A Robust Neighborhood Truncation Approach to Estimation of Integrated Quarticity

Torben Andersen, Dobrislav Dobrev, and Ernst Schaumburg

A Robust Neighborhood Truncation Approach to Estimation of Integrated Quarticity

Torben G. Andersen, Dobrislav Dobrev, Ernst Schaumburg [*]

This version: December 2012;   Initial draft: March 2010

Abstract

We provide a first in-depth look at robust estimation of integrated quarticity (IQ) based on high frequency data. IQ is the key ingredient enabling inference about volatility and the presence of jumps in financial time series and is thus of considerable interest in applications. We document the significant empirical challenges for IQ estimation posed by commonly encountered data imperfections and set forth three complementary approaches for improving IQ based inference. First, we show that many common deviations from the jump diffusive null can be dealt with by a novel filtering scheme that generalizes truncation of individual returns to truncation of arbitrary functionals on return blocks. Second, we propose a new family of efficient robust neighborhood truncation (RNT) estimators for integrated power variation based on order statistics of a set of unbiased local power variation estimators on a block of returns. Third, we find that ratio-based inference, originally proposed in this context by Barndorff-Nielsen and Shephard (2002), has desirable robustness properties in the face of regularly occurring data imperfections and thus is well suited for empirical applications. We confirm that the proposed filtering scheme and the RNT estimators perform well in our extensive simulation designs and in an application to the individual Dow Jones 30 stocks.

Keywords: Robust Neighborhood Truncation Estimator, Functional Filtering, Integrated Quarticity, Inference on Integrated Variance, Inference on Jumps, High-Frequency Data

JEL classification: C14; C15; C22; C80; G10.

1   Introduction

Important progress in measuring and forecasting return volatility has been obtained through techniques exploiting the information in intraday price movements. The use of high-frequency data is, however, not without its problems. The main complication is the pronounced inhomogeneity of the intraday return series as diurnal patterns interspersed with news events and market microstructure frictions complicate direct modeling of the high frequency dynamics and introduce a variety of idiosyncratic features that are largely irrelevant for inference about inter-daily volatility. The realized volatility (RV) approach "solves" this problem by aggregating the intraday return observations to a daily frequency in a manner that retains the majority of the inherent volatility information while mitigating the impact of noise and diurnal patterns. The RV approach has been widely adopted ever since its formal introduction as a nonparametric estimator of the return variation in Andersen and Bollerslev (1998).[1] In parallel, a large body of theoretical work on model-free estimation and inference for components of the realized return variation process has arisen. Initial econometric issues are addressed in Andersen, Bollerslev, Diebold and Labys (2001, 2003) and Barndorff-Nielsen and Shephard (henceforth BNS) (2002).[2]

Conceptually, realized volatility differs from the standard notion of volatility by focusing on ex-post measurement of the realization of the (stochastic) return variation rather than the (ex-ante) return variance. Once attention shifts to the actual volatility realizations, new questions arise. For example, how do we assess the accuracy of our (daily) ex-post measures of the integrated return variation and how do we identify the impact of jump components. Such features are critical for a variety of issues in real-time financial management, including volatility forecasting, analysis of the dynamic properties of jumps and news events, derivatives pricing, estimation of return correlations, determination of return-volatility asymmetries (the leverage effect), and developing insights into the interplay between return volatility and the macroeconomic environment.

The key ingredient for inference regarding the return variation and the presence of jumps is the so-called integrated quarticity (IQ). To illustrate the importance of accurate IQ measures we review a few results from the RV literature. We denote the continuously evolving log-price for a financial asset by Y_t. Under general conditions, the log-price constitutes a semi-martingale with respect to an underlying filtered probability space. The associated ex-post realized quadratic variation, QV, for Y_t over $[0,t]$ may be decomposed

---

[1]Corresponding measures were previously considered on an ad hoc basis, see, e.g., Schwert (1989), Hsieh (1991) and Zhou (1996).

[2]Additional work regarding inference on the jump and continuous components of the return variation is developed in BNS (2004b, 2006), while general methodological insights were provided by Jacod and Protter (1998). Subsequently, market microstructure complications were addressed as part of the drive to exploit tick level data, see, e.g., Zhang, Mykland and Ait-Sahalia (2005), Bandi and Russell (2008), and Barndorff-Nielsen, Hansen, Lunde and Shephard, henceforth BNHLS, (2008).

into an integrated (diffusive) volatility, IV, and a residual (jump) component, JV,

$$
\begin{aligned}
dY_t &= a_t\, dt + \sigma_t\, dB_t + dJ_t \\
QV_t &= \underbrace{\int_0^t \sigma_u^2\, du}_{IV} + \underbrace{\sum_{0 \le u \le t} (\Delta J_u)^2}_{JV}
\end{aligned}
\tag{1}
$$

where a_t and σ_t denote the instantaneous drift and diffusion coefficients, while B_t and J_t are adapted Wiener and finite activity jump processes, respectively.

For a given trading day, $t \in [0,1]$, we consider the ideal scenario in which we observe N equally-spaced (log) returns, $r_i = Y_{\frac{i}{N}} - Y_{\frac{i-1}{N}}, \quad i = 1, \cdots, N.$ In this case, the realized volatility (RV) is a consistent nonparametric estimator of QV, as the number of intraday observations diverges, $N \to \infty$ (in-fill asymptotics),

$$
RV_N = \sum_{i=1}^{N} r_i^2 \;\to\; QV\,.
$$

Moreover, *absent price jumps*, the limiting distribution is a Gaussian mixture,

$$
\sqrt{N}\,(RV_N - IV) \quad \to \quad \mathcal{N}(0, 2\,IQ)
$$

where $IQ = \int_0^1 \sigma_u^4\, du$, which, as observed by BNS (2002), can be consistently estimated from the high-frequency data themselves via the Realized Quarticity (RQ) statistic:

$$
RQ_N = \frac{N}{3} \sum_{i=1}^{N} r_i^4 \;\to\; IQ
$$

Clearly, accurate inference about the integrated variance hinges on reliable estimates for IQ. Unfortunately, IQ estimation is challenging. It involves estimating fourth order return moments from noisy intraday return series impacted by the confounding effects of market microstructure frictions, diurnal patterns, outliers, and other data irregularities. For example, it is well known that the RQ estimator is highly imprecise and non-robust to such features, even if jumps are absent. Moreover, when discrete price changes do occur, RV is no longer consistent for IV, and the RQ statistic diverges: $RQ_N \to \infty$ as $N \to \infty$. Given the compelling evidence for jumps, this is critical in practice. In response, various jump-robust IQ estimators have been developed, but they are subject to potentially serious finite sample biases. At present, there simply is no systematic evidence regarding the performance of alternative jump-robust procedures for empirically realistic scenarios.

Recognizing these issues, a variety of ad hoc IQ estimation procedures have been implemented in the empirical literature. Before the jump-robust theory was developed, the RQ statistic was used, but only with relatively coarse sampling. For example, BNS (2004a) exploit 10-minute foreign exchange returns, while Bandi and Russell (2008) recommend computing RQ from 15- or 20-minute returns, as sparse sampling mitigates the impact of outliers and microstructure noise. Later, BNS (2004b) and Huang and Tauchen (2005)

rely on 5-minute returns for constructing jump-robust estimators of IQ.[3] Finally, due to the distortions arising from market microstructure effects, Jiang and Oomen (2008) opt for simply squaring their jump-robust IV estimator to obtain an IQ estimator, thus settling for a substantial Jensen inequality bias, but aiming to reduce estimation uncertainty.

To illustrate the practical importance of jump robustness, consider drawing inference about the IV of IBM stock returns across three days in February 2008 using the non-jump robust RQ/RV measures versus a pair of jump robust measures, as shown on Figure 1.[4]

[Insert Figure 1 about here.]

The jump on 2/26/2008 is readily identified visually and easily detected using a jump robust test statistic. In fact, the robust MedRV estimates and associated standard error bands, based on MedRQ, suggest a relative stable volatility process across the three trading days. In contrast, the regular RV estimate for IV is greatly inflated on 2/26/2008, and the confidence band is huge, reflecting a diverging RQ statistic. Hence, the reliance on non jump-robust statistics has two consequences. First, when jumps are present the IV estimate is upward biased because the jump component in QV is attributed to IV.[5] Second, the associated confidence band is grossly overstated, indicating very poor estimation precision whereas, in fact, the robust estimate appears quite reliable. Hence, non-robust inference may produce excessively erratic IV estimates and convey a sense of exaggerated imprecision associated with these techniques. While the misleading inference afforded by the regular RV and RQ estimators is apparent in Figure 1, at least when contrasted with the robust inference and a depiction of the price path, it can be less obvious in cases with higher volatility levels and relatively smaller jumps. As such, it is important to develop feasible robust and efficient procedures for estimating IQ and conducting inference for IV.

One main contribution of this paper is to provide a first in-depth exploration of the virtues and drawbacks of alternative jump-robust estimation procedures for IQ, including their robustness to a variety of realistic features of the return generating process. A point of emphasis is the use of wide pre-averaging windows for controlling the impact of microstructure noise on the inference. This enhances robustness and simplifies the distribution theory as the impact of noise is annihilated asymptotically. A second contribution is the development of a new class of *robust neighborhood truncation* (RNT) estimators that generalize existing nearest neighbor and Quantile RV estimators. They involve the application of a second layer of order statistics to suitably chosen return functionals, thus robustifying the inference for IQ with only a minor loss of efficiency. We find such RNT estimators to perform

---

[3] Both studies explore the reliability of the procedures, but only under relatively ideal circumstances and not with a focus on the IQ estimator but rather the jump test statistic.

[4] We rely on the MedRV/MedRQ estimators of Andersen, Dobrev and Schaumburg (2009) here, but any other sensible pair of robust IV and IQ estimators would suffice in this particular case as the evidence for a single significant jump is compelling and the associated empirical inference problem thus very straightforward. The robust estimators will be introduced later in the paper.

[5] As a consequence, real-time predictions regarding the expected volatility over the following trading days will likely also be exaggerated on 2/26/2008 because the jump component of QV typically is much less persistent than the IV component, see, e.g., Andersen, Bollerslev and Diebold (2007).

admirably, especially when used in combination with the ratio statistic, IQ/IV^2, which is known to provide improved finite sample inference for IV. Moreover, these principles apply generally and can be used to enhance the robustness of inference from alternative classes of estimators. A third novelty is the use of an outlier filtering procedure that operates directly on an estimation functional of interest rather than on individual returns. This *functional filtering principle* adapts the filter to the specific assumptions underlying a given estimator. Hence, it controls the impact, and potential distortion, of abnormal outliers within the exact metric in which they contribute to the ultimate estimator. In applications to individual equity return data we find this filter indispensable for rendering entire classes of promising candidate IQ estimators viable. The unifying theme behind our new estimators and universal filtering procedure is to operate directly on the functional space of local power variation estimates rather than the individual returns. Nonetheless, the latter, and common, approach may be obtained as a special case of our procedure.

The remainder of the paper is structured as follows. Section 2 reviews the modern approach to robust estimation of integrated power variation. Section 3 develops our robust neighborhood truncation estimators. In Section 4, we discuss additional procedures applied to obtain robustness against jumps and noise. Section 5 illustrates the importance of common data features for IQ inference through an extensive simulation study. Finally, Section 6 provides evidence using high-frequency returns on the Dow Jones 30 stocks, while Section 7 concludes. All proofs are relegated to the Appendix.

2 Overview of Jump-Robust Power Variation Estimation

This section summarizes the modern approach to power variation estimation. We outline the theoretical setting and review some existing estimators which are later used in our simulation study and empirical investigation. In the process, we discuss practical trade-offs that must be confronted in estimating objects involving high powers of volatility.

2.1 The Theoretical Setting

We focus on a single asset traded continuously in a frictionless market over the period $[0, 1]$, referred to as a trading day. If it is a limited-liability asset with an expected positive payoff at some future date, the price will remain strictly positive. No-arbitrage conditions then ensure that the log-price process constitutes a semimartingale with respect to the underlying filtered probability space, see, e.g., Back (1991) and Andersen, Bollerslev and Diebold (2010). Hence, for most of our analysis we invoke the following conditions.

Assumption 1 The continuously compounded return process, r_t, is governed by a jump-diffusive semimartingale,

$$r_t \;=\; Y_0 \,+\, \int_0^t a_u \, du \,+\, \int_0^t \sigma_{u_-} \, dB_u \,+\, J_t \,, \tag{2}$$

4

where a is a locally bounded and predictable process, σ is an adapted cadlag process bounded away from zero, and J is a finite activity jump process.

Assumption 1A The volatility process, σ_t, follows a jump-diffusive semimartingale,

$$\sigma_t \,=\, \sigma_0 \,+\, \int_0^t \tilde{a}_u \, du \,+\, \int_0^t \tilde{\sigma}_{u_-} \, dB_u \,+\, \int_0^t \tilde{v}_{u_-} \, dW_u \,+\, \tilde{J}_t,$$

where $\tilde{a}$ is locally bounded and predictable, $\tilde{\sigma}$, $\tilde{v}$ are cadlag, the Brownian motions B, W are uncorrelated, and $\tilde{J}$ is a finite activity jump process.

If the Brownian component in Assumption 1 is non-zero, the return innovation is an order of magnitude larger than the expected return over short time intervals, implying that the drift term typically does not affect the asymptotic distribution of power variation estimators based on high-frequency data. Hence, we ignore the drift term in this section.[6]

Another key implication of Assumptions 1 and 1A is that we may derive the asymptotic properties of many relevant estimators assuming that the intraday returns are locally Gaussian. To operationalize this approach, the trading period is broken into K smaller blocks. For each block, we treat volatility as constant, even if the actual return variation evolves stochastically and the price path contains finite activity jumps. If N equally-spaced continuously compounded returns are available, and each block contains $m \geq 1$ returns, we assume, without loss of generality, that $N = K \cdot m$. Notice that each block covers $\frac{1}{K}^{\text{th}}$ of the trading period and each return reflects the price evolution over an interval length of $\frac{1}{N}$.

The above insight simplifies matters greatly, as nonparametric jump-robust estimators now are easy to devise. One simply selects a suitable unbiased estimator for the (power of) volatility within each block under the null hypothesis of i.i.d. Gaussian returns, and then cumulate the estimates across blocks to obtain the overall power variation. The distribution theory is developed using standard in-fill asymptotics, letting N grow indefinitely, while requiring $\frac{m}{K} \to 0$. In most cases, m is fixed and K diverges proportionally with N.

2.2   Estimating Power Variation under the Diffusive Null Hypothesis

We first consider the case where there is no jump component. Given the assumptions invoked above, we focus on the null hypothesis that the returns within a small block are i.i.d. Gaussian. A generic estimator of the pth order return variation, for p an even positive integer, is now obtained as the average of local estimates of σ^p based on a functional f operating on blocks of m adjacent returns. For each integer i, $1 \leq i \leq N - m + 1$, we have a return block, $\underline{r}_{i,m} = (r_i, r_{i+1}, \ldots, r_{i+m-1})$. Under the null, these returns are i.i.d. $\mathcal{N}(0, \sigma_i^2 / N)$. We let $f_i = f\left(\underline{r}_{i,m}\right)$ denote the functional exploited by a given estimator to obtain an unbiased estimate of σ_i^p for the i'th block. If Assumption 1 holds, the power

---

[6]Of course, the proofs for all our new results, provided in Appendix A, allow for a general drift term.

variation estimator is consistent. Heuristically, the law of large numbers implies, as $N \to \infty$,

$$\frac{1}{N-m+1} \sum_{i=1}^{N-m+1} f_i \approx \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} \sigma_i^p \to \int_0^1 \sigma_u^p \, du \,.$$

A corresponding central limit theory may typically be devised if we invoke Assumption 1A.

The simplest estimator within this framework is the realized Power Variation (PV) measure. It does not exploit multi-return blocks, so $m = 1$. It takes the form,

$$f_i^{PV}(p; N) = f(r_i) = \mu_p^{-1} \, N^{\frac{p}{2}} \, r_i^p \,, \quad \text{where} \quad \mu_p = E|Z|^p \,, \quad \text{for} \quad Z \sim \mathcal{N}(0,1) \,.$$

The normalization constant is given by $\mu_p = 2^{p/2} \frac{\Gamma((p+1)/2)}{\Gamma(1/2)}$, for any $p > 0$.

For $p = 2$, this produces the regular RV, or $\mathrm{PV}_N(2)$, estimator with $\mu_2 = 1$, while $p = 4$ yields the RQ, or $\mathrm{PV}_N(4)$, estimator from Section 1 with normalizing constant $\mu_4 = 3$.

A couple of comments are warranted. First, the setting ignores data errors and market microstructure frictions. Higher order return moments are particularly sensitive to faulty price observations or inappropriate assumptions regarding the evolution of the high frequency returns. We discuss these issues in the context of the simulation and empirical sections below. Second, in contrast to the realized power variation estimator, the functional f_i will in the following be designed to be jump-robust, i.e., provide valid asymptotic inference for the power variation, even in the presence of finite activity jumps. However, jumps often have a severe adverse effect on the finite sample properties of the estimators, especially for $p > 2$. Many of the practical complications below arise from this feature.

2.3 Jump-Robust Power Variation Estimation

We now outline the basic principles behind the construction of power variation estimators that are robust to the presence of finite activity jump processes. Asymptotically, as the block sizes shrink towards zero and the number of blocks grows indefinitely, there will be a finite number of blocks containing one single jump each. Hence, in the limit, the power variation associated with the blocks containing jumps is negligible. It follows that the power variation can be estimated consistently as long as the contribution from the "jump blocks" is an order of magnitude less than the overall power variation measure which, of course, is $O(1)$. However, the jumps are also of order $O(1)$, so the functional f must ensure that the jumps are dampened sufficiently to eliminate their impact asymptotically.

Formally, for any given sampling frequency, we denote the set of indices corresponding to returns for which the associated block contains a jump by I_J. Thus, for $i \in I_J$, there is a jump in the return block $\underline{r}_{i,m}$. We then write the generic power variation estimator as,

$$\frac{1}{N-m+1} \sum_{i=1}^{N-m+1} f_i = \frac{1}{N-m+1} \left(\sum_{i=1, i \notin I_J}^{N-m+1} f_i + \sum_{i \in I_J} f_i \right) .$$

The first term estimates the integrated power variation consistently, i.e.,

$$\frac{1}{N-m+1} \sum_{i=1,\, i\notin I_J}^{N-m+1} f_i \longrightarrow \int_0^1 \sigma_u^p \, du, \quad \text{as } N \longrightarrow \infty.$$

The contribution from the blocks containing jumps is negligible, in the limit, only if each such block is of order less than $O(1)$. Thus, the associated power variation estimator is consistent as long as $f_i \sim o(N)$, for $i \in I_J$.[7] This is accomplished in different ways by alternative jump-robust estimators. Moreover, their practical effectiveness is largely determined by the degree to which they accomplish sufficient dampening of the jump contributions in finite samples.

2.4 Alternative Jump-Robust Power Variation Estimators

2.4.1 Multi-Power Variation Estimators

The first (finite activity) jump-robust power variation estimators were the Realized Multi-Power Variation (MPV) statistics, inspired by BNS (2002). Expressed in terms of the functional applied to successive return blocks, the estimator takes the form,

$$f_i^{MPV}(m,p\,;N) \;=\; f\left(\underline{r}_{i,m}\right) \;=\; \left(\mu_{p/m}\right)^{-m} \; N^{\frac{p}{2}} \prod_{j=0}^{m-1} |r_{i+j}|^{\frac{p}{m}}.$$

For $m=1$ and $p=2$ or $p=4$, the estimator reduces to the (non jump-robust) RV or RQ estimator, respectively. Prominent (jump robust) special cases include $(m,p)=(2,2)$, which defines the bipower variation statistic, and various IQ estimators, such as tripower $(m,p)=(3,4)$, quadpower $(m,p)=(4,4)$, and quintpower $(m,p)=(5,4)$.

As described earlier, the actual estimator is now obtained by averaging the value of the functional across the available blocks,

$$MPV_N(m,p) \;=\; \overline{f}^{MPV}(m,p\,;N) \;=\; \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} f_i^{MPV}(m,p\,;N).$$

The MPV estimator is consistent and affords an associated CLT, as long as m is chosen sufficiently large relative to p. This produces an inevitable bias-variance trade-off. A larger m implies more dampening of the jump term, so the finite sample bias induced by the jump is alleviated. On the other hand, for a given sampling frequency, a larger block size, m, implies that the functional is less localized, so the constant volatility assumption provides a poorer approximation, and the estimator becomes less efficient.

