}_{p_3} + \widehat{\mathrm{IQ}}_{n_3}\right). \tag{7}$$

With the definitions

$$\mathrm{f}_{\widehat{\mathrm{IQ}}_1}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) := \frac{1}{2}\left(\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_1}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) + \mathrm{f}_{\widehat{\mathrm{IQ}}_{n_1}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right)\right),$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_2}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) := \frac{1}{2}\left(\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_2}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) + \mathrm{f}_{\widehat{\mathrm{IQ}}_{n_2}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right)\right),$$
$$\mathrm{f}_{\widehat{\mathrm{IQ}}_3}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) := \frac{1}{2}\left(\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_3}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) + \mathrm{f}_{\widehat{\mathrm{IQ}}_{n_3}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right)\right),$$

[2] With Riemann's Zeta function $\zeta(a) := \sum_{k=1}^{\infty} 1/k^a$.

the estimators read as

$$\widehat{\text{IQ}}_1 := \text{N} \sum_{i=1}^{\text{N}} \text{f}_{\widehat{\text{IQ}}_1} \left(x_{i,\text{N}}, y_{i,\text{N}}, z_{i,\text{N}} \right), \qquad \widehat{\text{IQ}}_2 := \text{N} \sum_{i=1}^{\text{N}} \text{f}_{\widehat{\text{IQ}}_2} \left(x_{i,\text{N}}, y_{i,\text{N}}, z_{i,\text{N}} \right),$$

$$\widehat{\text{IQ}}_3 := \text{N} \sum_{i=1}^{\text{N}} \text{f}_{\widehat{\text{IQ}}_3} \left(x_{i,\text{N}}, y_{i,\text{N}}, z_{i,\text{N}} \right).$$

Consistency of the estimators $\widehat{\text{IQ}}_1$, $\widehat{\text{IQ}}_2$ and $\widehat{\text{IQ}}_3$ is enough to guarantee consistency of the final (most efficient) estimator of IQ, which will be a convex combination of these three estimators.

Proposition 2 *Let the log price process* p *fulfill the assumptions in (A1) and* σ_t *is the never vanishing càdlàg (spot) volatility. Then we have for all* $j \in \{1, 2, 3\}$ *and* $\text{N} \to \infty$

$$\widehat{\text{IQ}}_j \xrightarrow{P} \int_0^1 \sigma_s^4 ds.$$

Proposition 2 can be proved by the help of an auxiliary estimator which depends on the high, low and close values of the standard Brownian motion W:

$$\widetilde{\widehat{V}} := \text{N} \sum_{i=1}^{\text{N}} (\xi_{\widehat{V}})_{i,\text{N}}, \quad (\xi_{\widehat{V}})_{i,\text{N}} := \sigma^4_{\frac{i-1}{\text{N}}} \, \text{f}_{\widehat{V}} \left(a_{i,\text{N}}, b_{i,\text{N}}, c_{i,\text{N}} \right) \tag{8}$$

with the definitions $a_{i,\text{N}} := (W^*)_{i,\text{N}} - W_{\frac{i-1}{\text{N}}}$, $b_{i,\text{N}} := (W_*)_{i,\text{N}} - W_{\frac{i-1}{\text{N}}}$, $c_{i,\text{N}} := W_{\frac{i}{\text{N}}} - W_{\frac{i-1}{\text{N}}}$, $(W^*)_1 := \sup_{t \in [0,1]} W_t$, $(W_*)_1 := \inf_{t \in [0,1]} W_t$ and $\widehat{V}$ denotes one of the estimator $\widehat{\text{IQ}}_1, \widehat{\text{IQ}}_2, \widehat{\text{IQ}}_3$. Then we have to show that

- the auxiliary estimator $\widetilde{\widehat{V}}$ converges to IQ in probability: $\widetilde{\widehat{V}} \xrightarrow{P} \text{IQ}$ and that
- the difference between the estimator $\widehat{V}$ and the auxiliary estimator vanishes: $\widehat{V} - \widetilde{\widehat{V}} \xrightarrow{P} 0$.

For the detailed steps see Web Appendix A.

Now we state a central limit theorem (CLT) for the estimators in (7). For this purpose, let us define the vectors

$$\widehat{\mathbf{IQ}}_{\boldsymbol{p}} := (\widehat{\text{IQ}}_{p_1}, \widehat{\text{IQ}}_{p_2}, \widehat{\text{IQ}}_{p_3})', \quad \widehat{\mathbf{IQ}}_{\boldsymbol{n}} := (\widehat{\text{IQ}}_{n_1}, \widehat{\text{IQ}}_{n_2}, \widehat{\text{IQ}}_{n_3})' \quad \text{and} \quad \widehat{\mathbf{IQ}} := (\widehat{\text{IQ}}_1, \widehat{\text{IQ}}_2, \widehat{\text{IQ}}_3)'.$$

Note that we need further assumptions on the volatility process and that we have to exclude price jumps for the CLT:

Theorem 1 (Central Limit Theorem) *Let the log price process* p *fulfill the assumptions in (A1) with* $\mathrm{J} = 0$ *and* σ_t *is the never vanishing spot volatility given by (A2). Then we have the following stable convergence in law:*

$$\sqrt{\mathrm{N}}\left(\widehat{\mathbf{IQ}} - \mathrm{IQ}\,\iota\right) \overset{d_S}{\to} R\int_0^1 \sigma_s^4 d\boldsymbol{B}_s^{(3)}$$

with

- $\iota := (1, 1, 1)'$,
- *a three-dimensional Brownian motion* $\boldsymbol{B} := (\boldsymbol{B_t^{(3)}})_{t\in[0,1]}$ *with standard Brownian motions* B^i, $i \in \{1, 2, 3\}$, *as entries, which are independent from each other and from the Brownian motion* W,
- *a matrix* $R \in \mathbb{R}^{3\times 3}$, *which consists of entries which depend on the moments of functions of* $((W^*)_1, (W_*)_1, W_1)$.

The proof of the CLT works in two steps, where the details are set out in the Web Appendix B.
First, we have to show the following convergence of the auxiliary estimator $\widetilde{\widetilde{\widehat{V}}}$:

$$\sqrt{\mathrm{N}}\left(\widetilde{\widetilde{\widehat{\boldsymbol{V}}}}^{(3)} - \mathrm{IQ}\,\iota\right) \overset{d_S}{\to} R\int_0^1 \sigma_s^4 d\boldsymbol{B_s^{(3)}}$$

with

- a 3-dimensional vector $\widetilde{\widetilde{\widehat{\boldsymbol{V}}}}^{(3)}$, consisting of entries of the form (8),
- $\widehat{V}$ denotes one of the estimators $\widehat{\mathrm{IQ}}_1, \widehat{\mathrm{IQ}}_2, \widehat{\mathrm{IQ}}_3$,
- the matrix $R \in \mathbb{R}^{3\times 3}$ and the three-dimensional Brownian motion $\boldsymbol{B}$ from the CLT.

Second, we have to prove that

$$\sqrt{\mathrm{N}}\left(\widehat{\mathbf{IQ}} - \widetilde{\widetilde{\widehat{\boldsymbol{V}}}}^{(3)}\right) \overset{P}{\to} \mathbf{0}$$

with the zero-vector $\mathbf{0} := (0, 0, 0)'$.

Now we can come to the construction of the final estimator $\mathrm{IQ}_{\mathrm{OHLC}}$. We want to derive the combination of the three jump-robust and consistent estimators $\widehat{\mathrm{IQ}}_1$, $\widehat{\mathrm{IQ}}_2$ and $\widehat{\mathrm{IQ}}_3$,

$$\mathrm{IQ}_{\mathrm{OHLC}} := w_1\widehat{\mathrm{IQ}}_1 + w_2\widehat{\mathrm{IQ}}_2 + w_3\widehat{\mathrm{IQ}}_3, \tag{9}$$

which has the smallest possible variance under the restriction $w'\iota = 1$ with the vector $w := (w_1, w_2, w_3)'$. From the CLT we have that the variance of $\mathrm{IQ}_{\mathrm{OHLC}}$ is given by

$$\mathrm{Var}\left(\mathrm{IQ}_{\mathrm{OHLC}}\right) = \Sigma \int_0^1 \sigma_s^8 ds \tag{10}$$

with $\Sigma := RR'$. Further, from the proof of Lemma 2, step (b) (see Web Appendix B) we know that the matrix Σ is given by

$$\Sigma := \mathbb{E}\left\{\widetilde{\boldsymbol{\chi}}^{(\mathbf{3})}_{\widehat{\boldsymbol{V}}} \widetilde{\boldsymbol{\chi}}^{(\mathbf{3})'}_{\widehat{\boldsymbol{V}}}\right\} \tag{11}$$

with a three-dimensional vector $\widetilde{\boldsymbol{\chi}}^{(\mathbf{3})}_{\widehat{\boldsymbol{V}}}$ consisting of entries of the form

$$\widetilde{\chi}_{\widehat{V}} := \mathrm{f}_{\widehat{V}}\left((W^*)_1, (W_*)_1, W_1\right) - 1 \quad \text{with} \quad \widehat{V} \in \{\widehat{\mathrm{IQ}}_1, \widehat{\mathrm{IQ}}_2, \widehat{\mathrm{IQ}}_3\}.$$

We can simplify the task of calculating the matrix Σ as it holds

Proposition 3 $\Sigma = \frac{1}{2}\Sigma_p - \frac{1}{2}\boldsymbol{E}_{3\times 3}$, *where all entries of the* 3×3*-matrix* $\boldsymbol{E}$ *are equal to 1.*

Thus it is enough to calculate the matrix $\Sigma_p := \mathbb{E}\left\{\widetilde{\boldsymbol{\chi}}^{(\mathbf{3})}_{\widehat{\boldsymbol{V}}_{\boldsymbol{p}}} \widetilde{\boldsymbol{\chi}}^{(\mathbf{3})'}_{\widehat{\boldsymbol{V}}_{\boldsymbol{p}}}\right\}$ with $\widehat{V}_p \in \{\widehat{\mathrm{IQ}}_{p_1}, \widehat{\mathrm{IQ}}_{p_2}, \widehat{\mathrm{IQ}}_{p_3}\}$. For the proof of the proposition see Web Appendix C.