---

[7] Naturally, additional dampening is required to establish a CLT for the power variation measures. In this case, $f_i \sim o(N^{\frac{1}{2}})$, for $i \in I_J$

2.4.2 Truncated Power Variation Estimators

An estimator closely related to the PV and MPV statistics is the Realized Truncated Power Variation (TPV) measure. Mancini (2009) introduces the *threshold* realized volatility and quarticity estimators, while Corsi, Pirino and Reno (2010), henceforth CPR, consider a bipower variant of these statistics. These estimators achieve jump robustness by truncating observations exceeding a pre-specified threshold. Under in-fill asymptotics, we may stipulate that the threshold converges toward zero slowly enough (slower than $\frac{1}{N}\log(N)$) that the limiting distribution of the resulting estimators is identical to their non-truncated counterparts. In particular, Truncated RV, or TRV, is asymptotically most efficient among all jump robust IV estimators, and similarly the Truncated RQ, or TRQ, is the most efficient jump-robust estimator for IQ. Moreover, it is evident that the (finite sample) jump distortion is determined by the size of the truncation threshold and thus is under direct control in designing the estimator. The block-functional defining the truncation multi-power variation estimator of order m with truncation threshold, $\vartheta_{i-j} > 0$, takes the form,

$$f_i^{TPV}(m, p\,; N) \;=\; \left(\mu_{p/m}\right)^{-m} \; N^{\frac{p}{2}} \prod_{j=0}^{m-1} |r_{i+j}|^{\frac{p}{m}} \cdot \mathbf{1}_{\{r_{i+j}^2 \,\le\, \vartheta_{i-j}\}}\,,$$

where $\mathbf{1}_{\{A\}}$ is an indicator function, taking the value of one if the statement A is true, and zero otherwise. As before, the actual TPV_N(m,p) estimator is obtained by averaging the functional values across the available return blocks for the trading period.

The choice of threshold can be delicate. It is beneficial to truncate aggressively to reduce the jump distortion by choosing a low threshold, but the non-jump returns are then also truncated with non-trivial probability. CPR suggest a finite-sample scaling to correct for this bias. They develop an iterative scheme aiming to obtain a fixed point at which the expectation of the truncated estimator equals the true (estimated) volatility under the null hypothesis. The approach is conceptually appealing, but has drawbacks. First, the modified estimator is no longer linear in the unobserved σ^2 and thus suffers from a downwards bias, due to Jensen's inequality, even in the ideal Brownian case. Second, they use a sizeable two-sided window (e.g., 50 observations) to obtain a local volatility estimate, thereby rendering it susceptible to an additional bias due to time variation in volatility across the block.

2.4.3 Neighborhood Truncation Estimators

Andersen, Dobrev and Schaumburg, henceforth ADS, (2012) introduce a couple of IV estimators, MinRV and MedRV, designed to improve on the trade-off between jump robustness and efficiency confronting the MPV estimators. MinRV and MedRV are based on an endogenous "nearest neighbor" truncation which is particularly helpful in alleviating the finite-sample impact of isolated large jumps. We now extend this theory to cover general power variation estimation. We start out by introducing notation that allows us to identify various order statistics associated with a given return block.

First, we denote the ith block, consisting of absolute returns raised to the pth power, by

$|\underline{r}_{i,m}^{p}| = (\,|r_i|^p,\,\cdots,\,|r_{i+m-1}|^p\,)$, $i = 1,\ldots,N-m+1$. Next, $q_j\left(|\underline{r}_{i,m}^{p}|\right)$, $j = 1,\cdots,m$, indicates the jth order statistic of the block $|\underline{r}_{i,m}^{p}|$, so $q_1(|\underline{r}_{i,m}^{p}|) \leq \ldots \leq q_m(|\underline{r}_{i,m}^{p}|)$. As the returns are assumed $i.i.d.\,\mathcal{N}(\,0,\,\sigma_i^2/N\,)$, we may also write $N^{\frac{p}{2}} \cdot |\underline{r}_{i,m}^{p}| = \sigma_i^{p} \cdot |\underline{Z}_{i,m}^{p}|$, highlighting the fact that all estimators for the block ultimately are functionals operating on the realization of an m-dimensional standard normal random vector.

We now readily obtain m separate unbiased estimators for σ_i^{p}, namely one for each order statistic. We denote these "Neighborhood Truncation" estimators, or $\mathrm{NT}_N(\mathrm{j,m,p})$. As before, we construct them by averaging the appropriate block functional across the trading period. The functional takes the form,

$$ f_i^{NT}(j,m,p\,;N) \;=\; \left(\mu_p^{(j,m)}\right)^{-1}\, N^{\frac{p}{2}}\, q_j\left(|\underline{r}_{i,m}^{p}|\right)\,, \quad j = 1,\,\ldots,\,m\,, $$

where $\mu_p^{(j,m)} = E\left[q_j\left(|Z_1|^p,\,\cdots,\,|Z_m|^p\right)\right]$, $Z_i \sim i.i.d.\,\mathcal{N}(0,1)$, $i = 1,\,\ldots,m$. This normalization ensures that the functional provides an unbiased estimator for σ_i^{p}. Since the scaling factors are inversely related to the expected value of the order statistics, we have the ranking, $\mu_p^{(1,m)} > \ldots > \mu_p^{(m,m)}$. For high values of p and m, these factors become quite large for the lower order statistics, while they are very small for the higher order statistics.[8]

The class of neighborhood truncation estimators generalize the MinRV and MedRV estimators, as we have, $MinRV = NT_N(1,2,2))$ and $MedRV = NT_N(2,3,2)$.[9] Under appropriate conditions the NT estimators are consistent and afford a CLT. We confirm this result when we introduce an even wider class of robust estimators in Section 3.

2.4.4 Combining Power Variation Estimators

It seems natural to combine some of the estimators introduced above to obtain superior asymptotic properties and, possibly, improved finite-sample performance. It is, however, outside the scope of the current paper to pursue this topic in depth. Nonetheless, we do develop the framework and notation to accommodate such combination estimators, as it is useful for our introduction of a new class of robust estimators in the next section.

Assume we have a candidate set of K separate jump-robust power variation estimators which all are unbiased, consistent and afford a CLT under the local Gaussian null hypothesis. We denote this set of estimators, $\mathcal{E} = \{\,\mathcal{E}_1,\,\ldots,\,\mathcal{E}_K\,\}$. Almost trivially, it is then, in theory, feasible to improve the performance of any single estimator by combining it with others.[10]

---

[8] For IQ estimation based on a block with five returns, we have $\mu_4^{(1,5)} = 35.0$, $\mu_4^{(2,5)} = 5.74$, $\mu_4^{(3,5)} = 1.44$, $\mu_4^{(4,5)} = 0.40$, $and\,\mu_4^{(5,5)} = 0.086$. Hence, the largest normalization constant exceeds the smallest one by a factor of several hundred; see Table 1 for additional information.

[9] For the integrated variance, these estimators may be interpreted as a redesigned version of the original Quantile Realized Variation, or QRV, estimator of Christensen, Oomen and Podolskij (COP) (2008), where it is applied to quantiles based on absolute rather than raw returns. The adaptation of QRV to cover the ADS (2012) nearest neighbor truncation estimators is discussed in COP (2010). This was also previously proposed during a conference discussion of COP (2010) by Kevin Sheppard, see also the comments in the realized_quantile_variance function in his Oxford MFE Toolbox.

[10] In practice, finite sample complications can render the procedure less successful. One may want to avoid estimators that are highly sensitive to microstructure noise or jumps. Likewise, high correlation across estimators can generate optimal combinations that are extreme which may induce a degree of instability.

We formalize the selection of a subset of the estimators by introducing a "selection" vector, identifying the elements in $\mathcal{E}$ used to construct a given (combination) estimator. Hence, we let the 1xH vector $\boldsymbol{I} = (k_1, \ldots k_H)$, $1 \leq H \leq K$, consisting of an ordered subset of integers from $\{1, \cdots, K\}$, indicate that the combination estimator is based on the set $\mathcal{E}_I = (\mathcal{E}_{k_1}, \cdots, \mathcal{E}_{k_H})$. Denoting the set of all possible selector vectors $\mathcal{I}$, any subset is now uniquely identified by $\boldsymbol{I} \in \mathcal{I}$, where $\boldsymbol{I}$ ranges from a scalar (using a single estimator), at the one extreme, to the full vector $\{1, \cdots, K\}$ (using them all), at the other extreme.

A natural way to preserve the desirable properties of the individual estimators is to exploit linear combinations with non-negative weights that sum to unity. For example, focusing exclusively on a set of NT estimators constructed using a return block of size m, we have $\mathcal{E} = \{\mathcal{E}_1, \ldots, \mathcal{E}_m\}$, where each element represents an unbiased estimator based on the corresponding (absolute return) order statistic. Picking a specific NT estimator amounts to, a priori, selecting a given integer $j \in \{1, \cdots, m\}$. There are a total of $2^m - 1$ distinct non-empty subsets of $\mathcal{E}$ from which to construct a combination estimator. It is a routine exercise to extend our asymptotic results to cover the case of any such linear combination of NT estimators.[11] Conceptually, it is likewise straightforward to derive corresponding results for linear combinations involving alternative types of jump-robust estimators.

In summary, within the ideal setting of Assumptions 1 and 1A, superior asymptotic performance can be obtained by combining the information associated with all available estimators. However, this must be weighed against the robustness objective of ensuring reliable finite sample inference in the presence of jumps as well as other potential sources of noise. Such robustness concerns motivate the introduction of an even broader class of combination estimators in the following section, obtained by nonlinearly combining suitably chosen unbiased estimators via a second layer of order statistics.

3 Robust Neighborhood Truncation Estimation

Our major objective is to develop a reliable jump-robust procedure for estimating measures associated with the integrated quarticity. Most of the estimators reviewed in the previous section were developed for IV, even if they can be adapted for higher order power variation measures. It is worth recognizing that the relative importance of factors impacting the trade-off between statistical efficiency and robustness changes substantially as we estimate higher order power variation measures. In fact, our simulation evidence demonstrates, quite strikingly, that the most suitable approach for IV estimation is unlikely to be preferable when estimating IQ. Consequently, we now introduce a novel inference procedure which enhances the robustness to common sources of finite-sample distortions and allows for a great deal of flexibility in implementation so that the estimator can be tailored to the specific features of a given return series and market environment.

---

[11]This is analogous to the use of linear combinations for the QRV estimator explored by COP (2008).

3.1   Theory

This section proposes estimating integrated power variation via a *nonlinear* combination of existing unbiased estimators, obtained by invoking an additional layer of order statistics. We develop the theory for neighborhood truncation estimators, but the principles apply for any set of unbiased estimators. The emphasis is on finite sample robustness to microstructure noise and jumps, so we label them "Robust Neighborhood Truncation," or RNT, estimators.

For a return block of size m, there are m distinct NT estimators, namely one for each order statistic. One may combine any subset of these to produce an estimator that exploits more sampling information than can be utilized by any individual one. The set of alternative selections is the set, $\mathcal{I}$, of non-empty subsets of $\{1, \cdots, m\}$. A specific choice is given by $\boldsymbol{I} = (k_1, \ldots k_H)$, $1 \leq H \leq m$. The corresponding NT estimators are defined via the functionals they apply to the underlying return blocks. To facilitate the exposition, we use the following short-hand notation for these functionals, applied to the ith return block,

$$\underline{f}^{NT}_{i,\boldsymbol{I}}(m, p\,; N) = \left(f^{NT}_i(k_1, m, p\,; N), \ldots, f^{NT}_i(k_H, m, p\,; N) \right).$$

The rationale behind the NT estimators is to alleviate the impact of extreme returns – large or small – which may be incompatible with the i.i.d. Gaussian assumption. The robust neighborhood truncation principle takes the reasoning one step further by producing an estimator for σ^p_i based on a suitable order statistic among the subset of selected NT estimators. Formally, we have,

$$f^{RNT}_i(\boldsymbol{I}; j, m, p\,; N) = \left(\mu^{(j,\boldsymbol{I})}_p \right)^{-1} q_j \left(\underline{f}^{NT}_{i,\boldsymbol{I}}(m, p\,; N) \right), \quad j = 1, \ldots, H,$$

where $\mu^{(j,\boldsymbol{I})}_p = E\left[q_j \left(|Z_1|^p, \cdots, |Z_H|^p \right) \right]$, $Z_h \sim i.i.d. \, \mathcal{N}(0,1)$, $h = 1, \ldots, H$.

Normalization is required, even if each NT estimator is individually unbiased, because selection *conditional* on observed realizations induces a bias. This is corrected by scaling with (the inverse of) the expected value of the corresponding order statistic for a standard normal Hx1 vector. This normalization factor is not available in closed form, but may be determined, to any degree of accuracy, by numerical integration or simulation.

As before, the actual estimator is obtained by averaging the estimates across all blocks in the trading periods, so we have,

$$RNT_N(\boldsymbol{I}; j, m, p) = \overline{f}^{RNT}(\boldsymbol{I}; j, m, p\,; N) = \frac{1}{N-m+1} \sum_{i=1}^{N-m+1} f^{RNT}_i(\boldsymbol{I}; j, m, p\,; N).$$

Notice that the RNT procedure involves two layers of order statistics: we first construct consistent NT estimators from the order statistics of a block of absolute returns (which readily may be extended to any set of consistent estimators), and then obtain the RNT estimator from another order statistic applied to a subset of these NT estimators. This provides a great deal of flexibility in alleviating the impact of extreme returns. In line with the logic behind the MinRV and MedRV measures, the RNT estimator is consistent

if we exclude the largest order statistic from *I*. Asymptotically, this ensures that none of the NT estimators are generated from a (scaled) jump return. Alternatively, this is also guaranteed if we avoid constructing the RNT estimator from the largest realization of the NT estimators, i.e., $j < H$ in the second step.

Proposition 1 *Let a family of H distinct NT estimators, indexed by* I*, be generated from absolute return blocks of size m. The largest order statistic used for constructing any of these NT estimators is denoted k_H, $k_H \leq m$. Next, consider the RNT estimator obtained from the jth order statistic applied to this family of NT estimators, $j \leq H$.*

If (i) Assumption 1 holds; (ii) $k_H < m$ and/or $j < H$; and (iii) p is a positive, even integer; then, as $N \to \infty$,

$$RNT_N(\boldsymbol{I}; j, m, p) \xrightarrow{P} \int_0^1 \sigma_u^p \, du \, .$$

If, in addition, Assumptions 1A applies, we obtain, for $\eta(\boldsymbol{I}; j, m, p)$ a known constant,

$$\sqrt{N} \left(RNT_N(\boldsymbol{I}; j, m, p) - \int_0^1 \sigma_u^p \, du \right) \xrightarrow{Stable \, \mathcal{L}} \mathcal{N} \left(0, \eta(\boldsymbol{I}; j, m, p) \int_0^1 \sigma_u^{2p} \, du \right).$$

The proposition warrants a few comments. First, the distributional convergence is stable with a mixed Gaussian limit, i.e., a normal distribution conditional on the realization of the integrated power variation, $PV(q) = \int_0^1 \sigma_u^q \, du$, where, importantly, the limiting normal variate is independent of the (random) power variation process.[12] Second, the convergence result is qualitatively similar to those established for existing power variation estimators, with the "efficiency factor" $\eta(\boldsymbol{I}; j, m, p)$ determining the relative asymptotic efficiency of the estimator. Third, the main objective is not efficiency per se, but good performance along with (finite sample) robustness to jumps, noise, and other data irregularities. Fourth, the results apply for the Neighborhood Truncation and Nearest Neighbor Truncation estimators, as these constitute special cases involving a particular choice for the vector $\boldsymbol{I}$. Fifth, the results are likely to extend to the infinite activity jump case, given suitably tight constraint on the size of the associated (jump) activity index.[13]

3.2 Illustration

The simulation and empirical work in the following sections exploit fairly small blocks of $m \leq 5$ in order to retain resiliency relative to rapidly changing volatility levels during the trading day. In addition, we find it useful to eliminate estimators that stem from the lowest

---

[12]See, e.g., Jacod and Shiryaev (2003), Chapter IX, for an introduction to stable convergence.

[13]We intend to apply these estimators in a setting with microstructure noise. As discussed in Section 4.2, this can be accommodated in the finite activity jump scenario via pre-averaging with a relatively large window size, resulting in a suboptimal convergence rate, but added robustness to noise. The literature on inference in the presence of infinite activity jumps *and* noise is limited. However, the findings for MPV estimators in Podolskij and Vetter (2009) and Hautsch and Podolskij (2010) suggest constructive results are feasible. Establishing formal results for this case falls outside the scope of the present paper.

order statistics of the absolute returns as these are relatively more affected by market microstructure noise such as price discreteness and bid-ask bounce. This is a particular concern, because these estimates are bias-corrected by scaling the original small returns by a large factor, implying that microstructure distortions may be amplified. Likewise, we typically satisfy the formal constraint on the order statistics by picking $j < H$, so we avoid basing the RNT estimator on the largest realization among the relevant NT estimators.

For a 5-dimensional return block, the construction of the $RNT_N((\mathbf{3,4,5}); 1, 5, 4)$ and $RNT_N((\mathbf{3,4,5}); 2, 5, 4)$ estimators are exemplified in Figure 2. The notation becomes quite involved so, for brevity, we refer to the two estimators in the diagram as RMinRQ and RMedRQ, respectively.[14] Both play a significant role in our subsequent exposition. For these estimators, the two smallest absolute returns are discarded, while the remaining three are used to compute the corresponding NT estimators. Among those, we pick the lowest, respectively median, realization and scale it to construct the associated RNT estimator.

[Insert Figure 2 about here.]

A few features are worth emphasizing. First, we display the NT estimators obtained from the first two order statistics along with the rest, even if they are excluded from the construction of the estimator. Hence, the third box displays all five returns taken to the fourth power. An extreme right skew is evident, with values spanning 0 to 915.1, even if the initial returns are not particularly scattered. A zero return is, of course, common due to the discreteness of the price grid. Second, between box 3 and 4 we apply the relevant scaling factors for the NT estimators, see Panel B of Table 1. Strikingly, the large factor $(5.74 = (0.1741)^{-1})$ for the second order statistic produces, by far, the largest realized estimator in box 4 (466.2). Finally, excluding the NT estimators originating from the two smallest absolute returns (0, 466.2), we pick the minimum and median of the remainder and scale these statistics (78.9 and 163.3) suitably with the scaling factors provided in Panel D of Table 1 ($2.611 = (0.38303)^{-1}$ and $1.214 = (0.82367)^{-1}$) to obtain the local RNT estimators for σ_i^4 of 206.1 given by RMinRQ, respectively 198.3 given by RMedRQ. These realizations happen to stem from the two largest order statistic of the original return block (5.5 and 4.5), but the low scaling factors for these order statistics ($0.086 = (11.59249)^{-1}$ and $0.398 = (2.51102)^{-1}$) imply that the associated (unbiased) NT estimators in box 4 are the smallest among the relevant subset. It reflects the relatively low spread between the three largest absolute return realizations of 4.0, 4.5, and 5.5. In general, these procedures tend to moderate the local estimates relative to estimators which rely more directly on the raw fourth powers in box 3.

[Insert Table 1 about here.]

---

[14]In this notation, the initial "R" references the RNT estimator, "Min" ("Med") signals that we exploit the smallest (median) realization of the underlying NT estimators, and "RQ" ("RV") indicates realized quarticity (realized variance) estimation, that is $p = 4$ ($p = 2$). It remains implicit that we take $m = 5$, but retain only the three largest absolute returns in constructing the estimator.

4 Robustification towards Noise and Errors

In estimating higher order return power variation measures, we deal with procedures that can be highly sensitive to erroneous outliers as well as the presence of noise. Hence, we adopt various techniques that mitigate the impact of such features on the inference. Our strategy includes standard pre-filtering for obvious data errors, pre-averaging to reduce the magnitude of the noise in the returns, and conducting inference on the ratio of IQ versus IV^2 rather than directly for IQ. However, most inference techniques continue to display excessive sensitivity to data irregularities. Consequently, we supplement the above steps with a novel filtering method, specifically designed for robust power variation estimators operating on return blocks. This section reviews the techniques we employ to enhance the robustness of our inference towards data errors and noise.

4.1 Eliminating Obvious Errors in the Tick-by-Tick Data

Any large set of raw transactions data is invariably subject to recording errors that infuse noise into the high-frequency returns. Most dramatically, faulty prices create artificial outliers, causing so-called "bounce-backs" in returns, as there is a "jump" both when the flawed price first appears and later, often shortly thereafter, when the price reverts to the correct level. Hence, the need for effective cleaning procedures has long been acknowledged. BNHLS (2009) lay out a systematic framework for dealing with trade data from NYSE-TAQ. In their terminology, we apply the filters P1-P3 and T1-T4.[15] These filters are arguably mild and uncontroversial and simply aim to eliminate obvious data errors.

4.2 Pre-Averaging

The assumption that (observed) high-frequency returns embody a diffusive component is systematically violated at the tick-by-tick level due to various market microstructure features, including the finite price grid and the bid-ask spread. As a result, tick-by-tick price changes are often an order of magnitude larger than what is consistent with a diffusive characterization. One effective approach to mitigating the impact of such noise is to apply pre-averaging, as originally suggested by Podolskij and Vetter (2009a). This is achieved by transforming the noisy observations on ultra high-frequency returns into a smaller set of kernel-averaged, and thus less erratic, "smoothed" returns. In particular, each of the m returns within a block are obtained via kernel-averaging based on separate, non-overlapping subsets of tick-by-tick returns. The benefit is a reduction in the impact of idiosyncratic noise and, especially, distortions induced by bounce-backs. The drawback is a substantial drop in the underlying sampling frequency. The latter impacts the choice of the window

---

[15]P1: Retain only observations with time stamps between 9:30am and 4:00pm. P2: Retain only trades with positive prices. P3: Retain only trades originating from the main exchange (NYSE for all stocks except MSFT and INTL for which it is NASDAQ). T1: Delete entries with corrected trades. T2: Delete entries with abnormal sale condition. T3: If multiple trades occurred with the same time stamp, use the median price. T4: Delete entries with prices above the ask (or under the bid) by more than one bid-ask spread.

width, m, as the (diffusive) volatility fluctuates more widely across longer blocks.

Our implementation of pre-averaging, detailed in Appendix C, is based on a relatively conservative choice of sampling frequency. This has the effect of emphasizing noise robustness over efficiency. Importantly, it also simplifies the analysis, as the impact of noise may be largely ignored in the asymptotic theory. First, the pre-averaging estimator has an asymptotic bias, but if the (pre-averaged) returns are not sampled at very high frequencies, the bias is, effectively, negligible. Second, if there are N original high-frequency returns, the optimal convergence rate for pre-averaged estimators in the presence of noise is typically $N^{1/4}$. The associated asymptotic variance reflects both the sampling variance of the true returns and the noise variance. This result is obtained if the number of returns per pre-averaging block, $2 \cdot K$, grows at the asymptotic rate $N^{1/2}$, so that the total number of pre-averaged returns without overlap, n, also is of order $N^{1/2}$. This allows the convergence rate – as usual – to equal the square-root of the number of (pre-averaged) returns, i.e., $\sqrt{n} = N^{1/4}$. But if K is larger, asymptotically rising at the rate N^q for $\frac{1}{2} < q < 1$, the number of pre-averaged returns grows more slowly, $n \sim N^{1-q}$, implying a convergence rate of $\sqrt{n} = N^{\frac{1-q}{2}}$, e.g., $N^{\frac{1}{5}}$ for $q = \frac{3}{5}$. At the same time, the noise will be averaged more aggressively and vanishes asymptotically at a faster rate. The bottom line is that, by appealing to a slower asymptotic convergence rate relative to the number of original high-frequency returns, the (asymptotic) efficiency is lower, but the asymptotic variance of the pre-averaging estimator becomes identical to the one for the no-noise case with n returns. However, this equivalence holds only for pre-averaged return series based on non-overlapping blocks without sub-sampling. The additional efficiency gain attainable by sub-sampling, as in Appendix C, is not identical with and without pre-averaging, differs from one estimator to another, and generally is not known in closed form.[16] Nonetheless, the efficiency of each pre-averaged and sub-sampled estimator in the presence of noise is very close to its efficiency in the absence of noise, as long as the pre-averaging window size is sufficiently large relative to sample size. We monitor the latter prediction in the simulations below to verify that it provides a useful characterization of the relevant features of the finite sample distribution.

In summary, we appeal to an asymptotic theory guided by a slightly *slower* convergence rate than the "optimal" $N^{1/4}$ for pre-averaged estimators. This enhances robustness to noise while allowing the asymptotic theory for the no-noise case to be the relevant benchmark. In practice, we choose a relatively large return block so our procedure is compatible with the theoretical setting in this regard. This has the convenient implication that the theory in Sections 2 and 3 provides the appropriate basis for assessing the limiting behavior of our estimators computed from pre-averaged returns, even if it ignores the presence of noise. Thus, henceforth, we simply treat the pre-averaged returns as if they were the original raw returns and, with slight abuse of notation, we redefine N to denote the relevant number of

---

[16]Analogous to the scaling factors induced by the pre-averaging kernel, the efficiency with subsampling depends on the pre-averaging scheme as well as the length of the pre-averaging window relative to sample size. Hence, applications involving subsampling must develop a suitable estimate of the terms involving the efficiency factor η in Section 3. One feasible approach is simulation.

(pre-averaged) returns, while accommodating the effect of sub-sampling in the conventional fashion.

4.3 Filtering via Truncation of Return Functionals

Even for returns based on pre-averaged tick data and sampled at moderate frequencies, microstructure features and other data irregularities may induce inhomogeneous and serially correlated observations that blatantly violate our distributional assumptions. For quarticity estimation, in particular, it is paramount to control the impact of this type of data imperfections to achieve a beneficial trade-off between robustness and efficiency.

This section briefly outlines a general truncation principle for return functionals that enhances the robustness of integrated power variation estimators operating on return blocks. It provides an extension of existing techniques that employ truncation to alleviate the impact of jumps or data errors. However, the philosophy and implementation are very different. Existing procedures truncate returns based on whether a single observation constitutes a significant outlier under the local Brownian null. Moreover, the truncation is an essential step in rendering the estimator robust as it dampens, and asymptotically eliminates, the distortion induced by price jumps on the estimated power variation. For this to be effective, the detection of larger jumps must be reliable, and it is common to apply a threshold for jumps that correspond to "three sigmas" or a p-value of about 0.3%. As a result, the procedure generates a non-trivial incidence of type I errors because diffusive returns based on high-frequency return data inevitably are subjected to unwarranted truncation.[17]

In contrast, we develop a filtering procedure that operates directly on the jump-robust functional and more broadly alleviates distortions induced by deviations from the null that the block consists of i.i.d. draws from a normal distribution. In this scenario, jump robustness is, in principle, already assured by the choice of an appropriate functional. Hence, the filtering is merely intended to eliminate truly excessive *ex post* estimates of local power variation, driven by functional values incompatible with the maintained null hypothesis. As such, we rely on an extremely conservative threshold for truncation, typically with p-values around 10^{-6} or below. This is sufficiently low that we expect, under the null hypothesis, to truncate less than a single realization of the return functional across our entire sample. In practice, the underlying assumptions are violated and truncation occurs with non-trivial frequency which helps control the associated distortion in the power variation estimators.

To introduce this filtering procedure, we recall that $f_i = f(\underline{r}_{i,m})$ is a functional providing an unbiased estimator of the local power variation, σ_i^p under the null hypothesis. Next, for a sufficiently small α, e.g., 10^{-6}, we let $Q^{(1-\alpha)}[g]$ denote the $(1-\alpha)^{\text{th}}$-quantile of the distribution of a random variable g. We then define the corresponding truncated functional

---

[17]Recognizing the downward bias, CPR develop an iterative rescaling procedure to alleviate this effect.

$f_i^{(1-\alpha)}$,

$$f_i^{(1-\alpha)} = \begin{cases} f_i & \text{if} \quad f_i\,\sigma_i^{-p} \leq Q^{(1-\alpha)}[\,f_i \cdot \sigma_i^{-p}\,] \\ 0 & \text{else.} \end{cases}$$

Accordingly, the realized truncated estimator based on $f^{(1-\alpha)}$ is given by,

$$\frac{1}{\sum_{i=1}^{N-m+1} \boldsymbol{I}\{\,f_i \cdot \sigma^{-p} < Q^{(1-\alpha)}[\,f_i \cdot \sigma^{-p}]\}} \quad \sum_{i=1}^{N-m+1} f_i^{(1-\alpha)}\,,$$

where $\boldsymbol{I}\{\boldsymbol{A}\}$ equals one if the expression $\boldsymbol{A}$ is true and zero otherwise. Setting $\alpha = 0$ we obtain the usual realized estimator based on f without truncation. Moreover, if $m = 1$ our functional filtering is equivalent to the usual return filtering at the significance level α.

A feasible version of the filter is developed in Andersen, Dobrev and Schaumburg (2011). The procedure exploits a local estimate of volatility based on preceding observations to provide the appropriate truncation level – exactly as done for the standard truncation RV estimator – while simulation is performed to obtain the critical values, taking into account the presence of estimation error for local volatility.

4.4 Using the Ratio $\sqrt{\widehat{IQ}}\,/\,\widehat{IV}$ for Robust IV Inference

The primary applications of IQ estimation is to draw inference about IV and to test for jumps under the null hypothesis of no jumps. For these procedures to perform well, it is essential that the IQ estimator has good efficiency and finite sample jump robustness.

Let $\widehat{IV}$, $\widehat{IQ}$ be suitable jump-robust estimators of IV and IQ. A natural approach for drawing inference about IV follows directly from its limiting distribution,

$$\frac{\sqrt{N}\left[\widehat{IV} - IV\right]}{\sqrt{\eta\,\widehat{IQ}}} \quad \overset{\text{Stable}\,\mathcal{L}}{\to} \quad \mathcal{N}(0,\,1)\,,$$

where the "efficiency" factor, η, depends on the specific choice of estimator.

Letting RV denote the realized volatility estimator, which is the efficient estimator of IV under the null, the natural Hausman test statistic for the presence of jumps, see BNS (2004) and Huang and Tauchen (2005), is given by

$$\frac{\sqrt{N}\left[\widehat{RV} - \widehat{IV}\right]}{\sqrt{(\eta - 2)\,\widehat{IQ}}} \quad \overset{\text{Stable}\,\mathcal{L}}{\to} \quad \mathcal{N}(0,\,1)\,.$$

An asymptotically equivalent set of test statistics with better finite sample properties, proposed by BNS (2002), can be derived by applying the delta method to the log-transform of the volatility measures. This has the benefit that IQ enters only in terms of the ratio $\sqrt{IQ}/IV$ which, as also demonstrated in our empirical investigation below, has a stabilizing

effect on the variance of log $\widehat{IV}$.

$$\frac{\sqrt{N}\,[\,\log(\widehat{IV}) - \log(IV)\,]}{\sqrt{\frac{\eta\,\widehat{IQ}}{\widehat{IV}^2}}} \quad \overset{\mathcal{L}}{\underset{\text{Stable}}{\to}} \quad \mathcal{N}(0,\,1)\,.$$

The corresponding Hausman test statistic for the presence of jumps is

$$\frac{\sqrt{N}\,[\,\log(\widehat{RV}) - \log(\widehat{IV})\,]}{\sqrt{(\eta-2)\,\frac{\widehat{IQ}}{\widehat{IV}^2}}} \quad \overset{\mathcal{L}}{\underset{\text{Stable}}{\to}} \quad \mathcal{N}(0,\,1)\,.$$

While the literature has documented superior performance of this ratio for jump-robust inference in a frictionless setting, it is evident that the ratio also will impact the way market microstructure noise affects the inference. In Appendix B, we provide an illustration based on computations involving the non-robust versions of the IQ and IV estimators. The findings point towards favorable properties of the ratio statistic relative to the raw statistic along this dimension as well. The intuition is as before: the realized IV^2 and IQ statistics tend to be impacted by noise in similar ways so the ratio provides a partial cancelation of errors. The issue is further pursued within the simulation set-up entertained in the following section.

5 Finite Sample Simulation Evidence

We design a series of Monte Carlo experiments, each focusing on a distinct feature of the data generating process that may affect the finite sample behavior of the estimators. The emphasis is on the qualitative impact of each feature in isolation. In reality, multiple features interact, creating complex patterns in tick-by-tick data. The joint presence of various factors, partially reinforcing or counteracting each other, render it difficult to infer the significance of individual features. Hence, our simulation design is not intended to replicate the empirical results in all dimensions, but to assist in identifying the features that create differential patterns in the results obtained from alternative estimators.

5.1 Estimators

We adopt the novel filtering procedure, based on (mildly) truncating the local power variation functional, for all estimators except those already truncating individual returns more aggressively, i.e., the truncation power variation estimators. For the latter, the impact of an additional layer of mild truncation is negligible.[18] For the truncation estimators, we follow the approach advocated by CPR, as they document it improves on existing implementations. Overall, we consider the following IV and IQ estimators,

1. **Multi-Power Variation**: We denote $\text{MPV}_N(m,2)$ by **MPVm** and $\text{MPV}_N(m,4)$ by **MPQm**. We include **MPV3** and **MPV5**, as well as **MPQ3** and **MPQ5** in our

---

[18]We confirm that our filtering procedure under the BM null, applied to pre-averaged returns, is active for about 1-in-10^6 return blocks, while the CPR filter is applied to roughly 1-in-10^3 returns on average.

analysis. This type of estimators was introduced by BNS (2004);

2. **Truncated Power Variation**: We use $\text{TRV} = \text{TPV}_N(1,2)$ and $\text{TRQ} = \text{TPV}_N(1,4)$, as well as Truncated Bipower Variation, $\text{TBV} = \text{TPV}_N(2,2)$ and Truncated Bipower Quarticity, $\text{TBQ} = \text{TPV}_N(2,4)$. They are discussed by, e.g., Mancini (2009) and CPR;

3. **Neighborhood Truncation**: We consider $\textbf{MedRV} = \text{NT}_N(2,3,2)$ and the corresponding quarticity estimator $\textbf{MedRQ} = \text{NT}_N(2,3,4)$. They are developed in ADS (2012) and in this paper;

4. **Robust Neighborhood Truncation**: We use $\textbf{RMinRV} = \text{RNT}_N(\mathbf{(3,4,5)};1,5,2)$, $\textbf{RMedRV} = \text{RNT}_N(\mathbf{(3,4,5)};2,5,2)$, $\textbf{RMinRQ} = \text{RNT}_N(\mathbf{(3,4,5)};1,5,4)$, and $\textbf{RMedRQ} = \text{RNT}_N(\mathbf{(3,4,5)};2,5,4)$. They are developed in this paper.

In addition, when investigating IV estimators, we include the standard RV estimator, serving as a non jump-robust benchmark, along with the QRV estimator of COP (2010). We omit an IQ counterpart of QRV from our analysis because we find the block size of 20 or more returns, necessary in order to establish the quantiles, to be prohibitively large for reliable inference on actual data subject to irregular sampling and pronounced intraday variation in volatility. Moreover, as discussed previously, one may interpret our NT estimators as modified QRV estimators based on the *absolute* returns over very small blocks.

Finally, taking into account the need to apply suitable noise-reduction technique when conducting inference in practice, we focus our Monte Carlo analysis exclusively on the pre-averaged implementation of all estimators, as defined in detail in Appendix C, including the efficiency gain from sub-sampling. The use of pre-averaging necessitates a somewhat conservative choice of sampling frequency. We report results based on 30, 120 and 600 second return observations.[19]

5.2 Simulation Results

We largely follow the comprehensive simulation design in ADS (2012) adopted for comparing IV estimators. We calibrate the unconditional daily IV to 0.000159, or roughly 20% per year, across all scenarios. For each scenario we simulate 250,000 trading days, corresponding to about 1,000 years, from 9:30 am to 4:00 pm with new prices arriving every 3 seconds on average, so we have 7,800 distinct prices each day.

We consider three major departures from the Gaussian benchmark: **(i) microstructure noise**: bid-ask bounce, recording errors, irregular trade intervals, and price decimalization (discreteness); **(ii) time-varying volatility**: stochastic and deterministic (diurnal) variation in volatility along with volatility jumps; **(iii) jumps in returns**: one or multiple intraday price jumps. Each scenario is briefly described below, with additional details available in ADS (2012). We focus on estimators of IV, IQ, and the ratio $\sqrt{IQ}/IV$, with the

---

[19]These return intervals span the range that is relevant for our empirical application. Findings based on a more comprehensive set of frequencies are reported in ADS (2011).

latter computed using the same type of estimator for the numerator and denominator.[20] For brevity, we often – including in the tables – refer to the ratio as estimated by, say, $MPQ3$, when it is estimated by $\sqrt{MPQ3}/MPV3$.

5.2.1 The Brownian Motion Benchmark

This is our baseline scenario with sampling on an equispaced time grid. It is an ideal setting in which we expect the finite sample performance of all estimators to closely mimic the underlying asymptotic theory. Panel A of Tables 2 and 3 verify that most of the estimators are unbiased for IV and IQ. The exception is the minor downward bias in TRV, TBV, TRQ and TBQ. This stems from noise in the truncation and bias-correction procedures applied in constructing these estimators. They truncate individual returns at three (estimated) standard deviations, so a scaling is needed to mitigate the impact of erroneous truncation of diffusive returns and this introduces some estimation error, even under the BM null. All other estimators rely on the very conservative truncation level (p-value of 10^{-6}) associated with our functional filtering, described in Section 4.3. As intended, the impact of this filter is negligible under the null hypothesis, so the estimators remain unbiased.

In terms of efficiency, the ordering of the jump-robust estimators is as prescribed by theory, with TRV superior in terms of RMSE for IV (matching RV), followed by RMedRV, QRV, RMinRV, and TBV. For IQ, TRQ is best, followed very closely by RMedRQ, and then RMinRQ, and TBQ at the higher frequencies.[21] In summary, the RNT estimators perform well and, in particular, come close to matching the efficiency of TRQ for IQ. It suggests that this type of estimator can enhance robustness to noise and jumps without any significant loss of efficiency in general.

Finally, turning to the estimates for $\sqrt{IQ}/IV$ in Table 4, Panel A, we notice a small downward bias at the lowest sampling frequency in almost all cases. This is caused by a finite sample Jensen effect.[22] More remarkably, in terms of efficiency, the TRQ estimator now performs relatively poorly. Instead, RMinRQ and RMedRQ outperform the alternatives by a substantial margin. The dramatic shift in relative efficiency reflects the fact that cancelation of outlier terms in the numerator and denominator is particularly effective under the robust neighborhood truncation principle where the largest return realizations are prevented from exerting any significant impact. Likewise, the MPQ5 estimator performs quite well due to the effective dampening of outliers. We conclude that, even under ideal circumstances, the RNT estimators provide an attractive alternative to existing procedures, especially for inference and jump tests based on the $\sqrt{IQ}/IV$ ratio.

[Insert Tables 2, 3, 4 about here.]

---

[20]This choice facilitates effective cancelation of noise and outliers across the numerator and denominator.

[21]Recall, we abandon the QRV style estimators for IQ due to the rapidly deteriorating performance in estimating higher order return variation under realistic market conditions.

[22]The improved performance for TBQ is due to a fortuitous cancelation of the biases of the numerator and denominator under the BM null.

5.2.2 Jumps in Returns

To assess finite sample jump-robustness, we augment the BM model with return jumps of the Poisson-Gaussian type that are independent of volatility and account for 20% of the daily QV (25% of IV). We focus on two cases, one with a single jump per day, the "BM + 1 Jump" scenario, and one with four jumps per day, the "BM + 4 Jumps" scenario, but the overall jump contribution to the daily variance is identical for the two cases.

Panels B and C in Table 2 show that TRV, TBV, QRV, RMinRV and RMedRV provide the best robust IV inference in terms of RMSE. Moreover, the relative performance within this group shifts as we move from a single to four jumps with, in particular, RMinRV and TBV improving their standing as the jump intensity increases. This tendency is even more pronounced for IQ estimation, where panels B and C in Table 3 reveal that RMinRQ is the best performer in both jump scenarios at the 2-minute frequency.

Finally, and very strikingly, panels B and C of Table 4 show that the pairing of RMinRQ and RMinRV dominate all other estimators by a significant margin in terms of estimating the ratio $\sqrt{IQ}/IV$, which governs the precision of log(IV). Thus, from the perspective of finite sample jump-robustness, the RNT estimators seem to offer attractive efficiency improvements, especially for the estimation of IQ and $\sqrt{IQ}/IV$.

Juxtaposing Panel B or C in Tables 3 and 4, we also note that the distortions induced by jumps are much less pronounced for the ratio statistic than for IQ. Again, the partial cancelation of the (upward) bias in the numerator and the denominator is operative. Thus, ratio-based inference is likely preferable regardless of the choice of estimator. In summary, estimating the ratio statistic using RMinRQ emerges as a natural part of practical jump-robust inference for IV or testing for jumps. Below, we explore whether this estimator is robust to other common "irregularities" in high-frequency return data as well.

5.2.3 Time-varying volatility

Pronounced intraday variation as well as seemingly abrupt changes (jumps) in spot volatility are prevalent in high-frequency returns. This poses a challenge for power variation estimation, as jump-robust estimators may not be able to distinguish sharply between rapidly shifting volatility and return jumps. For example, in the context of IV estimation, ADS (2012) document sizeable finite sample distortion in IV estimators when the intraday volatility is stochastic and subject to a diurnal U-shaped pattern. This section extends this analysis and draws broader conclusions for estimation of IQ and $\sqrt{IQ}/IV$ by exploring two distinct scenarios that violate the (locally) constant volatility assumption.

Our first scenario, "SV-U," is a modification of the corresponding design in ADS (2012). The diurnal pattern is calibrated to the average volatility of tick-time sampled trades of the stocks analyzed in the next section. In particular, we simulate a two-factor affine stochastic volatility model and superimpose an asymmetric diurnal pattern (Hasbrouck, 1999) for which the variance at the open is more than four times the midday and end-of-day variance.

The second scenario, "BM + 1 Volatility Jump," involves a six-fold spike in the intraday

variance at a random point in time, uniformly distributed across the trading day. Volatility is constant before the jump, and then constant at the new higher level following the jump. In this way, the scenario approximates the effect of sudden bursts in market activity that have inspired the development of alternative volatility jump specifications.