To calculate the matrix Σ_p exactly, we have to derive the common density of the high, low and close values of a standard Brownian motion from the mixed distribution/density function given in Borodin and Salminen (2002, 174, formula 1.15.8):

$$\begin{aligned}
&\mathrm{P}\left(\inf_{0\le s\le t} W_s \in da, \sup_{0\le s\le t} W_s \in db, W_t \in dz\right) \\
=&\frac{4}{\sqrt{2\pi}}\left(\sum_{k=-\infty}^{-1} k^2 A_k\left(a,b,z\right) + \sum_{k=1}^{\infty} k^2 A_k\left(a,b,z\right)\right. \\
&\left.- \sum_{k=-\infty}^{-2} k(1+k) B_k\left(a,b,z\right) - \sum_{k=1}^{\infty} k(1+k) B_k\left(a,b,z\right)\right)
\end{aligned} \tag{12}$$

with

$$\begin{aligned}
A_k\left(a,b,z\right) &:= \exp\left(-\frac{(z+2k(b-a))^2}{2}\right)\left((z+2k(b-a))^2 - 1\right) \\
B_k\left(a,b,z\right) &:= \exp\left(-\frac{(z-2a+2k(b-a))^2}{2}\right)\left((z-2a+2k(b-a))^2 - 1\right)
\end{aligned}$$

and $a < 0$, $b > 0$. For example, the (1,3)-entry of the matrix Σ_p,

$$\mathbb{E}\left\{\widetilde{\chi}_{\widehat{\mathrm{IQ}}_{p_1}} \widetilde{\chi}_{\widehat{\mathrm{IQ}}_{p_3}}\right\} = \mathbb{E}\left\{\left(\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_1}}\left((W^*)_1, (W_*)_1, W_1\right) - 1\right)\left(\mathrm{f}_{\widehat{\mathrm{IQ}}_{p_3}}\left((W^*)_1, (W_*)_1, W_1\right) - 1\right)\right\},$$

reads as

$$\mathbb{E}\left\{\left(d_1\left[((W^*)_1 - W_1)^4 + (W_*)_1^4\right]\mathrm{I}_{\{W_1>0\}} - 1\right)\left(d_3\left[((W^*)_1 - W_1)^2 (W_*)_1^2\right]\mathrm{I}_{\{W_1>0\}} - 1\right)\right\},$$

which equals

$$c\,\mathbb{E}\left\{\left[((W^*)_1 - W_1)^4 + (W_*)_1^4\right]\mathrm{I}_{\{W_1>0\}}\left[((W^*)_1 - W_1)^2 (W_*)_1^2\right]\mathrm{I}_{\{W_1>0\}}\right\} - 1$$

(with $c := d_1 d_3 = \frac{512}{9-6\zeta(3)}$), as the expectation value of the estimators $\widehat{\mathrm{IQ}}_{p_1}$, $\widehat{\mathrm{IQ}}_{p_2}$ and $\widehat{\mathrm{IQ}}_{p_3}$ is 1 under a standard Brownian motion (due to the constants d_j, which make the estimators unbiased). Now, the expectation value can be calculated by the help of the density (12). Accomplishing these steps for all entries yields the matrix

$$\Sigma_p := \begin{pmatrix} c_1 & c_2 & c_3 \\ c_2 & c_4 & c_5 \\ c_3 & c_5 & c_6 \end{pmatrix} \approx \begin{pmatrix} 10.88225 & 5.84777 & 4.22960 \\ 5.84777 & 7.90597 & 8.36909 \\ 4.22960 & 8.36909 & 9.64623 \end{pmatrix} \tag{13}$$

with constants

$$\begin{aligned}
c_1 &:= \frac{70}{3} - \frac{2}{3}\zeta(7) - \frac{8}{3}\zeta(5) - \frac{20}{3}\zeta(3) - 1, \\
c_2 &:= \frac{512\left(\frac{3945}{128} - 60\ln(2) + \frac{345}{1024}\zeta(7) + \frac{855}{512}\zeta(5) + \frac{3675}{512}\zeta(3)\right)}{3(54 + 9\zeta(3) - 96\ln(2))} - 1, \\
c_3 &:= \frac{512(\frac{105}{256} - \frac{15}{256}\zeta(7) - \frac{15}{128}\zeta(5) - \frac{45}{256}\zeta(3))}{9 - 6\zeta(3)} - 1, \\
c_4 &:= \left(\frac{32}{96\ln(2) - 54 - 9\zeta(3)}\right)^2 \left(\frac{105}{128} - \frac{21}{256}\zeta(7) - \frac{27}{128}\zeta(5) - \frac{105}{256}\zeta(3)\right) - 1, \\
c_5 &:= \left(\frac{32}{3 - 2\zeta(3)}\right)\left(\frac{32}{54 + 9\zeta(3) - 96\ln(2)}\right) \cdot \\
&\quad \left(\frac{1065}{256} + \frac{15}{512}\zeta(7) + \frac{135}{1024}\zeta(5) + \frac{735}{1024}\zeta(3) - \frac{15}{2}\ln(2)\right) - 1, \\
c_6 &:= \left(\frac{32}{3 - 2\zeta(3)}\right)^2 \left(-\frac{30451}{41472} - \frac{3}{256}\zeta(7) - \frac{3}{64}\zeta(5) - \frac{73}{384}\zeta(3) + \frac{40}{27}\ln(2)\right) - 1.
\end{aligned}$$

By the help of matrix (13), we can calculate the vector of optimal weights, $w^* := (w_1^*, w_2^*, w_3^*)'$ with the following Lemma.