The striking similarity between Panels D and E of Tables 2-4 indicates that these two distinct forms of time variation in volatility have a qualitatively similar impact on the estimators in terms of finite sample bias and RMSE. Effectively, both scenarios render neighboring returns inhomogeneous, resulting in a downward bias due to scaling factors that are incorrectly sized as well as inappropriate truncation of diffusive returns that are misclassified as jumps due to the fluctuating level of the return variance. The less "local" estimators are more exposed to such heterogeneity. This explains the ordering of the biases of the IV and IQ estimators in Panels D and E of Tables 2 and 3, with estimators relying on block size one to three being the least biased, those based on blocks of four or five returns being slightly more biased, and finally the estimators relying on substantially larger block sizes (such as 20 for QRV in Table 2) being most biased.

In summary, Panels D and E of Table 3 provide evidence against the use of sparser sampling frequencies, such as ten minutes or lower, for IQ estimation. This runs counter to suggestions in the literature, indicating that biases in IQ estimation may be alleviated through sparse sampling. We find, in contrast, that the bias is much lower, and quite tolerable, at the two minute frequency, regardless of the block size of the IQ estimator.

Most importantly, Panels D and E in Table 4, confirm that the biases for the ratio $\sqrt{IQ}/IV$ are less pronounced and more uniform across the full range of estimators, as may be expected given that we obtain partial cancelation of the downward biases which affect both the numerator and denominator. Moreover, the ratio estimator based on RMinRQ and RMinRV again performs best from an efficiency standpoint in spite of the block size of five. For comparison, increasing the block size of the MPQ and MPV estimators from three to five enlarges the RMSE for $\sqrt{IQ}/IV$ in the "BM + 1 Volatility Jump" scenario. Consequently, the superior efficiency in estimating $\sqrt{IQ}/IV$ stems from the design of the robust neighborhood truncation principle rather than from the increased block size.

We conclude that inference based on the ratio $\sqrt{IQ}/IV$ appears to be attractive also under time-varying volatility. Moreover, the RMinRQ estimator provides quite compelling performance in this setting as well for the higher sampling frequencies.

5.2.4   Microstructure Noise

There is a tradeoff in the choice of sampling frequency with jump robustness improving and resiliency to microstructure noise deteriorating as the return interval shrinks. We now explore the effectiveness in dealing with the adverse impact of various noise features by applying pre-averaging, as detailed in Appendix C, and sampling at moderate frequencies. In fact, as outlined in Section 4.2, our asymptotic theory is developed in a noise-free setting, so we seek to determine whether this provides a suitable approximation for practical inference.

We consider four separate market imperfections. First, in our "BM + IID Noise" scenario, Panel F, Tables 2-4, we simulate Gaussian i.i.d. noise with a noise-to-signal ratio of $\lambda = 0.25$, in line with what is typical for trade data on individual stocks. Second, we consider a "BM + 1 Bounceback" scenario, Panel G, Tables 2-4, in which (isolated) errors in the recorded price induce so-called "bounce-backs" in returns, i.e., two large adjacent jumps of opposite sign due to immediate price reversals. We calibrate the magnitude of the bounce-back to match 20% of the daily QV (25% of IV). The third source of noise is irregular sampling, and the associated results are captured by our "BM + Sparcity" scenario, Panel H, Tables 2-4. It is generated via random arrivals of the 7,800 distinct quotes by sampling without replacement from the numbers in the range of 1 through 23,400. While not necessarily realistic, this model is helpful in exploring the potential distortion of the estimators when applied on non-homogeneously sampled returns, effectively inducing spurious variations in their volatility. Finally, in our "BM + Discrete Pricing" scenario, Panel I, Tables 2-4, we mimic price decimalization by rounding all intraday prices to the nearest cent with a starting price of $50. Price discreteness is a major reason for the presence of multiple zero returns in high-frequency samples, leading to pronounced downward biases of many jump-robust estimators.[23]

First, the reported relative biases in Panels F, G, H, and I of Tables 2-4 reveal that, irrespective of the noise scenario, it is necessary to avoid sampling at the highest frequency, i.e., 30 seconds, to obtain reasonably unbiased estimates of IV and IQ, while the ratio $\sqrt{IQ}/IV$ is unbiased for all scenarios except the "sparsity" setting. However, once we reach 2 minutes, all the relevant quantities are estimated without bias, except for a minor bias for IQ in the sparsity scenario.

Second, for the 2-minute frequency, the MSE is nearly identical for the Brownian motion case and the various noise scenarios, highlighting the efficacy of the pre-averaging, filtering and subsampling procedures. In particular, for the ratio statistic, the MSE is literally identical across all scenarios and estimators at the 2-minute frequency except for the sparse sampling case. That is, noise has no discernible impact on the asymptotic errors of the estimators at this moderate frequency, apart from minor distortions arising from inhomogeneous sampling of the returns. In the empirical work below, we mitigate this effect by sampling in tick time which renders the return variability more uniform across observations.

Third, across all noise scenarios, the relative bias and MSE for the ratio $\sqrt{IQ}/IV$ are dramatically lower than for the IQ estimators. Thus, the cancelation of outliers in the numerator and denominator helps robustify the ratio statistic in the presence of noise.

Finally, we note that the pre-averaging is extremely effective for the "BM + 1 Bounceback" scenario. This is due to the near perfect cancelation of adjacent jumps of opposite sign when constructing the individual pre-averaged returns. In what follows, we rely on pre-averaging as implemented in our simulation experiments to suppress the impact of noise also in our empirical illustrations on real market data.

---

[23]MPV and MPQ, in particular, as they are based on products of adjacent (absolute) returns.

6 An Illustration for the Dow Jones 30 Stocks

Since the "true" values of IV and IQ are latent, there is no simple way to directly compare the performance of alternative estimators. Moreover, for IQ in particular, there is a great deal of uncertainty regarding the actual precision of existing estimation procedures. The preceding analysis has focused on bringing out the features that render estimation inaccurate as well as non-robust, and then developing new approaches that should improve the inference. We now seek to establish whether the issues we have identified actually do pose a challenge for practical estimation and if the suggestions and procedures we have proposed appear helpful.

Consequently, this section explores properties of competing estimators of IV, IQ, and $\sqrt{IQ}/IV$ for the Dow Jones 30 stocks using tick time sampling of NYSE/TAQ trade data.[24] We split our sample period into a low volatility period, January 2005 - May 2007, and a high volatility period, June 2007 - July 2009. This serves as a robustness check against different noise-to-signal ratios and liquidity levels in the two periods.[25] The focus is on the estimates for IQ and $\sqrt{IQ}/IV$, while the IV results provide a benchmark for assessing the RNT estimators as well as the impact of our new filtering scheme relative to prior findings.

6.1 Truncation of Return Functionals

As documented in Section 5.2.4, power variation estimators can be quite sensitive to deviations from the local diffusive null, arising from microstructure features or recording errors. Common data filtering procedures may eliminate extreme outliers, but they are not sufficient to ensure sensible IQ estimates in practice for many candidate estimators of interest.

[Insert Figure 3 about here.]

Panel A of Figure 3 depicts signature plots for a group of IQ estimators, obtained by averaging the daily IQ estimates across the entire sample period and all of the thirty stocks. By construction, the figure speaks to mean and bias effects rather than efficiency. It shows that the estimators generally are quite similar although it also reveals some significant variation. First, at the highest frequencies a few estimators, especially MPQ5, appear severely downward biased. Hence, the noise-reduction associated with pre-averaging, sub-sampling and filtering has been successful in stemming the upward bias of the raw estimators. The remaining effects are consistent with the impact of irregular or sparse sampling of the returns at the highest frequencies, see Table 3, Panel H. The other striking feature is the slow decline in the plots as we move towards lower frequencies. The most likely explanation is the impact of time-varying volatility within the sampling interval, as indicated by the results in Table 3, Panels D and E. We pursue this issue in detail in the following section.

---

[24]When using tick sampled data, we are implicitly converting the calendar time scale to a tick scale, where time evolves linearly in tick time. This implies that the estimators are consistent for IQ in tick time, but not in calendar time. Importantly, in this setting the (tick time) IQ represents the relevant quantity for assessing the asymptotic variance of the (tick time based) IV estimator and for inference regarding jumps.

[25]Ex post, we find no qualitative differences in results between the two samples, so we only report findings for the initial sample. The full set of results may be gleaned from ADS (2011).

Nonetheless, it is noteworthy that the signature plots are quite flat and largely coincide for the various estimators in the range of 90-150 seconds.

By contrast, Panel B of Figure 3 offers widely diverging results. Here, none of the estimators are subject to functional filtering, so only TRQ and TBQ provide uniform truncation of stark outliers. The consequence is apparent. Apart from the extreme dampening achieved by MPQ5, the remaining estimators are wildly upward biased at the higher frequencies. This speaks to the lack of robustness of IQ estimators that do not exploit direct truncation of individual returns. The problem of excessive variability, or noise, in IQ estimates has been noted sporadically in the empirical literature and it has motivated some authors to rely on low, and relatively inefficient but less error prone, frequencies for IQ estimation, e.g., BNS (2004a) and Bandi and Russell (2008). Similarly, Jiang and Oomen (2008) uses the squared IV as a simple approximation to IQ, thereby accepting a significant bias in exchange for variance reduction of the IQ estimator.

Thus, it is very encouraging that the functional filtering regularizes the IQ estimators. The nominal size of our filter is 10^{-6}, so only gross violations of the Gaussian null is flagged. However, obviously, the null hypothesis is not satisfied for actual high-frequency data, so the truncation frequency is substantially larger in practice. For our equity data, the fraction of observations filtered ranges from 0.01% to 0.05% for the 60-180 second range, depending on the frequency and sample period. Overall, more than 70% of the stock days are untouched by our functional truncation. In contrast, TBQ truncates at least one observation on 99.9% of the days.[26] Overall, the evidence is compatible with our objective, namely that the functional filtering should control major data irregularities while avoiding excessively intrusive, and potentially distorting, truncation of the underlying returns.

6.2 The Intraday Volatility Pattern

One benefit of tick-time sampling of transactions data is that it tends to mitigate the intra-day U-shape pattern in volatility. Figure 4 demonstrates that the tick-time sampling succeeds in straightening the volatility pattern across the main part of the trading day, but there is little impact on the elevated level of volatility in the first 60-90 minutes of trading. Our Monte Carlo experiments found such intraday volatility variation to be a potent source of systematic biases in power variation estimates, with the ratio statistic $\sqrt{IQ}/IV$ being less sensitive to such distortions than the raw IQ and IV measures.

[Insert Figure 4 about here.]

To systematically assess the empirical relevance of intra-day volatility fluctuations, and in particular the diurnal volatility pattern, we split our equity sample into sub-groups

---

[26]Of course, this reflects the different philosophy behind the mild functional truncation filtering relative to TRQ and TBQ. The latter employ truncation of single returns as the primary tool for achieving jump robustness, and thus need to ensure that - asymptotically - all jumps are prevented from impacting the IQ estimator. This requires aggressive truncation and, inevitably, some truncation of diffusive returns, motivating the CPR adjustment to mitigate the resulting finite-sample bias, as discussed in Section 5.

consisting of stock-day combinations representing the top and bottom deciles with respect to a simple scale-free proxy for the intraday variation in volatility. This proxy, denoted VoV (volatility-of-volatility), is constructed by splitting the trading day into 26 blocks, and then obtaining the median 30 second absolute returns within each block after first having eliminated all zero returns. Our VoV proxy is then defined as the coefficient of variation (standard deviation divided by mean) of these 26 medians.[27] Importantly, the VoV measure is designed to pick up *any* significant variation in intra-day volatility and not just the commonly occurring U-shape. In the following section, we provide signature plots that broadly support our interpretation of the simulation evidence.

6.3 Signature Plots for Integrated Power Variation

Panel A of Figure 5 provides signature plots for the IV estimates averaged across all days and stocks. They are plotted as a function of the sampling frequency implied by the size of the pre-averaging window and the intensity of the transactions (in tick time). The RV estimator is included as a reference point and, as expected, lies above the jump-robust IV estimators which are bundled closely together across most of the frequencies.[28] The main outliers are MPV5 at frequencies higher than 120 seconds and QRV at frequencies lower than 90 seconds. At high frequencies, the drop in MPV5 and, to a lesser extent, MPV3 may be explained by the presence of zero returns, due to discreteness, which has a pronounced impact on these estimators. Consistent with this explanation, RV is the only estimator not to display a tendency to fall off at the highest frequencies. At lower frequencies, the striking downward bias of QRV is in line with the evidence from the simulation experiments. While all the jump-robust measures feature downward sloping signature plots, QRV is, by far, subject to the most significant distortion.

The IQ estimators, shown in Panel A of Figure 6, display a similar but more pronounced pattern of decline at the lowest sampling frequencies, consistent with the simulation scenarios that incorporate time-varying volatility. In particular, the estimators are roughly ordered by block size, with the 5-block estimators at the bottom and the 3-block estimators at the top. The truncation based IQ estimators, although nominally based on a short block, are disproportionately impacted by time variation in volatility due to the wide window necessary for determining the truncation thresholds, explaining the relatively sharp decline in their signature plots. We do not include an IQ analogue to the QRV estimator as it displays downward biases that are an order of magnitude larger than for the others.

The signature plot for the $\sqrt{IQ}/IV$ ratio in Panel A of Figure 7 displays a relatively flat but distinct monotonically declining pattern. This suggests that the microstructure effects, afflicting estimators at the highest frequencies, cancel out quite effectively at moderate frequencies consistent with the simulation evidence. The multi-power variation estimators display clear abnormalities at the highest frequencies, primarily due to an imperfect offset

---

[27]We confirm that alternative robust volatility-of-volatility measures lead to qualitatively similar results.

[28]The distance from the (non-functional filtered) RV curve to the set of robust IV measures provides an estimate of the average jump contribution to the quadratic return variation.

of the zero returns in the denominator. At sampling frequencies of 60 seconds or lower, however, all the estimators are tightly clustered and downward sloping in accordance with the findings from the simulation scenarios with time-varying volatility.

To further explore the impact of time-varying volatility, Panels B and C of each figure depict signature plots, respectively, for the top and bottom deciles with respect to our volatility-of-volatility proxy, VoV, across the combined stock-day sample. It is evident that both the IQ and IV signature plots on high VoV days are dramatically more steeply sloped than on low VoV days, corroborating the hypothesis that the slope is caused primarily by time-varying volatility due to factors like the intraday U-shape pattern, volatility jumps, and volatility bursts associated with news effects. For the ratio statistic in Figure 7, the signature plot for low VoV days is essentially flat, with the exception of the multi-power estimators, when viewed on the scale of the high VoV days, highlighting the fact that time variation in volatility also has a significant impact on the ratio statistic.

While the above evidence makes a fairly compelling case that VoV has a pronounced impact on the slope of the signature plots, it is based on grand averages across stocks and days and does not allow for formal statements about statistical significance. To provide a more rigorous analysis, controlling for the potential impact of a few outlier stocks, we have also run a series of panel regressions with stock fixed effects, capturing the average effect of VoV on the slope of the signature plot for each individual stock. These regressions confirm the strong significance of VoV on the slope of the signature plots.[29]

Given the accumulated evidence we consider a sampling frequency around 90-120 seconds as a sensible choice for inference about IV and jump tests based on the $\sqrt{IQ}/IV$ ratio across most of the estimators. At higher frequencies, the microstructure effects start impacting the estimators quite severely, while employing lower frequencies entails a significant loss of efficiency and ultimately also generates a severe downward bias, especially for days with turbulent market conditions when volatility-of-volatility fluctuates greatly. Furthermore, among the set of estimators we consider, the RNT estimators, RMinRQ and RMedRQ, appear to possess advantages in both efficiency and robustness.

[Insert Figures 5, 6, 7 about here.]

7 Conclusion

We provide a first in-depth look at robust estimation of integrated quarticity (IQ) based on high frequency data. The sensitivity of many existing IQ estimators to pervasive data irregularities inspires us to introduce a novel set of jump-robust estimators that are defined in terms of order statistics of suitable return functionals and generalize the existing nearest neighbor truncation estimators of ADS (2012). This new class of robust neighborhood

---

[29]For each individual robust estimator, we regressed the difference of the daily realized power variation estimate across the distinct frequencies against the daily VoV measure, allowing for a separate intercept term but enforcing a common slope across the stocks. The detailed results are available in ADS (2011).

truncation (RNT) estimators can be designed to enhance the robustness properties vis-à-vis microstructure noise features of the data as well as reducing the finite sample sensitivity to outliers. The identical principle can more generally be applied also to other consistent estimators to enhance jump- and noise-robustness. We find that the novel RNT estimators outperform existing estimators by a considerable margin in terms of finite sample efficiency in estimating the key ratio of IQ/IV^2. This quantity is extremely useful for robust inference regarding IV and for testing for price jumps.

In the empirical implementation, we emphasize the importance of appropriate filtering for gross violations of the particular null hypothesis associated with a given estimation procedure. In particular, we apply a novel functional filtering scheme for local power variation estimators, which generalizes truncation of individual returns to truncation of return functionals and is easy to apply for a broad range of popular estimators. By invoking this approach at the level of the local power variation estimators, the threshold can be set very conservatively, thereby avoiding systematic biases arising from aggressive truncation and thus eliminating the need for ex-post bias correction. We also emphasize the use of pre-averaging based on a wide pre-averaging window. This allows for important robustness to extreme outliers, like the so-called bounce-backs, and it simplifies the associated distribution theory as the impact of noise vanishes asymptotically.

The unifying theme behind the new class of estimators, as well as the universal filtering procedure applied to them, is to operate directly on the functional space of local power variation estimates instead of restricting attention to the underlying individual returns. In fact, we may view the latter as a special case, arising from a block size of unity. Combining the functional filtering with the novel RNT class of estimators enables efficient inference in an extensive simulation design and generates supportive evidence from an empirical application using the Dow Jones 30 stocks. Overall, the study provides a set of new guidelines for the construction of practical robust and efficient estimation and inference regarding IV and IQ.

References

Andersen, T. G. and T. Bollerslev (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review 39*, 885–905.

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. *Review of Economics and Statistics 89*, 701–720.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2010). *Parametric and Nonparametric Volatility Measurement.* North Holland. In Handbook of Financial Econometrics, Volume 1 - Tools and Techniques; Yacine Aït-Sahalia, Lars P. Hansen, and Jose A. Scheinkman (Eds.).

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*, 42–55.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2003). Modeling and forecasting realized volatility. *Econometrica 71*, 579–625.

Andersen, T. G., T. Bollerslev, and D. Dobrev (2007). No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d. noise: Theory and testable distributional implications. *Journal of Econometrics 138*, 125–180.

Andersen, T. G., D. P. Dobrev, and E. Schaumburg (2008). Duration based volatility estimation. *Manuscript, Northwestern University*.

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

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

Back, K. (1991). Asset pricing for general processes. *Journal of Mathematical Economics 20*, 371–395.

Bandi, F. M. and J. R. Russell (2008). Microstructure noise, realized variance, and optimal sampling. *Review of Economic Studies 75*, 339–369.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, M. Podolskij, and N. Shephard (2006). A central limit theorem for realized power and bipower variations of continuous semimartingales. *in From stochastic calculus to mathematical finance : the Shiryaev Festschrift, Springer Verlag*.

Barndorff-Nielsen, O. E., S. E. Graversen, J. Jacod, and N. Shephard (2006). Limit theorems for bipower variation in financial econometrics. *Econometric Theory 22*, 677–719.

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

Barndorff-Nielsen, O. E. and N. Shephard (2002). Estimating quadratic variation using realized variance. *Journal of Applied Econometrics 17*, 457–477.

Barndorff-Nielsen, O. E. and N. Shephard (2004a). Econometric analysis of realized covariation: High frequency based covariance, regression, and correlation in financial economics. *Econometrica 72*, 885–925.

Barndorff-Nielsen, O. E. and N. Shephard (2004b). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics 2*, 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*, 1–30.

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*, 796–806.

Christensen, K., R. Oomen, and M. Podolskij (2008). Realized quantile-based estimation of integrated variance. *Working paper; CREATES, University of Aarhus, Denmark*.

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

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

Hasbrouck, J. (1999). The dynamics of discrete bid and ask quotes. *Journal of Finance 54*, 2109–2142.

Hautsch, N. and M. Podolskij (2010). Pre-averaging based estimation of quadratic variation in the presence of noise and jumps: Theory, implementation, and empirical evidence. *Journal of Business and Economic Statistics forthcoming.*

Hsieh, D. A. (1991). Chaos and nonlinear dynamics: Applications to financial markets. *Journal of Finance 46*, 1839–1878.

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

Jacod, J., Y. Li, P. Mykland, M. Podolskij, and M. Vetter (2009). Mirostructure noise in the continuous case: the pre-averaging approach. *Stochastic Processes and Their Applications 26*, 2803–2831.

Jacod, J. and P. Protter (1998). Asymptotic error distributions for the euler method for stochastic differential equations. *Annals of Probability 26*, 267–307.

Jacod, J. and A. N. Shiryaev (2003). *Limit theorems for stochastic processes* (2nd ed.). Grundlehren der mathematischen Wissenschaften. Berlin; New York: Springer.

Jiang, G. J. and R. C. Oomen (2008). Testing for jumps when asset prices are observed with noise - a "swap variance" approach. *Journal of Econometrics 144*, 352–370.

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

Mykland, P. A. and L. Zhang (2009). Inference for continuous semimartingales observed at high frequency. *Econometrica 77*, 1403–1445.

Podolskij, M. and M. Vetter (2009a). Bipower-type estimation in a noisy diffusion setting. *Stochastic Processes and Their Applications 119*, 2803–2831.

Podolskij, M. and M. Vetter (2009b). Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli 15*, 634–658.

Schwert, G. W. (1989). Why does stock market volatility change over time? *The Journal of Finance 44*, 1115–1153.

Zhang, L., P. A. Mykland, and Y. Aït-Sahalia (2005). A tale of two time scales: Determining integrated volatility with noisy high-frequency data. *Journal of the American Statistical Association 100*, 1394–1411.

Zhou, B. (1996). High-frequency data and volatility in foreign-exchange rates. *Journal of Business and Economic Statistics 14*, 45–52.

A   Proofs of Propositions

This appendix provides proofs of Proposition 1 and related results for the general case of estimating the integrated power variation of order p, where p is a positive and even integer. The proofs are initially given for the MinPV and MedPV type estimators and subsequently shown to extend to the RNT estimators in Lemma 9 below.

The (integrated) power variation, $PV(p)$, is formally defined as,

$$PV(p) = \int_0^1 \sigma_u^p \, du. \tag{3}$$

Obviously, $p = 4$ corresponds to the theoretical quantity relevant for the $MinRQ$ and $MedRQ$ estimators while $p = 2$ refers to the integrated variance underlying the $MinRV$ and $MedRV$ estimators. The higher order integrated power variation estimators are less commonly used but do appear in the recent literature. For example, $p = 8$ is required to assess the (asymptotic) precision of integrated quarticity estimators.

A.1   Basic Setting

Let Y_t be the log price process following a Brownian semimartingale

$$Y_t \;=\; Y_0 + \int_0^t a_u \, du + \int_0^t \sigma_{u_-} \, dB_u, \tag{4}$$

where a is a locally bounded and predictable process and σ is adapted, cadlag and bounded away from zero. Without loss of generality, we further assume that the functions a, σ are uniformly bounded and $\inf_{t>0} \sigma_t > 0$ a.s.[30] The extension allowing for finite activity jumps in Y_t is dealt with Section A.6 below.

When discussing central limit theorems (CLTs) we require in addition that the volatility process follows a generalized Itô process:

Assumption (A1) : $\sigma_t = \sigma_0 + \int_0^t \tilde{a}_u \, du + \int_0^t \tilde{\sigma}_{u_-} \, dB_u + \int_0^t \tilde{v}_{u_-} \, dW_u$,

where $\tilde{a}$ is locally bounded and predictable and $\tilde{\sigma}$, $\tilde{v}$ are cadlag and the Brownian motions B, W are uncorrelated. We impose, without loss of generality, that the functions $\tilde{a}$, $\tilde{\sigma}$, and $\tilde{v}$ are uniformly bounded as well as $\inf_{t>0} \tilde{\sigma}_t > 0$ and $\inf_{t>0} \tilde{\nu}_t > 0$ a.s. We further note that, when the volatility process $\{\sigma_t\}_{t\geq 0}$ satisfies Assumption A1, then the power variation process, $\{\sigma_t^2\}_{t\geq 0}$, also conforms to this general characterization.

We assume Y is observed at $N+1$ evenly spaced time points spanning the interval $[0;1]$. Below, we denote these observations by $Y_{i/N}$, $i = 0, \ldots, N$, and the associated log-returns by $\Delta_i^N Y = Y_{i/N} - Y_{(i-1)/N}$, $i = 1, \ldots, N$. The proofs involve sequences of standardized return observations and corresponding approximating sequences for which volatility is fixed across one or more returns. Hence, we introduce non-overlapping blocks of $M \geq 1$ returns for which the volatility process is constant. We assume we have $K = N/M$ such blocks in the sample. Consequently, we define the quantities,

$$\chi_i^N \;=\; \sqrt{N} \, \Delta_i^N Y, \tag{5}$$

$$\beta_i^{N,M} \;=\; \sqrt{N} \, \sigma_{\lfloor \frac{(i-1)/M \rfloor M}{N} \rfloor} \, \Delta_i^N B \;=\; \sqrt{N} \, \sigma_{\lfloor \frac{(i-1)/M \rfloor}{K} \rfloor} \, \Delta_i^N B, \tag{6}$$

where $\lfloor \cdot \rfloor$ indicates the integer part of an expression. Hence, for each of the K return blocks, corresponding to $\beta_i^{N,M}$, the volatility remains fixed at the value it attains at the beginning of the block.

---

[30]As argued in Barndorf-Nielsen, Graversen, Podolskij, Jacod and Shephard (2006), henceforth BNGJPS, this follows from working with the stopped versions of the processes: $T_t^{(k)} = Y_{t \wedge T_k}$ and $\sigma_t^{(k)} = \sigma_{t \wedge T_k}$ where $T_k = \inf\{t \, | \, |a_t| + |\sigma_{t_-}| \geq k\}$ and $T_k \nearrow \infty$ a.s.

A.2 The Min and Med Power Variation Estimators

Let p be a fixed positive even integer and let $g_{min,p} : \mathbb{R}^2 \mapsto \mathbb{R}_+$ be given by,

$$
\begin{aligned}
g_{min,p}(\boldsymbol{a}) &= d_{min,p} \, \min\left(\,|a_1|^p, |a_2|^p\,\right), \quad \text{where the scaling constant takes the form,} \tag{7} \\
d_{min,p} &= \left(\, E\left[\,\min\left(\,|Z_1|^p, |Z_2|^p\,\right)\right]\right)^{-1}, \quad \text{and} \quad Z_1, Z_2 \sim i.i.d.\,\mathcal{N}(0,1)
\end{aligned}
$$

For example, we have $d_{min,2} = \frac{\pi}{\pi-2}$ and $d_{min,4} = \frac{\pi}{3\pi-8}$.

Similarly, we define the median-based function $g_{med,p} : \mathbb{R}^3 \mapsto \mathbb{R}_+$ and scaling factors,

$$
\begin{aligned}
g_{med,p}(\boldsymbol{a}) &= d_{med,p} \, \text{med}\left(\,|a_1|^p, |a_2|^p, |a_3|^p\,\right), \quad \text{where} \\
d_{med,p} &= \left(\, E\left[\,med\left(\,|Z_1|^p, |Z_2|^p\,\right), |Z_3|^p\,\right)\right]\right)^{-1}, \quad \text{and} \quad Z_1, Z_2, Z_3 \sim i.i.d.\,\mathcal{N}(0,1).
\end{aligned}
$$

In this case, $d_{med,2} = \frac{\pi}{6-4\sqrt{3}+\pi}$ and $d_{med,4} = \frac{3\pi}{9\pi+72-52\sqrt{3}}$.

For any even positive integer, p, we define the nearest neighbor truncation estimators of the p'th order power variation by,

$$
\text{MinPV}_N(p) = d_{min,p} \, \frac{1}{N-1} \sum_{i=1}^{N-1} \min\left((\chi_i^N)^p, (\chi_{i+1}^N)^p\right) = \frac{1}{N-1} \sum_{i=1}^{N-1} g_{min,p}\left(\chi_i^N, \chi_{i+1}^N\right),
$$

$$
\text{MedPV}_N(p) = d_{med,p} \, \frac{1}{N-2} \sum_{i=1}^{N-2} \min\left((\chi_i^N)^p, (\chi_{i+1}^N)^p, (\chi_{i+2}^N)^p\right) = \frac{1}{N-2} \sum_{i=1}^{N-2} g_{med,p}\left(\chi_i^N, \chi_{i+1}^N, \chi_{i+2}^N\right).
$$

For the cases of primary interest, i.e., $p=2$ and $p=4$, these estimators are identical to the *Min* and *Med* estimators introduced in Section 2.1. Specifically, we have

$$
\begin{aligned}
\text{MinPV}_N(2) &= \text{MinRV}_N, & \text{MedPV}_N(2) &= \text{MedRV}_N, \\
\text{MinPV}_N(4) &= \text{MinRQ}_N, & \text{MedPV}_N(2) &= \text{MedRQ}_N.
\end{aligned}
$$

A.3 Additional Notation and Preliminary Results

We provide a detailed proof of the results in Propositions 1 and 2 concerning the $\text{MinPV}_N(p)$ estimator. The proofs for $\text{MedPV}_N(p)$ may be derived similarly. Moreover, we henceforth consider a fixed even, positive integer, p, so the $g_{\min,p}$ function is uniquely defined. We refer to it simply as g below.

First, we observe that, for any bivariate vectors, $\boldsymbol{a} = (a_1, a_2)$ and $\boldsymbol{b} = (b_1, b_2)$, we have the following useful bound,

$$
|\,g(\boldsymbol{a}) - g(\boldsymbol{b})\,| \leq d_{min,p}\left(\,|\,a_1^p - b_1^p\,| + |\,a_2^p - b_2^p\,|\right), \tag{8}
$$

and furthermore that, except on the null set $\{(a_1, a_2) \in \mathbb{R}^2 | a_1 = a_2\}$, we have

$$
\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\left[\min(a_1^p, a_2^p + \varepsilon z) - \min(a_1^p, a_2^p)\right] = \begin{cases} z & \text{if} \quad |a_2| < |a_1| \\ 0 & \text{if} \quad |a_2| > |a_1| \end{cases} \tag{9}
$$

The proofs of Propositions 1 and 2 revolve around the sequences,

$$
\boldsymbol{V}_N = \frac{1}{N} \sum_{i=1}^{N-1} g\left(\chi_i^N, \chi_{i+1}^N\right) \quad \text{and} \quad \boldsymbol{U}_N^M = \frac{1}{N} \sum_{i=1}^{N-1} g\left(\beta_i^{N,M}, \beta_{i+1}^{N,M}\right).
$$

Since $\text{MinPV}_N(p) = \frac{N}{N-1} V_N$, the V_N sequence is asymptotically equivalent to our *MinPV* estimator, while U_N^M is an approximating sequence as, for large N, $\chi_i^N \approx \beta_i^{N,M}$.

For any adapted, integrable, d-dimensional cadlag process, X, and for $N \geq j > i - 1 \geq 0$, we define the expectation conditional on information at time $\frac{i-1}{N}$:

$$E_{i-1} \left[X_{\frac{j}{N}} \right] = E \left[X_{\frac{j}{N}} | \mathcal{F}_{\frac{i-1}{N}} \right] \tag{10}$$

A useful implication of our ability to focus on the case with uniformly bounded drift and volatility functions is that, using the Burkholder-Davis-Grundy inequalities, we have,

$$E_{i-1} \left[|\chi_i^N|^q \right] \leq C \quad \text{and} \quad E_{i-1} \left[|\beta_i^{N,M}|^q \right] \leq C, \tag{11}$$

where $q > 0$ and C denotes a generic positive constant which we (with slight abuse of notation) allow to take on disparate values in different places.

We decompose our estimators for the power variation, $PV(p)$, into a sum of conditional expectations and the associated martingale difference sequences: $V_N = V_{1N} + V_{2N}$ and $U_N^M = U_{1N}^M + U_{2N}^M$ where,

$$\boldsymbol{V}_{1N} = \frac{1}{N} \sum_{i=1}^{N-1} E_{i-1} \left[g\left(\chi_i^N, \chi_{i+1}^N\right) \right], \qquad \boldsymbol{V}_{2N} = \frac{1}{N} \sum_{i=1}^{N-1} \left\{ g(\chi_i^N, \chi_{i+1}^N) - E_{i-1} \left[g(\chi_i^N, \chi_{i+1}^N) \right] \right\}$$

$$\boldsymbol{U}_{1N}^M = \frac{1}{N} \sum_{i=1}^{N-1} E_{\lfloor \frac{i-1}{M} \rfloor M} \left[g\left(\beta_i^{N,M}, \beta_{i+1}^{N,M}\right) \right], \quad \boldsymbol{U}_{2N}^M = \frac{1}{N} \sum_{i=1}^{N-1} \left\{ g(\beta_i^{N,M}, \beta_{i+1}^{N,M}) - E_{\lfloor \frac{i-1}{M} \rfloor M} \left[g(\beta_i^{N,M}, \beta_{i+1}^{N,M}) \right] \right\}.$$

When $M = 1$ we will use the shorthand $\beta_i^N \equiv \beta_i^{N,1}$, $\boldsymbol{U}_N \equiv \boldsymbol{U}_N^1$ and similarly for the individual pieces $\boldsymbol{U}_{1N}$ and $\boldsymbol{U}_{2N}$. These definitions allow us to decompose the main estimator:

$$\boldsymbol{V}_N = \boldsymbol{U}_{1N} + \boldsymbol{U}_{2N} + (\boldsymbol{V}_{1N} - \boldsymbol{U}_{1N}) + (\boldsymbol{V}_{2N} - \boldsymbol{U}_{2N}) \tag{12}$$

Consistency of V_N can then be obtained by showing consistency of the estimator applied to the approximating Brownian path with piecewise constant volatility ($\boldsymbol{U}_N = \boldsymbol{U}_{1N} + \boldsymbol{U}_{2N}$) and then showing that the difference $V_N - U_N$ (the last two terms in (12) above) is asymptotically negligible. This is what we do in Section A.4 below. To prove a CLT, we exploit a different decomposition (similar to Mykland and Zhang (2007)), in which we show the CLT for our estimator applied to an approximating Brownian motion for which volatility is piecewise constant over blocks of length M. We then proceed to show that the difference between the original estimator and the estimator applied to the approximating process is negligible. This analysis is carried out in Section A.5 based on the decomposition:

$$\begin{aligned} \sqrt{N} \left(\boldsymbol{V}_N - PV(p) \right) &= \sqrt{N} \left(\boldsymbol{V}_{1N} - \boldsymbol{U}_{1N} \right) + \sqrt{N} \left(\boldsymbol{V}_{2N} - \boldsymbol{U}_{2N} \right) + \sqrt{N} \left(\boldsymbol{U}_{1N} - PV(p) \right) \\ &\quad + \sqrt{N} \left(\boldsymbol{U}_{2N} - \boldsymbol{U}_{2N}^M \right) + \sqrt{N} \, \boldsymbol{U}_{2N}^M \end{aligned} \tag{13}$$

A.4   Proposition 2: Consistency

We proceed by analyzing equation (12) term by term through a series of lemmas. For brevity, we focus on the features that are specific to our estimator, while referring to proofs in the extant literature when feasible. This also serves to highlight the underlying structural similarities between our $PV(p)$ measure and previously proposed power variation estimators and, in particular, IV and IQ estimators.

Lemma 2 *Under the maintained assumptions we have,*

$$\boldsymbol{U}_{1N} \xrightarrow{P} PV(p) \tag{14}$$

Moreover, if Assumption (A1) holds we obtain,

$$\sqrt{N} \left(U_{1N} - PV(p) \right) \xrightarrow{P} 0 \tag{15}$$

Proof. First, note that

$$g\left(\beta_i^N, \beta_{i+1}^N\right) \;=\; \left[g\left(\beta_i^N, \beta_{i+1}^N\right) - g\left(\beta_i^N, \sqrt{N}\sigma_{\frac{i-1}{N}}\Delta_{i+1}^N B\right)\right] + g\left(\beta_i^N, \sqrt{N}\sigma_{\frac{i-1}{N}}\Delta_{i+1}^N B\right)$$

so we may write

$$\boldsymbol{U}_{1N} \;=\; \frac{1}{N}\sum_{i=1}^{N-1}\mathbb{E}_{i-1}\left[g\left(\beta_i^N, \beta_{i+1}^N\right) - g\left(\beta_i^N, \sqrt{N}\sigma_{\frac{i-1}{N}}\Delta_{i+1}^N B\right)\right] + \frac{1}{N}\sum_{i=1}^{N-1}\sigma_{\frac{i-1}{N}}^p \tag{16}$$

The first sum in (16) tends to zero in probability. To see this, note that the bound (8) implies the following limit in L_2-norm:

$$\mathbb{E}\left|\frac{1}{N}\sum_{i=1}^{N-1}\mathbb{E}_{i-1}\left[g\left(\beta_i^N, \beta_{i+1}^N\right) - g\left(\beta_i^N, \sqrt{N}\sigma_{\frac{i-1}{N}}\Delta_{i+1}^N B\right)\right]\right|^2 \;\leq\; \frac{C}{N}\,\mathbb{E}\left[\sum_{i=1}^{N-1}\left|\sigma_{\frac{i}{N}}^p - \sigma_{\frac{i-1}{N}}^p\right|^2\right] \;\to\; 0 \tag{17}$$

where the convergence (17), and thus also convergence in probability, follows from the fact that σ_t^p has finite quadratic variation (since σ_t is a cadlag semimartingale). In addition, since $\{\sigma_t^p\}_{t\geq 0}$ is uniformly bounded and cadlag, the pointwise dominated convergence of $\left(\sigma_u^p - \sigma_{\frac{\lfloor uN\rfloor}{N}}^p\right) \to 0$ for $u \in [0;1]$ follows and Lebesgue's theorem yields

$$\sum_{i=1}^{N-1}\left[\int_{\frac{(i-1)}{N}}^{\frac{i}{N}}\left(\sigma_u^p - \sigma_{\frac{(i-1)}{N}}^p\right)du\right] \;\overset{a.s.}{\to}\; 0 \tag{18}$$

Together (17) and (18) imply $PV(p) - U_{1N} \to 0$, which establishes (14). To show (15) we need the stronger assumption (A1). Define the sequence of independent standard normals $Z_i = \sqrt{N}\Delta_i^N B$, then Assumption (A1) yields

$$\mathbb{E}_{i-1}\left[\left(\sigma_{\frac{i}{N}}^p - \sigma_{\frac{i-1}{N}}^p\right)Z_{i+1}^p\mathbf{1}_{|Z_{i+1}|<|Z_i|}\right] \;=\; \mathbb{E}_{i-1}\left[\left(\sigma_{\frac{i}{N}}^p - \sigma_{\frac{i-1}{N}}^p\right)\varphi(Z_i)\right] \;=\; O_P\left(1/N\right) \tag{19}$$

since $\varphi(Z_i) = \mathbb{E}_i[Z_{i+1}^p\mathbf{1}_{|Z_{i+1}|<|Z_i|}]$ is an even function of the Brownian path $\{B_t\}_{(i-1)/N<t<i/N}$. Now the property (9) yields

$$\mathbb{E}_{i-1}\left[g\left(\beta_i^N, \beta_{i+1}^N\right) - g\left(\beta_i^N, \sqrt{N}\sigma_{\frac{i-1}{N}}\Delta_{i+1}^N B\right)\right] \;=\; O_P\left(1/N\right) \tag{20}$$

This ensures that the first term in (16) is asymptotically negligible, even when scaled up by $\sqrt{N}$. Hence, the remaining task is to show,

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N-1}\sigma_{\frac{i-1}{N}}^p - PV(p)\right) \;\overset{P}{\to}\; 0\,.$$

However, this is analogous to the common task of showing that

$$\sqrt{N}\left(\frac{1}{N}\sum_{i=1}^{N-1}\sigma_{\frac{i-1}{N}}^2 - IV\right) \;\overset{P}{\to}\; 0\,.$$

in the IV literature and the method of proof is, by now, well established; see, e.g., BNGJPS where the result is shown in a general setting (allowing for infinite activity jumps) of which the current framework is a special case. A more intuitive and detailed exposition is provided by Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006), henceforth BNGJS. ∎

Lemma 3 *Under the maintained assumptions, we have*

$$\boldsymbol{U}_{2N} \overset{P}{\to} 0 \tag{21}$$

Proof. To simplify notation, define the martingale difference sequence $\left\{ \frac{1}{N} \eta_i^N, \mathcal{F}_{\frac{i}{N}} \right\}_{i \geq 0}$:

$$\eta_i^N = g(\beta_i^N, \beta_{i+1}^N) - E_{i-1}\left[g(\beta_i^N, \beta_{i+1}^N) \right]$$

Note that $\mathbb{E}[(\eta_i^N)^2 | \mathcal{F}_{\frac{i-1}{N}}] \leq C$, so applying the Cauchy-Schwartz inequality,

$$\boldsymbol{V}\left[\frac{1}{N} \sum_{i=1}^{N} \eta_i^N \right] = \frac{1}{N} \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^{N} \left((\eta_i^N)^2 + 2\eta_i^N \eta_{i+1}^N \right) \right] \leq \frac{C}{N} \mathbb{E}\left[\frac{1}{N} \sum_{i=1}^{N} \mathbb{E}[(\eta_i^N)^2 | \mathcal{F}_{\frac{i-1}{N}}] \right] \leq \frac{C}{N} \to 0 \,.$$