Lemma 1 *Given a matrix A of full rank and a right-hand side c, minimizing $w'\Sigma w$ under the condition $Aw = c$ is achieved by*

$$w^* = \Sigma^{-1} A' \left(A\Sigma^{-1}A'\right)^{-1} c \quad \textit{with} \quad (w^*)'\Sigma w^* = c'\left(A\Sigma^{-1}A'\right)^{-1} c.$$

Choosing $A = \Sigma_p$ and $c = (1,1,1)'$ (as we require $w'\tau = 1$) we obtain $w^* = (0.49349, -0.18630, 0.69281)'$ and $(w^*)'\Sigma_p w^* = 7.55353$. Thus, our final estimator is given by

$$\mathrm{IQ}_{\mathrm{OHLC}} = w_1^* \widehat{\mathrm{IQ}}_1 + w_2^* \widehat{\mathrm{IQ}}_2 + w_3^* \widehat{\mathrm{IQ}}_3 \tag{14}$$

with $(w^*)'\Sigma w^* = \frac{1}{2}7.55353 - \frac{1}{2} = 3.27676$ (see Proposition 3). Note that despite the negative weight, the entire estimator $\mathrm{IQ}_{\mathrm{OHLC}}$ is P almost surely greater or equal to 0, which can be checked by showing that the global minimum of the function is zero. We used MAPLE to do this.

Implication 1 *Let the log price process* p *fulfill the assumptions in (A1) with* $\mathrm{J} = 0$, σ_t *is the never vanishing spot volatility given by (A2) and* $\mathrm{IQ}_{\mathrm{OHLC}} := 0.49349\widehat{\mathrm{IQ}}_1 - 0.18630\widehat{\mathrm{IQ}}_2 + 0.69281\widehat{\mathrm{IQ}}_3$. *Then we have the following stable convergence in law:*

$$\sqrt{\mathrm{N}}\left(\mathrm{IQ}_{\mathrm{OHLC}} - \mathrm{IQ}\right) \xrightarrow{d_S} \mathbb{MN}\left(0,\ 3.27676 \int_0^1 \sigma_s^8 ds\right).$$

Note that the efficiency factor 3.27676 is smaller than the factor of the optimal estimator proposed in Jacod and Rosenbaum (2012), see formula (3.7) and example (3.9) therein, as our approach exploits the full range of data.

1.1 The role of zero returns

In Monte-Carlo studies, intraday zero increments constitute no problem, as the assumed model for the log prices rules them out theoretically. Nevertheless, they can occur in empirical applications with positive probability, as the market is not perfect. In fact, we have to deal with microstructure noise effects or stocks which are not perfectly liquid. This can lead to intraday zero increments. Consequently, neglecting them could distort our estimator of integrated quarticity (which takes just positive and negative increments into account so far) in empirical applications. Therefore, we additionally consider the estimators

$$\widehat{\mathrm{IQ}}_{m_1} := \mathrm{N}\sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{m_1}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right), \qquad \widehat{\mathrm{IQ}}_{m_2} := \mathrm{N}\sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{m_2}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right),$$

$$\widehat{\mathrm{IQ}}_{m_3} := \mathrm{N}\sum_{i=1}^{\mathrm{N}} \mathrm{f}_{\widehat{\mathrm{IQ}}_{m_3}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right)$$

with

$$\begin{aligned}
\mathrm{f}_{\widehat{\mathrm{IQ}}_{m_1}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) &:= d_1\left((x_{i,\mathrm{N}} - z_{i,\mathrm{N}})^4 + y_{i,\mathrm{N}}^4\right)\mathrm{I}_{\{z_{i,\mathrm{N}}=0\}}, \\
\mathrm{f}_{\widehat{\mathrm{IQ}}_{m_2}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) &:= d_2\left((x_{i,\mathrm{N}} - z_{i,\mathrm{N}})^3(-y_{i,\mathrm{N}}) + (x_{i,\mathrm{N}} - z_{i,\mathrm{N}})(-y_{i,\mathrm{N}})^3\right)\mathrm{I}_{\{z_{i,\mathrm{N}}=0\}}, \\
\mathrm{f}_{\widehat{\mathrm{IQ}}_{m_3}}\left(x_{i,\mathrm{N}}, y_{i,\mathrm{N}}, z_{i,\mathrm{N}}\right) &:= d_3\left(x_{i,\mathrm{N}} - z_{i,\mathrm{N}}\right)^2 y_{i,\mathrm{N}}^2\,\mathrm{I}_{\{z_{i,\mathrm{N}}=0\}}.
\end{aligned}$$