The L_2 convergence implies $\frac{1}{N} \sum_{i=1}^{N} \eta_i^N \overset{P}{\to} 0$. ∎

Lemma 4 *Under the maintained assumptions, we have,*

$$(\boldsymbol{V}_{1N} - \boldsymbol{U}_{1N}) \overset{P}{\to} 0 \,. \tag{22}$$

Under Assumption (A1), we obtain,

$$\sqrt{N} \, (\boldsymbol{V}_{1N} - \boldsymbol{U}_{1N}) \overset{P}{\to} 0 \,. \tag{23}$$

Proof. We must show,

$$V_{1N} - U_{1N} = \frac{1}{N} \sum_{i=1}^{N-1} \mathbb{E}_{i-1}\left[\left(g(\chi_i^N, \chi_{i+1}^N) - g(\beta_i^N, \beta_{i+1}^N) \right) \right] \to 0 \text{ as } N \to \infty \,. \tag{24}$$

Using the bound (8), it follows that,

$$
\begin{aligned}
V_{1N} - U_{1N} &\leq \frac{1}{N} \mathbb{E}\left[\sum_{i=1}^{N-1} | g(\chi_i^N, \chi_{i+1}^N) - g(\beta_i^N, \beta_{i+1}^N) | \right] \\
&\leq \frac{C}{N} \mathbb{E}\left[\sum_{i=1}^{N} | (\chi_i^N)^p - (\beta_i^N)^p | \right] = \frac{C}{N} \sum_{i=1}^{N} \left(\mathbb{E}_{i-1} \left| h(\sqrt{N}\Delta_i^N Y) - \sigma_{\frac{i-1}{N}}^p \right| \right)
\end{aligned}
$$

where we have defined the function $h(x) = x^p$. This formulation maps directly into the setting of BNGJPS where the results of this lemma are proven in a more general setting and for a generic $h(x)$ function subject to regularity conditions. In particular, our h function trivially satisfies the continuous differentiability and polynomial growth conditions necessary for the applicability of their analysis. An accessible, albeit lengthy, account of the steps of the argument may be found in BNGJS (2006, pp. 713-719). So while this proof is quite involved, the above reformulation of the relevant inequalities, as they arise within our specific setting, allows us to simply refer to previously published work for the result. ∎

Lemma 5 *Under the maintained assumptions, we have,*

$$(\boldsymbol{V}_{2N} - \boldsymbol{U}_{2N}) \overset{P}{\to} 0 \,. \tag{25}$$

Moreover, we may strengthen this result further to obtain,

$$\sqrt{N} \, (\boldsymbol{V}_{2N} - \boldsymbol{U}_{2N}) \overset{P}{\to} 0 \,. \tag{26}$$

Proof. In order to demonstrate the second result of the lemma, which obviously implies the first, we define,

$$\xi_i^N = (1/\sqrt{N}) \left[g\left(\chi_i^N, \chi_{i+1}^N \right) - g\left(\beta_i^N, \beta_{i+1}^N \right) \right],$$

and we must then prove that,

$$\sum_{i=1}^{N-1} \left(\xi_i^N - \mathbb{E}_{i-1}[\xi_i^N] \right) \xrightarrow{P} 0.$$

This is a martingale difference sequence with respect to the filtration $\mathcal{F}_{\frac{i}{N}}$, so it suffices to show,

$$\sum_{i=1}^{N-1} \mathbb{E}\left[(\xi_i^N)^2 \right] = \mathbb{E}\left[\sum_{i=1}^{N-1} \mathbb{E}_{i-1}\left[(\xi_i^N)^2 \right] \right] \to 0 \quad \text{as} \quad N \to \infty.$$

Mimicking the type of steps undertaken in the proof of the previous lemma, including application of the uniform bound on moments of χ_i^N and β_i^N, we obtain,

$$
\begin{aligned}
\sum_{i=1}^{N-1} \mathbb{E}\left[(\xi_i^N)^2 \right] &= \frac{1}{N} \mathbb{E}\left[\sum_{i=1}^{N-1} \mathbb{E}_{i-1} \left| g(\chi_i^N, \chi_{i+1}^N) - g(\beta_i^N, \beta_{i+1}^N) \right|^2 \right] \\
&\leq \frac{C}{N} \mathbb{E}\left[\sum_{i=1}^{N} \mathbb{E}_{i-1}\left[\left(h(\chi_i^N) - h(\beta_i^N) \right)^2 \right] \right].
\end{aligned}
$$

As for the previous lemma, our reformulation of the task maps the problem into the corresponding task in BNGJPS (2006) who prove a corresponding lemma in a more general setting. A detailed account of the requisite steps to complete this part of the proof may again be gleaned from BNGJS (2006, pp. 704-706). ■

Taken together, Lemma 2 - 3 and the first parts of Lemma 4 - 5 imply the consistency of our estimator under the minimal maintained assumptions. The second parts of Lemmas 4 - 5 are critical for the proof of the central limit theorem below.

A.5 Proposition 3: The CLT

Lemma 6 *Under assumption (A1), we have*

$$\sqrt{N}\, \boldsymbol{U}_{2N}^M \xrightarrow{stable\ \mathcal{D}} \mathcal{N}\left(0,\, \nu \int_0^1 \sigma^{2p}\, du \right) \tag{27}$$

where the constant $\nu = \mathbb{V}ar\left[g(Z_0, Z_1) + 2\,\mathbb{C}ov\left[g(Z_0, Z_1), g(Z_1, Z_2) \right] \right.$ *for* $Z_0, Z_1, Z_2 \sim i.i.d.\mathcal{N}(0,1)$.

Proof. Consider splitting the N scaled return observations into K blocks, the kth of which is the vector $\boldsymbol{\chi}_k^M = \{\sqrt{N}\Delta_i^N Y\}_{i \in \{(k-1)M+1,\ldots,kM\}}$. The corresponding vector of observations from the approximating Brownian motion where volatility is held constant over the block is $\boldsymbol{\beta}_k^{N,M} = \{\beta_i^{N,M}\}_{i \in \{(k-1)M+1,\ldots,kM\}}$. Next, define by $g_M(\cdot) : \mathbb{R}^M \mapsto \mathbb{R}$ the block estimator of volatility:

$$g_M(\boldsymbol{\beta}_k^{N,M}) = \frac{1}{M} \sum_{i=(k-1)M+1}^{kM-1} g(\beta_i^{N,M}, \beta_{i+1}^{N,M}) \tag{28}$$

We wish to apply Theorem IX.7.28 in Jacod and Shiryaev (2003) to $\sqrt{N}\,\boldsymbol{U}_{2N}^M$. Defining the martingale difference sequence $\psi_k^{N,M} = \sqrt{M}\left(g_M(\boldsymbol{\beta}_k^{N,M}) - \frac{M-1}{M}\sigma^2_{\frac{(k-1)}{K}}\right)$ we can write

$$
\begin{aligned}
\sqrt{N}\,\boldsymbol{U}_{2N}^M &= \frac{1}{\sqrt{K}}\sum_{k=1}^{K}\psi_k^{N,M} + \frac{1}{\sqrt{N}}\sum_{k=1}^{K-1}\left(g(\beta_{kM}^{N,M},\beta_{kM+1}^{N,M}) - \mathbb{E}_{\frac{k}{K}}\left[g(\beta_{kM}^{N,M},\beta_{kM+1}^{N,M})\right]\right) \\
&= \frac{1}{\sqrt{K}}\sum_{k=1}^{K}\psi_k^{N,M} + o_P(1)
\end{aligned}
\tag{29}
$$

The last equality follows from the fact that each term in the second sum is centered and has bounded variance (given the uniform bound on σ_t). Thus the sum divided by $\sqrt{N}$ will tend to zero provided $K = o_P(N)$.

We must now verify conditions (7.27)-(7.31) of Theorem IX.7.28. First note that $E[\,\psi_k^{N,M}\,|\,\mathcal{F}_{\frac{k-1}{K}}\,] = 0$ so that condition (7.27) is trivially satisfied. Condition (7.28) follows from the fact that

$$
\frac{1}{K}\sum_{k=1}^{K} E\left[\left\{\sqrt{M}\left(g_M(\boldsymbol{\beta}_k^M) - \frac{M-1}{M}\sigma^P_{\frac{(k-1)}{K}}\right)\right\}^2 \Big|\,\mathcal{F}_{\frac{k-1}{K}}\right] = \frac{\nu}{K}\sum_{k=1}^{K}\sigma^{2p}_{\frac{k-1}{K}} \xrightarrow{P} \nu\int_0^1 \sigma_u^{2p}\,du \tag{30}
$$

where the convergence in probability (and in fact a.s.) is a consequence of the volatility process being cadlag and uniformly bounded. Next, we turn to condition (7.29). Let $\Delta_k^M \boldsymbol{B} = \left(B_{\frac{k}{K}} - B_{\frac{(k-1)}{K}}\right)$, then $E\left[\psi_k^{N,M}\Delta_k^M \boldsymbol{B}\,\big|\,\mathcal{F}_{\frac{k-1}{K}}\right] = 0$, which follows from the fact that the variables $\psi_k^{N,M}$ are centered and that g_M is an even function. Condition (7.30), stating that $E\left[(\psi_k^{N,M})^2\,\mathbf{1}_{|\psi_k^{N,M}|>\varepsilon}\right] \xrightarrow{P} 0$, follows straightforwardly from the fact that σ is uniformly bounded.

Finally, let $\{N_t\}_{t\in[0;1]}$ be a bounded martingale orthogonal to B (i.e. the covariation $\langle B,N\rangle_t = 0$ a.s.). We want to show that, for each block k, $E[\psi_k^{N,M}\left(N_{\frac{k}{K}} - N_{\frac{(k-1)}{K}}\right)|\mathcal{F}_{\frac{k-1}{K}}] = 0$. For $t > \frac{k-1}{K}$ consider the martingale difference sequence $M_t = E\left[\psi_k^{N,M}\,\big|\,\mathcal{F}_t\right]$. By the martingale representation theorem, $M_t = M_{\frac{k-1}{K}} + \int_{\frac{k-1}{K}}^{\frac{k}{K}}\varphi_u\,dB_u$ for some predictable process φ_u. Therefore the processes $\{M_t\}_{t>\frac{k-1}{K}}$ and $\{N_t - N_{\frac{k-1}{K}}\}_{t>\frac{k-1}{K}}$ are orthogonal and the product, $\{M_t(N_t - N_{\frac{k-1}{K}})\}$ is again a martingale which must then have mean zero. This verifies condition (7.31) and Theorem IX.7.28 in Jacod and Shiryaev (2003) states that as N (and hence K and M) tend to infinity:

$$
\sqrt{N}\,\boldsymbol{U}_{2N}^M \xrightarrow{\text{stable}} \mathcal{N}\left(0,\nu\int_0^1 \sigma^{2p}\,du\right) \tag{31}
$$

$\blacksquare$

Lemma 7 *Under the maintained assumptions, we have*

$$
\sqrt{N}\left(\boldsymbol{U}_{2N} - \boldsymbol{U}_{2N}^M\right) \xrightarrow{P} 0 \tag{32}
$$

Proof. Defining $\eta_i^{N,M} = g(\beta_i^{N,M},\beta_{i+1}^{N,M}) - \mathbb{E}_{\lfloor(i-1)/M\rfloor M}\left[g(\beta_i^{N,M},\beta_{i+1}^{N,M})\right]$, we note that $\left\{\frac{1}{\sqrt{N}}(\eta_i^N - \eta_i^{N,M})\right\}_{i\geq 1}$ is a martingale difference sequence with respect to the filtration $\{\mathcal{F}_{i/N}\}$. To show that $\sqrt{N}\left(\boldsymbol{U}_{2N} - \boldsymbol{U}_{2N}^M\right) = \sum_{i=1}^{N-1}(\eta_i^N - \eta_i^{N,M})/\sqrt{N} \to 0$ in probability, it therefore suffices (by Doobs inequality, e.g. Revuz and Yor (1999)) to show that

$$
\frac{1}{N}\mathbb{E}\left[\sum_{i=1}^{N-1}\left|g(\beta_i^N,\beta_{i+1}^N) - g(\beta_i^{N,M},\beta_{i+1}^{N,M})\right|^2\right] \to 0 \tag{33}
$$

By the bound of $g(\cdot)$ we have

$$\frac{1}{N} \mathbb{E}\left[\sum_{i=1}^{N-1} \left| g(\beta_i^N, \beta_{i+1}^N) - g(\beta_i^{N,M}, \beta_{i+1}^{N,M}) \right|^2\right] \leq \frac{C}{N} \mathbb{E}\left[\sum_{i=1}^{N} \mathbb{E}_{i-1} \left| (\beta_i^N)^p - (\beta_i^{N,M})^p \right|^2\right]$$

$$\leq \frac{C}{N} \mathbb{E}\left[\sum_{i=1}^{N} \left| \sigma_{\frac{i-1}{N}}^p - \sigma_{\frac{\lfloor (i-1)/M \rfloor M}{N}}^p \right|^2\right] = C\, \mathbb{E}\int_0^1 \left(\sigma_{\frac{\lfloor uN \rfloor}{N}}^p - \sigma_{\frac{\lfloor uK \rfloor}{K}}^p \right)^2 du = o_P(1), \qquad (34)$$

where the last inequality follows from the uniform boundedness of σ_t and Lebesgues theorem. ∎

Importantly, the specification of the volatility process in Assumption (A1) may be extended to include finite as well as infinite activity jump processes subject only to the regularity conditions stipulated in BNGJPS. This follows from the fact that the only terms in (13) affected by the inclusion of jumps are the terms $\sqrt{N}(V_{1N} - U_{1N})$ and $\sqrt{N}(V_{2N} - U_{2N})$ which map into the corresponding terms in BNGJPS as outlined in the proofs above. As such, the distributional results of the paper cover a wide range of underlying return generating processes.

A.6 The Asymptotic Distribution under Jump Alternatives

Suppose now the log price process is given as $X = Y + J$, where Y is a Brownian semimartingale of the form (4) while J is a finite activity jump process. We show below that the above results continue to hold.[31] The key is that $\sqrt{N}|Y_{\frac{i}{N}} - Y_{\frac{i-1}{N}}| = O_P(|\log(N)|^{1/2})$, which follows readily from Levy's modulus of continuity theorem for Brownian motion. This immediately yields:

Proposition 8 *When J is a **finite activity** jump process, the asymptotic distribution of the $minPV(p)$ and $MedPV(p)$ estimators applied to the processes $\{X_t\}$ and $\{Y_t\}$ are identical.*

Proof. As before, we deal only with the $MinPV$ case as the $MedPV$ case is analogous. On a given realization of the path there is a finite number of jumps, so (asymptotically) at most one of the terms $|X_{\frac{i}{N}} - X_{\frac{i-1}{N}}|$ or $|X_{\frac{i+1}{N}} - X_{\frac{i}{N}}|$ includes a jump. Therefore, each term in the estimator (up to a normalizing constant) is

$$\min\left(|X_{\frac{i}{N}} - X_{\frac{i-1}{N}}|^p, |X_{\frac{i+1}{N}} - X_{\frac{i}{N}}|^p\right) = O_P\left(\frac{(\log N)^{p/2}}{N}\right)$$

regardless of whether a (single) jump occurred or not over $[\frac{i-1}{N}, \frac{i+1}{N}]$. Since only finitely many terms differ,

$$\sum_{j=1}^{N}\left[\min\left(|X_{\frac{i}{N}} - X_{\frac{i-1}{N}}|^p, |X_{\frac{i+1}{N}} - X_{\frac{i}{N}}|^p\right) - \min\left(|Y_{\frac{i}{N}} - Y_{\frac{i-1}{N}}|^p, |Y_{\frac{i+1}{N}} - Y_{\frac{i}{N}}|^p\right)\right]$$

$$= O_P\left(\frac{(\log N)^{p/2}}{N}\right) = o_P\left(\frac{1}{\sqrt{N}}\right)$$

so neither consistency nor convergence in distribution is affected by the presence of finite activity jumps. ∎

A.7 Robust Neighborhood Truncation Estimators

We consider the family of robust neighborhood truncation (RNT) estimators on a block of m i.i.d. $\mathcal{N}(0, \sigma^2)$ returns, $\{Z_1, \ldots, Z_m\}$. The estimator is then constructed by taking the jth quantile of H unbiased estimators of σ^p on the block. Denoting these primitive estimators by $\mathcal{E}_1, \ldots, \mathcal{E}_H$, we can write the RNT estimator as

$$RNT_N^{(j,I)}(p) = d_{(j,I)}(p)\, \frac{1}{(N-m+1)} \sum_{i=1}^{N-m+1} q_j\, [\mathcal{E}_1, \ldots, \mathcal{E}_H]$$

---

[31] As for the volatility process, the specification may be generalized to infinite activity jump processes along the lines of Barndorff-Nielsen et al (2006c).

where the $d_{(j,I)}(p)$ is a scaling factor.

Lemma 9 *Let p be a positive even integer and assume that the estimators $\mathcal{E}_1, \ldots, \mathcal{E}_H$ satisfy the conditions of Proposition 1 and Proposition 2, then the robust neighborhood truncation estimator, $RNT_N^{(j,\mathbf{I})}(p)$, defined in (35) is consistent for σ^p and satisfies a CLT.*

Proof. We need to verify the three properties,(8)-(9) and symmetry, of the $g(\cdot)$ function used in the theorem are satisfied when $g(\cdot) = RNT_N^{(j,I)}(p)$. We deal with each condition in turn.

Clearly, if each primitive estimator $\mathcal{E}_i$ is symmetric, so is $RNT_N^{(j,I)}(p)$. Moreover, if each $\mathcal{E}_i$ satisfies a bound of the type (8), so will $RNT_N^{(j,I)}(p)$ as it is simply an order statistic of such bounded functionals. Finally, assume that each $\mathcal{E}_i$ satisfies (9). Except on a null set, there exists a neighborhood around each m-tuple $(z_1, \ldots, z_m)$, on which $RNT_N^{(j,I)}(p) = \mathcal{E}_i(z_1^p, \ldots, z_m^p)$ for some $1 \leq i \leq H$. Therefore it follows that also $RNT_N^{(j,I)}(p)$ satisfies (9). ∎

Remark 10 *Since the NT estimators (up to a scale factor) essentially are a special case of RNT, the lemma applies to these as well.*

B   Noise Robustness Properties of the Ratio IQ/IV^2

The ratio IQ/IV^2 plays an important role for both IV inference and jump attribution in finite samples. This section extends the analysis of Huang and Tauchen (2005) to show that the ratio IQ/IV^2 has certain desirable robustness features in the presence of microstructure noise. Following Ait-Sahalia and Mykland (2005), we assume that the true price process (p_i^*) is observed at $N+1$ discrete points in time with an independent stationary (possibly autocorrelated) measurement error (u_i) that results in an MA error structure in observed returns:

$$p_i = p_i^* + u_i \;\Rightarrow\; r_i = r_i^* + \varepsilon_i \;\; \text{where} \;\; \varepsilon_i = u_i - u_{i-1}$$

For simplicity, we focus the discussion here on the (RV,RQ) pairing and denote:

$$\widehat{IV} = \sum_{i=1}^N r_i^2 , \; IV^* = \sum_{i=1}^N r_i^{*2} , \; \widehat{IQ} = \frac{N}{3}\sum_{i=1}^N r_i^4 , \; IQ^* = \frac{N}{3}\sum_{i=1}^N r_i^{*4}$$

The presence of microstructure noise produces a bias in both the IQ and IV estimators of the form:

$$
\begin{aligned}
E[\widehat{IQ}] &= E[IQ^*] + \frac{N^2}{3}E[\varepsilon_i^4] + 2NE[IV^*]E[\varepsilon_i^2] \\
E[\widehat{IV}^2] &= E[IV^{*2}] + NE[\varepsilon_i^4] + N(N-1)(E[\varepsilon_i^2])^2 + 2(N+2)E[IV^*]E[\varepsilon_i^2] \\
&\qquad\qquad\qquad + 2\sum_{m=1}^{N-1}(N-m)Cov(\varepsilon_1^2, \varepsilon_{1+m}^2)
\end{aligned}
\tag{35}
$$

In the special case where $u_i \sim \mathcal{N}(0, \sigma_u^2)$ is serially uncorrelated and denoting the noise to signal ratio $\lambda = \sigma_u^2/(\frac{1}{N}E[IV^*])$, the expressions above simplify to,

$$
\begin{aligned}
E[\widehat{IQ}] &= E[IQ^*] + 4E[IV^{*2}](\lambda + \lambda^2) \\
E[\widehat{IV}^2] &= E[IV^{*2}] + 4E[IV^{*2}](\lambda + \lambda^2) + O\left(\frac{1}{N}\right)
\end{aligned}
\tag{36}
$$

Under the null of no jumps, $\hat{IQ}, \hat{IV}$ are asymptotically unbiased and consistent and,

$$\frac{\widehat{IQ}}{\widehat{IV}^2} \;\xrightarrow{P}\; \frac{IQ^* + 4IV^{*2}(\lambda + \lambda^2)}{IV^{*2} + 4IV^{*2}(\lambda + \lambda^2)} , \quad \text{as } N \to \infty$$

The downward bias of the limiting ratio depends on the noise-to-signal ratio λ and preaveraging of

returns therefore plays an important role in reducing λ and the associated distortions.[32] Moreover, for sufficiently pre-averaged returns, there is very little evidence of serial correlation, as pointed out by COP (2010), and the serially uncorrelated noise case considered above is therefore the empirically most relevant case.[33] In finite samples, this downward bias is further compounded by a pure Jensen (concavity) effect as readily seen from the Monte Carlo results for the Brownian motion scenario Path by path, the Cauchy-Schwartz inequality of course implies that $\widehat{IQ}/\widehat{IV}^2 \geq \frac{1}{3}$ must hold regardless of the noise structure or other imperfections.