Thus, estimators accounting for zero returns are given by

$$\widehat{\mathrm{IQ}}_{1(m)} := \frac{1}{2}\left(\widehat{\mathrm{IQ}}_{p_1} + \widehat{\mathrm{IQ}}_{n_1}\right) + \widehat{\mathrm{IQ}}_{m_1}, \quad \widehat{\mathrm{IQ}}_{2(m)} := \frac{1}{2}\left(\widehat{\mathrm{IQ}}_{p_2} + \widehat{\mathrm{IQ}}_{n_2}\right) + \widehat{\mathrm{IQ}}_{m_2},$$

$$\widehat{\mathrm{IQ}}_{3(m)} := \frac{1}{2}\left(\widehat{\mathrm{IQ}}_{p_3} + \widehat{\mathrm{IQ}}_{n_3}\right) + \widehat{\mathrm{IQ}}_{m_3}.$$

Hence, we use the estimator

$$\mathrm{IQ}_{\mathrm{OHLC}(m)} = w_1^* \widehat{\mathrm{IQ}}_{1(m)} + w_2^* \widehat{\mathrm{IQ}}_{2(m)} + w_3^* \widehat{\mathrm{IQ}}_{3(m)}$$

in empirical applications.

1.2 Bounce-Backs

We speak about so-called bounce-backs when the price jumps in one direction but immediately returns to its old level. These movements are a form of market microstructure noise (see Aït-Sahalia et al. (2006) and Andersen et al. (2012)) and result in a severe distortion of the estimator $\mathrm{IQ}_{\mathrm{OHLC}}$. The reason is that $\mathrm{IQ}_{\mathrm{OHLC}}$ uses only the information contained in the i-th intraday interval in order to obtain the i-th estimated value of IQ. To illustrate the effect of bounce-backs on the estimator let us assume that in some intraday interval the price jumped down and immediately returned to its former level and that the corresponding intraday return was negative. Consequently, the corresponding lowest value is comparatively far off the closing value (as the price bounced back). Thus, the candlestick $z_{i,\mathrm{N}} - y_{i,\mathrm{N}}$ will be very large, resulting in a bias in $\mathrm{IQ}_{\mathrm{OHLC}}$.

2 Simulation Study

We investigate the ohlc-based estimator under five models and compare its performance with that of realized quarticity. The models considered are:

- Standard Brownian motion (SBM),
- Standard Brownian motion with small adjacent jumps,
- Standard Brownian motion with bounce-backs,
- Standard Brownian motion with large jumps,
- A stochastic volatility (SV) model without jumps.

The required processes are simulated by the Euler scheme. We simulate 10000 days, where every day is based on a simulation step size of 390 (N) times 50000 (f) in order to create an equally spaced frequency of 1 minute (a sampling frequency of 1 minute corresponds to $\mathrm{N} = 390$ subdivisions per day for a presumed trading day of 6.5 hours). The simulation requires a high number of discretization steps in order to establish intraday high and low values which are as unbiased as possible. Lower frequencies are derived from the original sample space by aggregation.

For the case of four small adjacent jumps we choose a deterministic jump height of $3\sqrt{1/\mathrm{N}}$ with $\mathrm{N} = 390$ (inspired by Kolokolov and Renò (2012)). We randomly place the first jump in one of the first N-3 intraday intervals. The three remaining jumps are forced to appear in the three following subsequent intervals (if the first jump is set, for example, in the third intraday interval, the other three occur in the fourth, fifth and sixth). For the case of bounce-backs, we simulate a standard Brownian motion with one randomly placed jump of normal distributed jump height (which ensures a jump contribution to total daily volatility of 10%, as in Huang and Tauchen (2005)). The second jump is placed in the same 1-minute interval and is forced to have the same jump height but to be of opposite sign. For the case of large jumps, we place one jump of deterministic height of $10\sqrt{1/\mathrm{N}}$, $\mathrm{N} = 390$, randomly in one of the intraday intervals. For the simulation of the stochastic volatility model without jumps, we follow the suggestions in Jacod and Todorov (2009), Veraart (2010), Andersen and others, Chernov and others and Andersen et al. (2002). Thus, we consider the process

$$\begin{aligned} d\,\mathrm{p}_t &= \mu dt + \exp(\beta_0 + \beta_1 \sigma_t) dW_t^{\mathrm{p}} \\ d\sigma_t &= \alpha_\sigma \sigma_t dt + dW_t^\sigma \end{aligned} \tag{15}$$

with standard Brownian motions $W^{\mathrm{p}} = (W_t^{\mathrm{p}})_{t\in[0,1]}$ and $W^\sigma = (W_t^\sigma)_{t\in[0,1]}$ with $\mathrm{corr}(W^{\mathrm{p}}, W^\sigma) = \rho$, $(\sigma_t)_{t\in[0,1]}$ is the stochastic volatility process and α_σ describes the rate of mean reversion of the volatility. The parameters are given by $\mu = 0$, $\beta_0 = 0$, $\beta_1 = 0.125$ and $\alpha_\sigma = -0.1$ (medium mean reversion, half life just over one week) and $\rho = -0.62$. The mean of the approximated integrated quarticity of the simulated data is 1.06, where IQ is approximated by $\frac{1}{\mathrm{N}\cdot f}\sum_{i=1}^{\mathrm{N}\cdot f} \sigma^4_{\frac{i-1}{\mathrm{N}\cdot f}}$ with $\mathrm{N} = 390$. Its minimum is 0.41 and its maximum is 3.07.