In the presence of other deviations from the Brownian null, forming the ratio $\widehat{IQ}/\widehat{IV}^2$ may have a stabilizing effect provided that the resulting distortion is uniform and roughly proportional to squared returns since this will lead to a cancelation in numerator and denominator. We see this effect at work in the simulations with price discreteness but in other instances, e.g., sparsity, it clearly fails. In cases involving additive distortions such as jumps, there will be no cancelation of biases and the ratio will tend to (in the case of upward biases) diverge at high frequencies due to the scaling by N in the numerator.

C $\,$ IV and IQ estimators based on pre-averaged returns

Given N equally-spaced (log) returns, $r_i = Y_{\frac{i}{N}} - Y_{\frac{i-1}{N}}$, $i = 1, \cdots, N$, we define corresponding pre-averaged returns for any pre-averaging window size $2K \leq N$:

$$\overline{r}_i \;=\; \frac{1}{K} \sum_{j=K}^{2K-1} Y_{\frac{i+j}{N}} - \frac{1}{K} \sum_{j=0}^{K-1} Y_{\frac{i+j}{N}}, \; i = 1, \cdots, N - 2K + 1 \; . \tag{37}$$

An equivalent definition with analytically more tractable expression is given by:

$$\overline{r}_i \;=\; 2 \sum_{j=1}^{2K-1} g\left(\frac{j}{2K}\right) r_{i+j} \; , \; i = 1, \cdots, N - 2K + 1 \; , \tag{38}$$

where $g(x) = x \wedge (1-x)$, $x \in [0,1]$ is the pre-averaging kernel. We further define $\psi_K = \frac{1}{2K} \sum_{j=1}^{2K-1} 4 g\left(\frac{j}{2K}\right)^2$ as the finite sample analog of the variance scaling factor $\psi = \int_0^1 4 g(u)^2 du = \frac{1}{3}$ induced by the pre-averaging kernel.[34]

Consider the following $2K$ sub-samples of non-overlapping pre-averaged returns:

$$\begin{aligned}
\overline{S}_1 &= \left\{ \overline{r}_{1+2(s-1)K} \; : \; s = 1, ..., \left\lfloor \frac{N}{2K} \right\rfloor \right\} \\
\overline{S}_2 &= \left\{ \overline{r}_{2+2(s-1)K} \; : \; s = 1, ..., \left\lfloor \frac{N-1}{2K} \right\rfloor \right\} \\
&\quad ... \\
\overline{S}_{2K} &= \left\{ \overline{r}_{2K+2(s-1)K} \; : \; s = 1, ..., \left\lfloor \frac{N-2K+1}{2K} \right\rfloor \right\}
\end{aligned}$$

Let $\widehat{IV}\left[\overline{S}_i\right]$ and $\widehat{IQ}\left[\overline{S}_i\right]$, $i = 1, \cdots, 2K$ denote the raw IV and IQ estimates obtained on each sub-sample of pre-averaged returns. Then the pre-averaged (and sub-sampled) estimators $\widehat{IV}\left[\overline{r}_\cdot\right]$ and $\widehat{IQ}\left[\overline{r}_\cdot\right]$ for

---

[32]By Hölder's inequality, $IQ \geq IV^2$, so that the distortion due to microstructure noise in general will result in a downward bias in the limiting ratio.

[33]In the general case where ε_i is a MA(q) process: $\varepsilon_i = \nu_i + \sum_{k=1}^q \theta_k \nu_{i-k}$, we have

$$\begin{aligned}
E[\widehat{IQ}] &= E[IQ^*] + E[IV^2]\left(\theta_1^2 + \ldots + \theta_q^2 + 1\right)^2 \lambda^2 + 2E[IV^2]\left(\theta_1^2 + \ldots + \theta_q^2 + 1\right)\lambda \\
E[\widehat{IV}^2] &= E[IV^{*2}] + E[IV^2]\left(\theta_1^2 + \ldots + \theta_q^2 + 1\right)^2 \lambda^2 + 2E[IV^2]\left(\theta_1^2 + \ldots + \theta_q^2 + 1\right)\lambda + O\left(\frac{1}{N}\right)
\end{aligned}$$

[34]Note that the factor 4 arises from the multiplication by 2 in equation (38).

the full set of pre-averaged returns $\overline{r}. = \{\overline{r}_i\}_{i=1}^{N-2K+1}$ can be defined as follows:

$$\widehat{IV}\left[\overline{r}.\right] = \frac{1}{2K}\sum_{i=1}^{2K}\frac{1}{\psi}\widehat{IV}\left[\overline{S}_i\right]$$

$$\widehat{IQ}\left[\overline{r}.\right] = \frac{1}{2K}\sum_{i=1}^{2K}\frac{1}{\psi^2}\widehat{IQ}\left[\overline{S}_i\right]$$

After incorporating finite sample bias correction, these take the following final form that we use in our pre-averaged implementation of all estimators:

$$\widehat{IV}^{Adj}\left[\overline{r}.\right] = \frac{1}{2K}\sum_{i=1}^{2K}\frac{1}{\psi_K}\frac{\frac{N}{2K}}{\left\lfloor\frac{N-i+1}{2K}\right\rfloor}\widehat{IV}\left[\overline{S}_i\right]$$

$$\widehat{IQ}^{Adj}\left[\overline{r}.\right] = \frac{1}{2K}\sum_{i=1}^{2K}\frac{1}{\psi_K^2}\frac{\left(\frac{N}{2K}\right)^2}{\left\lfloor\frac{N-i+1}{2K}\right\rfloor^2}\widehat{IQ}\left[\overline{S}_i\right]$$

Consistency and asymptotic normality are clearly preserved by pre-averaging and sub-sampling, while noise-robustness improves. In particular, bounce-backs are near perfectly annihilated given that adjacent returns are subject to almost identical kernel weights. For further details on pre-averaging please refer to Podolskij and Vetter (2009) and Jacod, Li, Mykland, Podolskij and Vetter (2009) among others.

D Figures and Tables

Figure 1: **IV inference using non-jump robust RQ/RV measures versus jump-robust MedRQ/MedRV measures.** We plot prices (blue line), the IV point estimate (red line), the interquartile range (blue box) as well as two standard deviation IV confidence bands (black whiskers) for IBM for three trading days in February 2008.
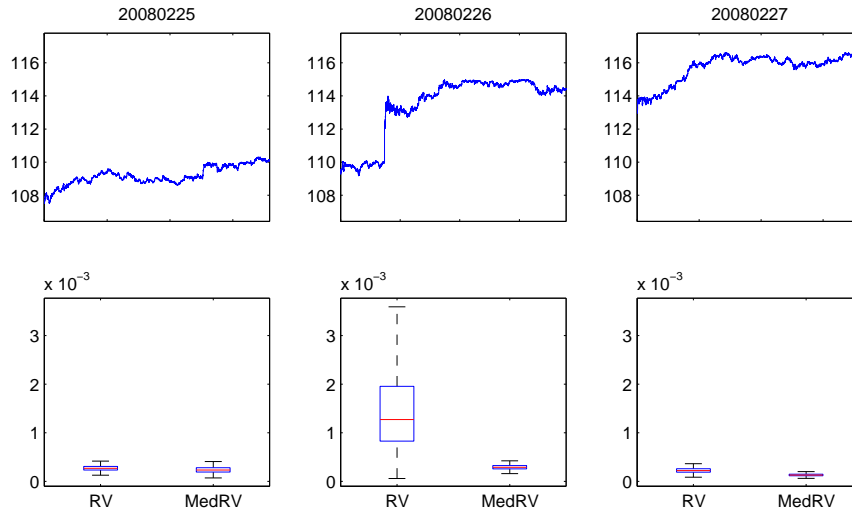
Figure 2: **Schematic representation of the construction of the RMinRQ and RMedRQ estimators of σ^4 on a block of five adjacent returns.**
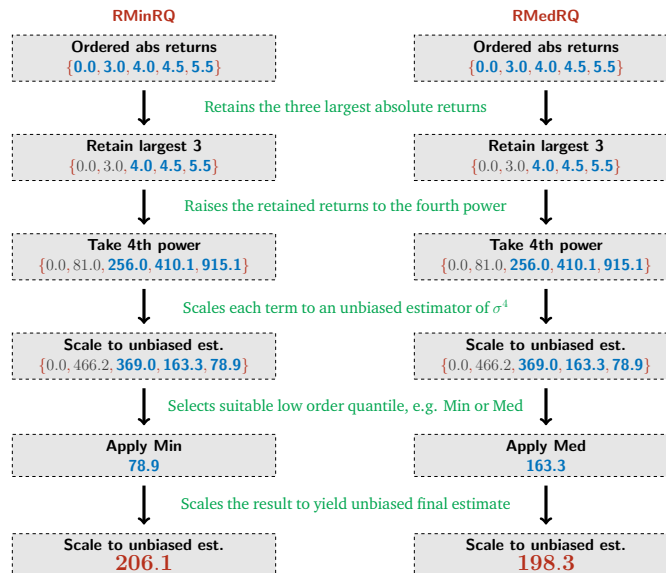
Figure 3: **IQ signature plots.** In Panel A, the estimation is performed as described in the paper and, in particular, all estimators, except TRQ and TBQ, are subject to the functional filtering procedure detailed in Section 4.3. In contrast, Panel B depicts the estimators without functional filtering. TRQ and TBQ are identical in the two panels, as they are based solely on truncation of individual returns. The average is across all stocks in the DJ30 index during January 2005-May 2007 and all estimates are pre-averaged and sub-sampled based on tick-time sampling.

Panel A: Mean of IQ for filtered estimators (truncated local functionals).

Panel B: Mean of IQ for non-filtered estimators (non-truncated local functionals).
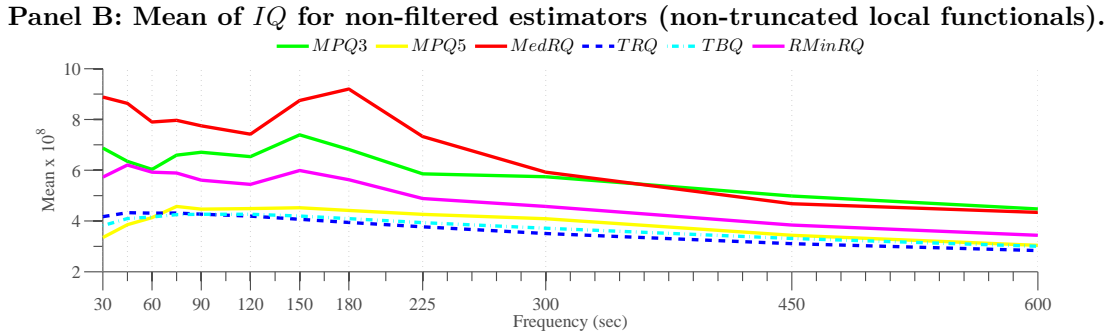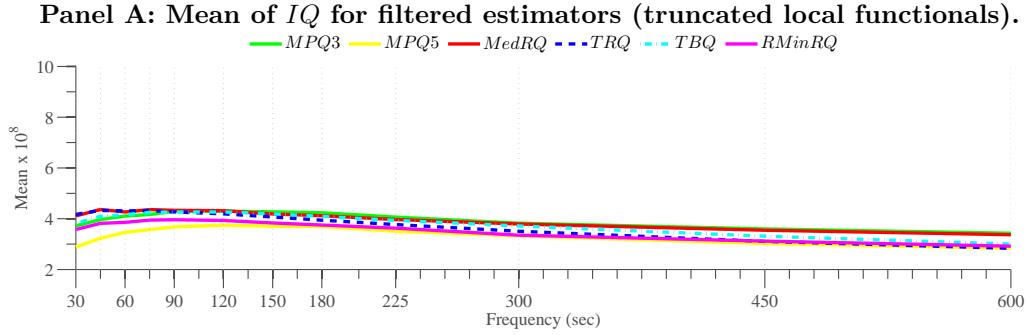
Figure 4: **Diurnal volatility pattern for intraday trade data across the DJ30 stocks between January 1, 2005 and May 31, 2007.** We plot the diurnal U-shape variance factors across stock-days based on local estimates of σ^2 in one minute buckets using tick time (Panel A) or calendar time (Panel B) sampling. On each stock-day, the factor in each one-minute bucket is computed by normalizing by the average of the 390 variance estimates on that day. The average variance factor is then computed by averaging across all stock-days.
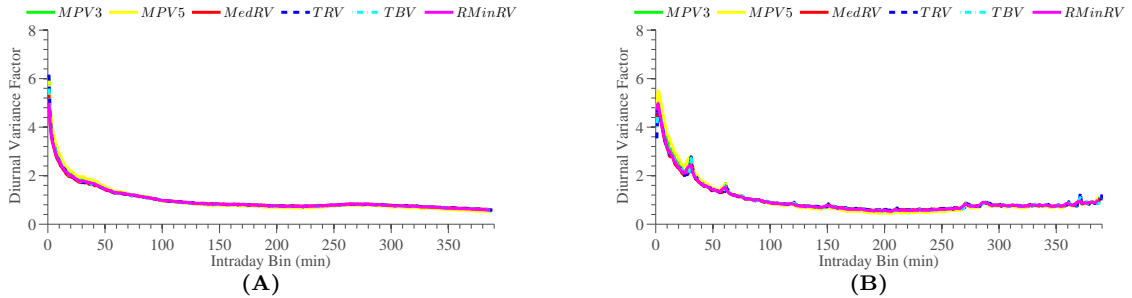
Figure 5: **Average estimates of IV across the DJ30 stocks between January 1, 2005 and May 31, 2007.**
We provide signature plots for the mean of each pre-averaged estimator of IV as a function of pre-averaging window
size matching the sampling frequency (measured in seconds on the x-axis). Panel A plots the mean across all days.
Panel B plots the mean across the top 10% days with respect to intraday variation in volatility. Panel C plots the
mean across the bottom 10% days with respect to intraday variation in volatility. Intraday variation in volatility is
measured by the VoV measure of volatility of volatility described in Section 6.
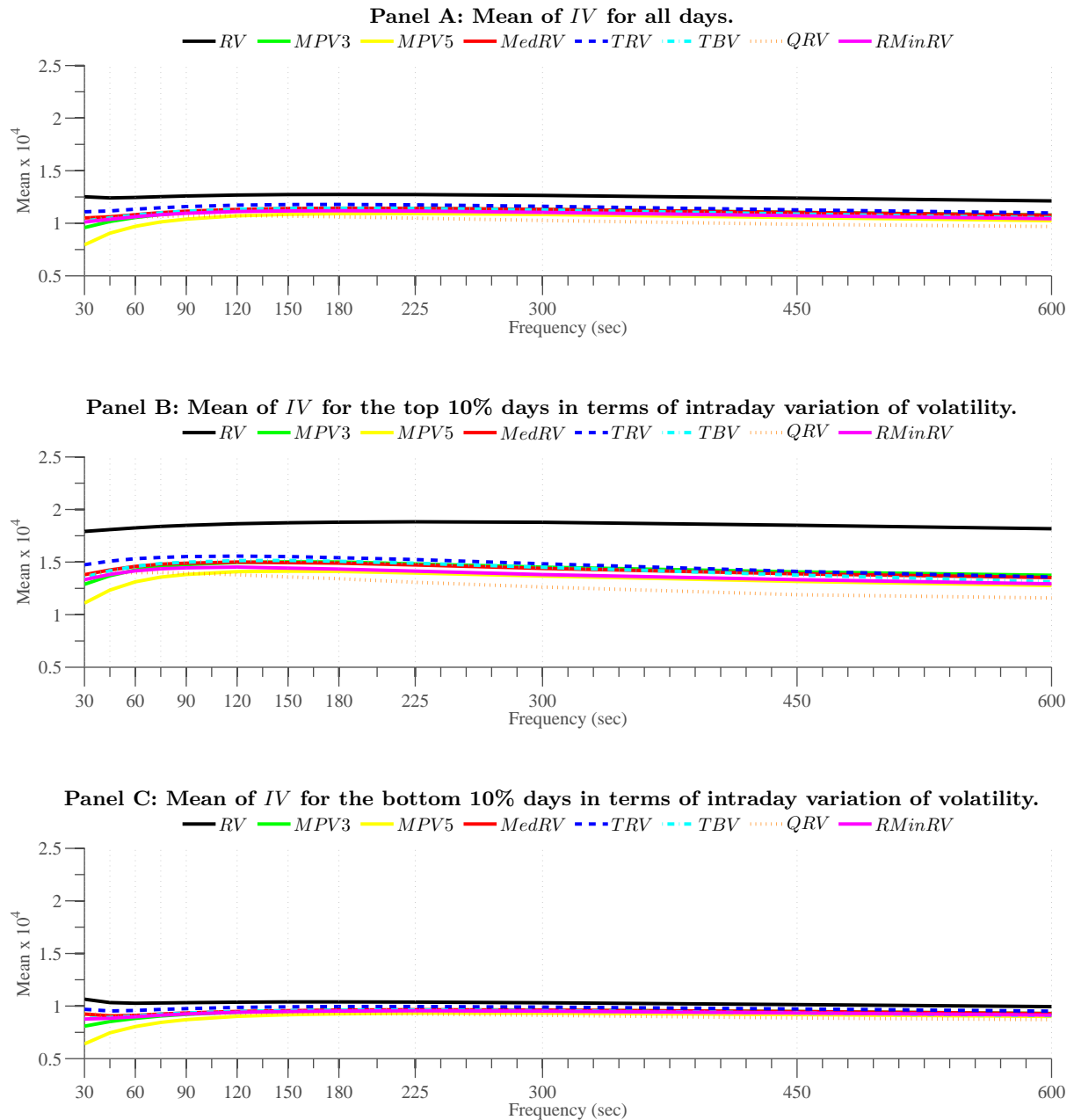
Figure 6: **Average estimates of $\sqrt{IQ}$ across the DJ30 stocks between January 1, 2005 and May 31, 2007.** We provide signature plots for the mean of each pre-averaged estimator of $\sqrt{IQ}$ as a function of pre-averaging window size matching the sampling frequency (measured in seconds on the x-axis). Panel A plots the mean across all days. Panel B plots the mean across the top 10% days with respect to intraday variation in volatility. Panel C plots the mean across the bottom 10% days with respect to intraday variation in volatility. Intraday variation in volatility is measured by the VoV measure of volatility of volatility described in Section 6.

Panel A: Mean of $\sqrt{IQ}$ for all days.

Panel B: Mean of $\sqrt{IQ}$ for the top 10% days in terms of intraday variation of volatility.

Panel C: Mean of $\sqrt{IQ}$ for the bottom 10% days in terms of intraday variation of volatility.

Figure 7: **Average estimates of $\sqrt{IQ}/IV$ across the DJ30 stocks between January 1, 2005 and May 31, 2007.** We provide signature plots for the mean of each pre-averaged estimator of $\sqrt{IQ}/IV$ as a function of pre-averaging window size matching the sampling frequency (measured in seconds on the x-axis). Panel A plots the mean across all days. Panel B plots the mean across the top 10% days with respect to intraday variation in volatility. Panel C plots the mean across the bottom 10% days with respect to intraday variation in volatility. Intraday variation in volatility is measured by the VoV measure of volatility of volatility described in Section 6.