Table 2 states the (averaged) variance of $\sqrt{\mathrm{N}}\left(\widehat{\mathrm{IQ}} - \mathrm{IQ}\right)/\sqrt{\int_0^1 \sigma_\tau^8 d\tau}$ for all simulated data sets for a frequency of 1 minute. We know from Implication 1 that this expression should be around 3.28 for $\widehat{\mathrm{IQ}} = \mathrm{IQ}_{\mathrm{OHLC}}$ and around 10.66 (see, for example, Andersen et al. (2011)) when $\widehat{\mathrm{IQ}} = \mathrm{RQ}$, at least for the models which meet the theoretical assumptions (A1) and (A2). In this way, the correctness of both the calculations and the simulations could be checked.

For the evaluation, we state the bias and the mean square error (MSE). The real unknown IQ is 1 in the case of a standard Brownian motion and is approximated by the formula mentioned above for the SV model. For each day, we also compute the confidence interval (CI) for integrated volatility $\mathrm{IV} := \int_0^1 \sigma_\tau^2 d\tau$ (estimated by RV) with a significance level of $\alpha = 0.05$. We use IQ_{OHLC} and RQ as estimates for the required integrated quarticity and state the length (calculated as the median over all days) and the empirical size of the resulting intervals. The size is the proportion of days on which the real IV (in the case of SBM plus jumps, IV equals 1. In the case of stochastic volatility, IV is approximated by $\frac{1}{\mathrm{N}\cdot f}\sum_{i=1}^{\mathrm{N}\cdot f}\sigma^2_{\frac{i-1}{\mathrm{N}\cdot f}}$) lies in the CI. Optimally, the size should be around 1-α. Table 1 states all findings.

2.1 Standard Brownian motion

The first two columns of Table 1 show the bias and the MSE. The bias is small with a distortion less than 1% for RQ and around 1% for IQ_{OHLC}. The MSE confirms the theoretical result that IQ_{OHLC} fluctuates not as heavily as RQ. We obtain quite precise coverage rates of the confidence intervals using the new ohlc-based estimator. Up to 10-minute frequency, the size is more or less around 95% whereas the confidence intervals using RQ cover the unknown integrated volatility in less than 94% of the days already for frequencies lower or equal to 5 minutes. Comparing the lengths of the confidence intervals with nearly correct coverage rates, we see that the numbers are quite similar. At first glance this seems to be surprising since the MSE of IQ_{OHLC} is much smaller than the MSE of RQ. But we have to bear in mind that RQ fluctuates in both directions. In fact, in roughly half of the cases, the value of RQ is smaller than the value of IQ_{OHLC}, creating a CI of smaller length. As we take the median over all lengths, the values of the lengths do not differ. As realized quarticity and realized variance are similar by construction, it is likely that on days when RV deviates a lot from the true value of IV, RQ is biased also, likely in the same direction. Consequently, when RV differs considerably from the unknown value of integrated volatility, the less fluctuating estimator IQ_{OHLC} cannot create a CI which is superior to the confidence interval including RQ in terms of covering the unknown IV.

2.2 Standard Brownian motion with adjacent jumps

RQ and RV are known to be biased in the case of adjacent jumps. Thus, not surprisingly, the bias and the MSE quickly deteriorate with decreasing sampling frequency. However, the influence of adjacent jumps diminishes with lower frequencies since we placed the jumps on a 1-minute frequency. By aggregating the data, the probability increases that all jumps lie in the same interval. This can still interfere with RV and RQ as they are not jump-robust. Regarding IQ_{OHLC} the bias is quite small for high frequencies. Compared to the results under SBM, the bias is worse for lower frequencies when jumps start to aggregate in some interval. However, the MSE remains small. Considering the size and the length of the confidence intervals, we observe the following: in the case of RQ, the size of the CI exceeds the value 0.95 in general, whereas the size is too small in the case of IQ_{OHLC}. We attribute this to the fact that both RV and RQ are upward biased because of jumps (for example, the mean of RV is 1.28 for 5-minute frequency). Consequently, we have a higher chance to cover the unknown integrated volatility using RQ (which is highly upward biased too).

2.3 Standard Brownian motion with bounce backs

As already discussed, the ohlc-based estimator is not robust to bounce-backs. As can be seen from the bias and the MSE, IQ_{OHLC} is severely affected, where the performance improves with decreasing frequency when the effect of the bounce-backs on the returns diminishes and the continuous part of the price process becomes more influential. Since IQ_{OHLC} is highly upward biased and fluctuates heavily, the size of the confidence intervals exceeds the desired confidence level. The lengths are surprisingly close to the lengths of the confidence intervals using RQ. But we have to be aware of the fact that we state the median of the lengths. In fact, in the case of bounce-backs, the length of the CI can reach 11.88 using IQ_{OHLC} (for 30-minute frequency) compared to 5.41 using RQ. In general, the results for RQ are in line with the results found under SBM. This confirms that RQ is robust to bounce-backs.

2.4 Standard Brownian motion and one large jump per day

Not surprisingly, the bias and the MSE is very high for RQ, since it is not robust to jumps. Consequently, the length of the confidence interval is comparably large (larger than in the case of small jumps for 1- and 3-minute frequency). The coverage rates are quite high, which – except for 30-minute frequency – exceed the value 0.97. Regarding the new ohlc-based estimator, the findings confirm small fluctuations and high robustness to jumps. The lengths of the confidence intervals are comparable to the lengths under SBM. However, the intervals fail to cover the value of IV too often. We attribute this again to the fact that RV is highly upward biased.

2.5 SV model without jumps

In general, both estimators work well in the case of stochastic volatility. The bias is negligible and the MSE is small. IQ_{OHLC} fluctuates less around the true value than RQ, especially for lower frequencies. The sizes of the confidence intervals using IQ_{OHLC} are more or less around 0.95 (except for a frequency of 30 minutes). Regarding RQ, the coverage rate of the confidence interval is acceptable for a frequency of 1-minute. For lower frequencies, the real integrated volatility is covered on less than 94% of the days. Using IQ_{OHLC}, we produce narrower intervals compared to using RQ, where the lengths do not differ much in their median but in the range of lengths. Thus, using RQ, the lengths vary from 0.28 to 1.44, whereas we observe lenghts from 0.34 to 1.17 using IQ_{OHLC}.

3 Conclusion

Integrated quarticity describes the fluctuations in volatility and is a crucial ingredient in constructing confidence intervals for volatility. In this work, we propose a new estimator of integrated quarticity based on the intraday open, high, low and close (ohlc) values of a log price process. Thus, the whole observed process enters into the estimation, leading to an estimator which is very efficient. We investigate the performance of the new estimator under various scenarios which can potentially decrease its accuracy. In particular, we investigate the size and the length of confidence intervals for integrated volatility. The results are compared with the findings obtained by using the well-known realized quarticity. It turns out that the new proposed estimator is robust against large and adjacent jumps. We also briefly discuss drawbacks involved in using intraday ohlc data. More Monte-Carlo studies have to be conducted to examine alternative estimators of integrated quarticity provided by the literature.

References

Aït-Sahalia, Y., P. A. Mykland, and L. Zhang (2006). Comment. *Journal of Business & Economic Statistics 24*(2), 162–167.

Andersen, T. G., L. Benzoni, and J. Lund (2002). An empirical investigation of continuous-time equity return models. *The Journal of Finance 57*(3), 1239–1284.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2007). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *The Review of Economics and Statistics 89*(4), 701–720.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2011). A functional filtering and neighborhood approach to integrated quarticity estimation. NBER Working Paper 17152.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2012). Jump-robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics 169*(1), 75–93.

Barndorff-Nielsen, O. E. and N. Shephard (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society: Series B 64*(2), 253–280.

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

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

Barndorff-Nielsen, O. E., N. Shephard, and M. Winkel (2006). Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and their Applications 116*(5), 796–806.

Barndorff-Nielsen, O. E. and A. E. D. Veraart (2012). Stochastic volatility of volatility and variance risk premia. *Journal of Financial Econometrics 11*(1), 1–46.

Bollerslev, T. and Mikkelsen (1999). Long-term equity anticipation securities and stock market volatility dynamics. *Journal of Econometrics 92*(1), 75–99.

Busse, J. A. (1999). Volatility timing in mutual funds: evidence from daily returns. *Review of Financial Studies 12*(5), 1009–1041.

Carr, P. and L. Wu (2009). Variance risk premiums. *Review of Financial Studies 22*(3), 1311–1341.

Christensen, K., R. Oomen, and M. Podolskij (2011). Fact or friction: jumps at ultra high frequency. CREATES Research Paper 2011-19.

Christensen, K. and M. Podolskij (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics 141*(2), 323–349.

Clément, E., S. Delattre, and A. Gloter (2013). An infinite dimensional convolution theorem with applications to the efficient estimation of the integrated volatility. *Stochastic Processes and their Applications 123*(7), 2500–2521.

Corsi, F., N. Fusari, and D. La Vecchia (2013). Realizing smiles: Options pricing with realized volatility. *Journal of Financial Economics 107*(2), 284–304.

Corsi, F., D. Pirino, and R. Renò (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics 159*(2), 276–288.

Fleming, J., C. Kirby, and B. Ostdiek (2003). The economic value of volatility timing using "realized" volatility. *Journal of Financial Economics 67*(3), 473–509.

Garman, M. B. and M. J. Klass (1980). On the estimation of security price volatilities from historical data. *Journal of Business 53*(1), 67–78.

Huang, X. and G. Tauchen (2005). The relative contribution of jumps to total price variance. *Journal of Financial Econometrics 3*(4), 456–499.

Hull, J. and A. White (1987). The pricing of options on assets with stochastic volatilities. *The Journal of Finance XLII*, 281–300.

Jacod, J. and M. Rosenbaum (2012). Quarticity and other functionals of volatility: efficient estimation. *Annals of Statistics 41*(3), 1462–1484.

Jacod, J. and V. Todorov (2009). Testing for common arrivals of jumps for discretely observed multidimensional processes. *Annals of Statistics 37*(4), 1792–1838.

Klößner, S. (2009). Separating risk due to diffusion, positive jumps and negative jumps. Working Paper.

Kolokolov, A. and R. Renò (2012). Efficient multipowers. Working Paper.

Mancini, C. (2009). Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics 36*(2), 270–296.

Mancino, M. and S. Sanfelici (2008). Robustness of fourier estimator of integrated volatility in the presence of microstructure noise. *Computational Statistics & Data Analysis 52*(6), 2966–2989.

Mancino, M. E. and S. Sanfelici (2012). Estimation of quarticity with high-frequency data. *Quantitative Finance 12*(4), 607–622.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business 53*(1), 61–65.

Veraart, A. E. D. (2010). Inference for the jump part of quadratic variation of Itô semimartingales. *Econometric Theory 26*(2), 331–368.

		bias		MSE		size of CI		length of CI	
Estimator		RQ	IQ_{OHLC}	RQ	IQ_{OHLC}	RQ	IQ_{OHLC}	RQ	IQ_{OHLC}
SBM	1 min	0.0013	-0.0162	0.0282	0.0081	0.9459	0.9449	0.2790	0.2780
	3 min	0.0033	-0.0102	0.0819	0.0249	0.9441	0.9506	0.4771	0.4808
	5 min	0.0051	-0.0114	0.1417	0.0397	0.9316	0.9424	0.6105	0.6178
	10 min	0.0001	-0.0062	0.2694	0.0803	0.9164	0.9437	0.8413	0.8686
	30 min	-0.0011	-0.0094	0.7836	0.2271	0.8662	0.9285	1.3341	1.4664
SBM+4jumps	1 min	0.4630	0.0231	0.3221	0.0116	0.8744	0.7325	0.3343	0.2827
	3 min	1.9459	0.0061	6.4391	0.0266	0.9535	0.5765	0.7741	0.4848
	5 min	2.5693	0.0156	13.8908	0.0476	0.9718	0.5794	1.0351	0.6252
	10 min	2.0524	0.0383	10.0801	0.1104	0.9793	0.6608	1.3534	0.8838
	30 min	1.2520	0.1892	7.0691	0.5742	0.9573	0.7777	1.9031	1.5491
SBM+bb	1 min	-0.0002	16.7150	0.0274	2895.3638	0.9444	0.9785	0.2788	0.4023
	3 min	0.0006	5.3096	0.0819	303.7859	0.9344	0.9723	0.4773	0.5742
	5 min	-0.0036	3.1659	0.1327	109.2183	0.9311	0.9708	0.6082	0.7177
	10 min	-0.0076	1.6155	0.2528	27.7174	0.9155	0.9673	0.8368	0.9891
	30 min	-0.0170	0.6135	0.7390	4.7996	0.8631	0.9475	1.3210	1.6228
SBM+large jump	1 min	9.0711	-0.0059	95.3213	0.0088	0.9940	0.0958	0.8661	0.2792
	3 min	3.3507	0.0209	16.5606	0.0339	0.9812	0.4818	0.9522	0.4867
	5 min	2.2534	0.0462	9.0887	0.0677	0.9788	0.6270	1.0391	0.6323
	10 min	1.3712	0.0908	4.9071	0.1692	0.9770	0.7493	1.2211	0.8965
	30 min	0.8200	0.2088	4.1028	0.5667	0.9447	0.8238	1.7360	1.5689
SV model	1 min	-0.0019	-0.0172	0.0338	0.0106	0.9428	0.9447	0.2806	0.2790
	3 min	-0.0027	-0.0069	0.1053	0.0305	0.9388	0.9449	0.4791	0.4843
	5 min	-0.0013	-0.0082	0.1683	0.0501	0.9301	0.9433	0.6106	0.6229
	10 min	-0.0057	-0.0048	0.3470	0.0975	0.9137	0.9422	0.8394	0.8750
	30 min	0.0127	-0.0015	1.1967	0.2950	0.8650	0.9246	1.3250	1.4740

Table 1: Obtained results for RQ and IQ_{OHLC} and the corresponding confidence intervals (CI) for SBM, SBM with 4 small adjacent jumps, SBM with bounce-backs (bb), SBM with one large jump and the stochastic volatility model.

SBM		SBM+4jumps		SBM+bb		SBM+large jump		SV model	
RQ	IQ_{OHLC}	RQ	IQ_{OHLC}	RQ	IQ_{OHLC}	RQ	IQ_{OHLC}	RQ	IQ_{OHLC}
11.00	3.06	42.01	4.32	10.70	1020332	5085	3.41	10.69	3.19

Table 2: Efficiency factors of RQ and IQ_{OHLC} for the simulated data sets for 1-minute frequency.