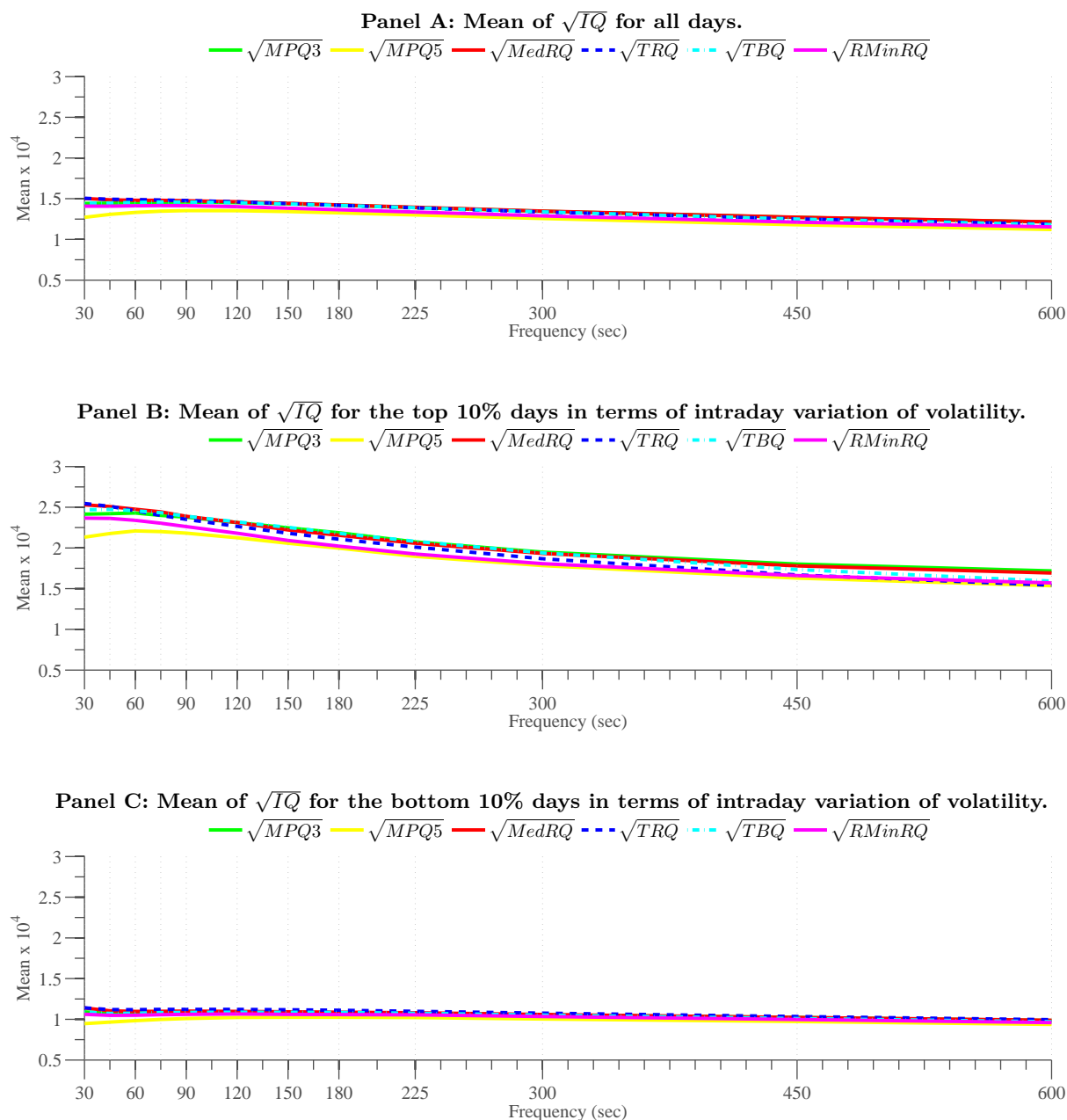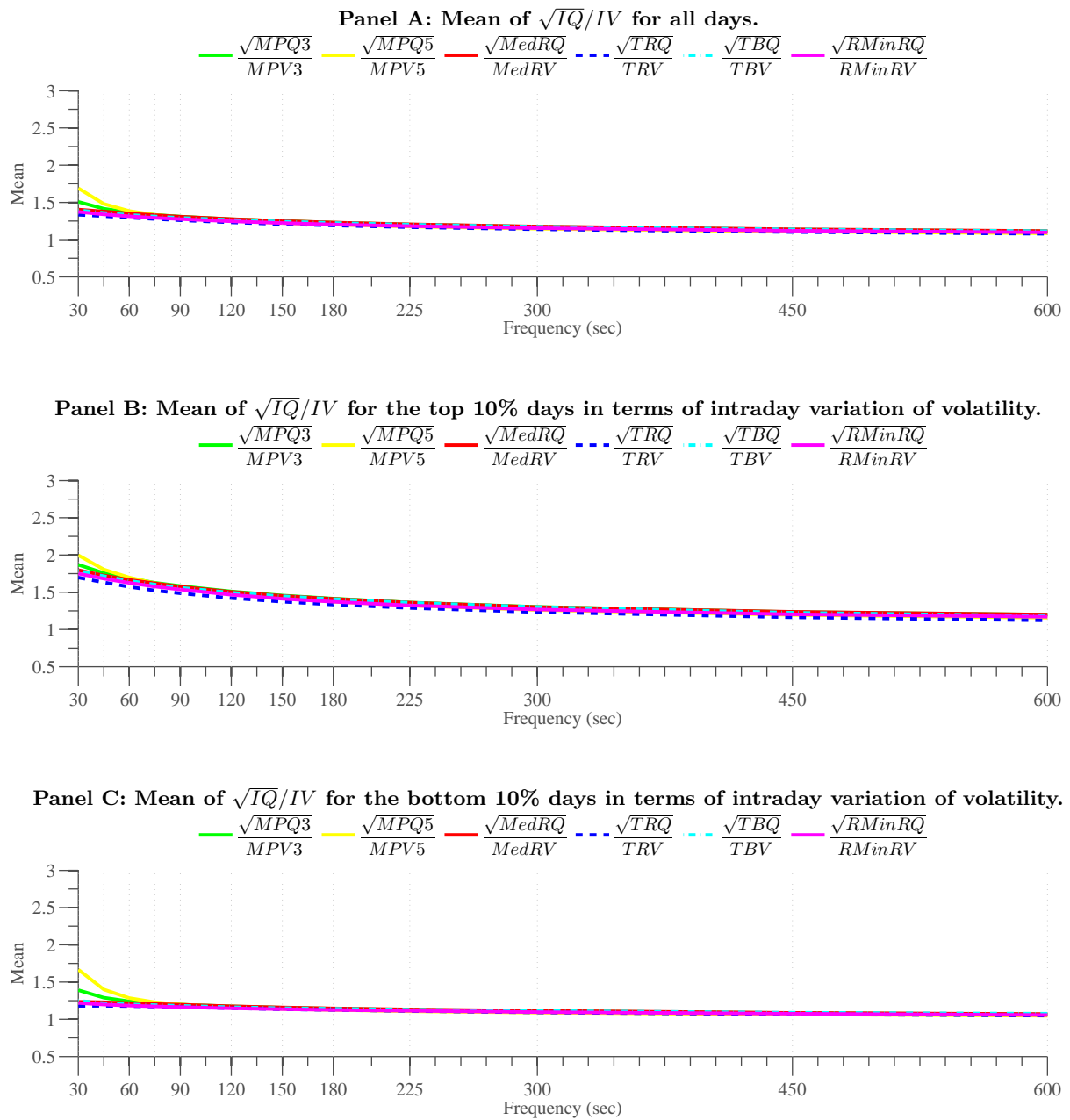Table 1: **Tabulation of Moments of Order Statistics for Standard Gaussian Return Blocks of up to Five Returns.** We compute the second and fourth moments of order statistics based on blocks of powers of independent standard normals, $Z_i \sim \mathcal{N}(0,1)$, whose inverse represent the scaling factors of the NT and RNT estimators defined in sections 2.4.3 and 3.1. Panel A: Expectation of order statistics of squared normals (NTV estimators). Panel B: Expectation of order statistics of normals raised to the 4th power (NTQ estimators). Panel C: Expectation of quantiles of rescaled squared order statistics of normals (RNTV estimators). Panel D: Expectation of quantiles of rescaled order statistics of normals raised to the 4th power (RNTQ estimators).

Panel A: 2nd moments defining the inverse scaling factors for corresponding NTV estimators

Block Size	$Z_{(1)}^2$	$Z_{(2)}^2$	$Z_{(3)}^2$	$Z_{(4)}^2$	$Z_{(5)}^2$
2	$\mu_2^{(1,2)} \approx 1.6366198$ $= \frac{\pi-2}{\pi}$	$\mu_2^{(2,2)} \approx 0.36338023$ $= \frac{2+\pi}{\pi}$			
3	$\mu_2^{(1,3)} \approx 0.19279847$ $= \frac{-6+2\sqrt{3}+\pi}{\pi}$	$\mu_2^{(2,3)} \approx 0.70454374$ $= \frac{6-4\sqrt{3}+\pi}{\pi}$	$\mu_2^{(3,3)} \approx 2.1026578$ $= 1 + \frac{2\sqrt{3}}{\pi}$		
4	$\mu_2^{(1,4)} \approx 0.12070214$ $= 1 + \frac{4(4\sqrt{3}-9)}{3\pi}$	$\mu_2^{(2,4)} \approx 0.40908747$ $= \frac{12-8\sqrt{3}+\pi}{\pi}$	$\mu_2^{(3,4)} = 1$	$\mu_2^{(4,4)} \approx 2.4702104$ $= 1 + \frac{8}{\sqrt{3}\pi}$	
5	$\mu_2^{(1,5)} \approx 0.083077313$	$\mu_2^{(2,5)} \approx 0.271201456$	$\mu_2^{(3,5)} \approx 0.61591649$	$\mu_2^{(4,5)} \approx 1.2560557$	$\mu_2^{(5,5)} \approx 2.7737491$

Panel B: 4th moments defining the inverse scaling factors for corresponding NTQ estimators

Block Size	$Z_{(1)}^4$	$Z_{(2)}^4$	$Z_{(3)}^4$	$Z_{(4)}^4$	$Z_{(5)}^4$
2	$\mu_4^{(1,2)} \approx 0.45352091$ $= 3 - \frac{8}{\pi}$	$\mu_4^{(2,2)} \approx 5.5464791$ $= 3 + \frac{8}{\pi}$			
3	$\mu_4^{(1,3)} \approx 0.13874649$ $= 3 + \frac{26-24\sqrt{3}}{\sqrt{3}\pi}$	$\mu_4^{(2,3)} \approx 1.0830697$ $= \frac{72-52\sqrt{3}+9\pi}{3\pi}$	$\mu_4^{(3,3)} \approx 7.7781838$ $= 3 + \frac{26}{\sqrt{3}\pi}$		
4	$\mu_4^{(1,4)} \approx 0.057664089$ $= 3 + \frac{4(4(13\sqrt{3}-27)\pi-9)}{9\pi^2}$	$\mu_4^{(2,4)} \approx 0.38199370$ $= 3 + \frac{4(9+(36-26\sqrt{3})\pi)}{3\pi^2}$	$\mu_4^{(3,4)} \approx 1.7841458$ $= 3 - \frac{12}{\pi^2}$	$\mu_4^{(4,4)} \approx 9.7761964$ $= 3 + \frac{4(9+26\sqrt{3}\pi)}{9\pi^2}$	
5	$\mu_4^{(1,5)} \approx 0.028554808$	$\mu_4^{(2,5)} \approx 0.17410122$	$\mu_4^{(3,5)} \approx 0.69383242$	$\mu_4^{(4,5)} \approx 2.5110214$	$\mu_4^{(5,5)} \approx 11.592490$

Panel C: 2nd moments defining the inverse scaling factors for corresponding RNTV estimators

Block Size	$\min\left(\frac{Z_{(3)}^2}{\mu_2^{(3,5)}}, \frac{Z_{(4)}^2}{\mu_2^{(4,5)}}, \frac{Z_{(5)}^2}{\mu_2^{(5,5)}}\right)$	$\mathrm{med}\left(\frac{Z_{(3)}^2}{\mu_2^{(3,5)}}, \frac{Z_{(4)}^2}{\mu_2^{(4,5)}}, \frac{Z_{(5)}^2}{\mu_2^{(5,5)}}\right)$
5	$\mu_2^{(1,(\mathbf{3,4,5}))} \approx 0.62084$	$\mu_2^{(2,(\mathbf{3,4,5}))} \approx 0.94544$

Panel D: 4th moments defining the inverse scaling factors for corresponding RNTQ estimators

Block Size	$\min\left(\frac{Z_{(3)}^4}{\mu_4^{(3,5)}}, \frac{Z_{(4)}^4}{\mu_4^{(4,5)}}, \frac{Z_{(5)}^4}{\mu_4^{(5,5)}}\right)$	$\mathrm{med}\left(\frac{Z_{(3)}^4}{\mu_4^{(3,5)}}, \frac{Z_{(4)}^4}{\mu_4^{(4,5)}}, \frac{Z_{(5)}^4}{\mu_4^{(5,5)}}\right)$
5	$\mu_4^{(1,(\mathbf{3,4,5}))} \approx 0.38303$	$\mu_4^{(2,(\mathbf{3,4,5}))} \approx 0.82367$

Table 2: **Relative bias and relative mean squared error (MSE) factors for pre-averaged estimators of IV.** We report the relative bias (left half of the table) and MSE (right half of the table) factors for different estimators of IQ for the following models: BM (Panel A), BM + 1 Jump (Panel B), BM + 4 Jumps (Panel C), SV-U (Panel D), BM + 1 Volatility Jump (Panel E), BM + IID Noise (Panel F), BM + 1 Bounceback (Panel G), BM + Sparcity (Panel H), BM + Discrete Pricing (Panel I). The rows in each panel correspond to each of the following pre-averaging window sizes: 30, 120, 600 seconds. The table columns represent the different estimators, with the abbreviations defined in the text. The relative MSE is computed using baseline window size of 120-seconds as the sample mean of $195(\widehat{IV} - IV)^2/IV^2$, where IV is the true simulated integrated variance on each day. The MSE factor for 120-sec pre-averaged RV in Panel A has the known theoretical value of $151/140 \approx 1.08$, see Jacod et al. (2009).

Estimator:	Relative Bias									Relative MSE								
	RV	MPV3	MPV5	TRV	TBV	QRV	MedRV	RMinRV	RMedRV	RV	MPV3	MPV5	TRV	TBV	QRV	MedRV	RMinRV	RMedRV
Panel A: BM																		
30 sec	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.26	0.35	0.38	0.27	0.32	0.29	0.34	0.31	0.29
120 sec	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.08	1.42	1.56	1.08	1.29	1.26	1.41	1.26	1.18
600 sec	1.00	1.00	1.00	0.99	0.97	1.00	1.00	1.00	1.00	5.52	7.51	8.44	5.53	6.60	7.91	7.47	6.87	6.42
Panel B: BM + 1 Jump																		
30 sec	1.25	1.01	1.01	1.01	1.00	1.01	1.00	1.00	1.01	37.02	0.41	0.41	0.28	0.32	0.30	0.35	0.32	0.30
120 sec	1.25	1.03	1.03	1.03	1.01	1.02	1.01	1.02	1.02	38.37	1.87	1.81	1.30	1.38	1.37	1.50	1.39	1.36
600 sec	1.25	1.08	1.06	1.08	1.03	1.06	1.05	1.06	1.07	45.63	10.71	10.79	8.47	7.66	9.93	8.80	8.79	8.85
Panel C: BM + 4 Jumps																		
30 sec	1.25	1.03	1.03	1.03	1.02	1.02	1.01	1.02	1.02	18.64	0.62	0.53	0.42	0.38	0.40	0.39	0.38	0.40
120 sec	1.25	1.07	1.06	1.07	1.05	1.06	1.04	1.05	1.06	19.95	2.78	2.48	2.47	1.94	2.16	1.93	1.99	2.20
600 sec	1.25	1.14	1.12	1.16	1.10	1.14	1.12	1.13	1.15	27.04	14.30	14.26	14.30	10.79	15.36	12.55	13.04	13.86
Panel D: SV-U																		
30 sec	1.00	1.00	0.99	1.00	0.99	0.98	1.00	0.99	0.99	0.30	0.39	0.42	0.30	0.36	0.41	0.39	0.35	0.33
120 sec	0.99	0.98	0.98	0.99	0.98	0.94	0.98	0.97	0.98	1.21	1.56	1.70	1.21	1.46	1.90	1.55	1.42	1.33
600 sec	0.98	0.95	0.93	0.96	0.92	0.90	0.95	0.93	0.93	5.78	7.38	8.25	5.74	7.13	8.31	7.34	6.89	6.48
Panel E: BM + 1 Volatility Jump																		
30 sec	1.00	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	0.40	0.53	0.58	0.41	0.48	0.51	0.52	0.47	0.44
120 sec	1.00	0.99	0.99	1.00	0.98	0.97	0.99	0.99	0.99	1.64	2.13	2.35	1.66	1.98	2.67	2.13	1.94	1.82
600 sec	0.99	0.96	0.94	0.93	0.92	0.90	0.96	0.95	0.96	8.23	10.34	11.81	8.17	9.89	16.15	10.38	10.10	9.53
Panel F: BM + IID Noise																		
30 sec	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	1.03	0.44	0.52	0.56	0.41	0.46	0.47	0.52	0.47	0.46
120 sec	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.08	1.42	1.56	1.08	1.29	1.26	1.42	1.27	1.19
600 sec	1.00	1.00	1.00	0.99	0.97	1.00	1.00	1.00	1.00	5.49	7.47	8.40	5.50	6.57	7.88	7.43	6.83	6.39
Panel G: BM + 1 Bounceback																		
30 sec	1.01	1.01	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.39	0.36	0.38	0.27	0.32	0.30	0.35	0.31	0.29
120 sec	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.08	1.42	1.56	1.08	1.29	1.26	1.42	1.26	1.19
600 sec	1.00	1.00	1.00	0.99	0.97	1.00	1.00	1.00	1.00	5.55	7.53	8.47	5.55	6.62	7.96	7.51	6.90	6.46
Panel H: BM + Sparcity																		
30 sec	1.00	0.96	0.96	0.99	0.97	0.97	0.96	0.97	0.98	0.30	0.61	0.75	0.31	0.54	0.43	0.66	0.53	0.42
120 sec	1.00	0.99	0.99	1.00	0.99	0.99	0.99	0.99	0.99	1.11	1.43	1.56	1.11	1.33	1.27	1.43	1.28	1.20
600 sec	1.00	1.00	1.00	0.99	0.97	1.00	1.00	1.00	1.00	5.55	7.49	8.41	5.56	6.60	7.94	7.45	6.87	6.44
Panel I: BM + Discrete Pricing																		
30 sec	1.02	1.00	0.94	1.02	1.01	1.02	1.02	1.01	1.02	0.34	0.36	1.08	0.32	0.35	0.37	0.43	0.34	0.36
120 sec	1.00	1.00	1.00	1.00	0.99	1.00	1.00	1.00	1.00	1.08	1.42	1.56	1.08	1.29	1.26	1.41	1.27	1.19
600 sec	1.00	1.00	1.00	0.99	0.97	1.00	1.00	1.00	1.00	5.49	7.47	8.40	5.50	6.57	7.87	7.43	6.83	6.39

Table 3: **Relative bias and relative mean squared error (MSE) factors for pre-averaged estimators of IQ.** We report the relative bias (left half of the table) and MSE (right half of the table) factors for different estimators of IQ for the following models: BM (Panel A), BM + 1 Jump (Panel B), BM + 4 Jumps (Panel C), SV-U (Panel D), BM + 1 Volatility Jump (Panel E), BM + IID Noise (Panel F), BM + 1 Bounceback (Panel G), BM + Sparcity (Panel H), BM + Discrete Pricing (Panel I). The rows in each panel correspond to each of the following pre-averaging window sizes: 30, 120, 600 seconds. The table columns represent the different estimators, with the abbreviations defined in the text. The relative MSE is computed using a baseline window size of 120-seconds as the sample mean of $195(\widehat{IQ} - IQ)^2/IQ^2$, where IQ is the true simulated integrated quarticity for each day.

Estimator:	Relative Bias							Relative MSE						
	MPQ3	MPQ5	TRQ	TBQ	MedRQ	RMinRQ	RMedRQ	MPQ3	MPQ5	TRQ	TBQ	MedRQ	RMinRQ	RMedRQ
Panel A: BM														
30 sec	1.00	1.00	0.99	0.99	1.00	1.00	1.00	1.55	1.68	1.22	1.41	1.55	1.29	1.23
120 sec	1.00	1.00	0.98	0.99	1.00	1.00	1.00	6.37	6.96	4.84	5.71	6.40	5.32	5.05
600 sec	1.00	1.00	0.97	0.96	1.00	1.00	1.00	33.81	37.87	24.11	28.37	33.72	29.06	27.46
Panel B: BM + 1 Jump														
30 sec	1.09	1.04	1.02	1.01	1.01	1.01	1.02	5.34	2.47	1.38	1.51	1.66	1.42	1.45
120 sec	1.16	1.08	1.09	1.05	1.04	1.05	1.07	22.87	11.30	8.30	7.23	7.74	7.00	7.69
600 sec	1.26	1.17	1.28	1.13	1.13	1.16	1.20	100.90	66.83	71.15	48.15	53.36	52.29	59.93
Panel C: BM + 4 Jumps														
30 sec	1.16	1.08	1.09	1.05	1.04	1.05	1.07	9.35	3.77	3.33	2.19	2.20	2.08	2.69
120 sec	1.26	1.16	1.27	1.15	1.13	1.15	1.20	31.64	16.95	24.29	13.47	14.69	13.13	19.65
600 sec	1.39	1.31	1.54	1.32	1.33	1.35	1.42	120.66	93.47	142.43	86.62	113.01	94.73	121.22
Panel D: SV-U														
30 sec	0.98	0.97	0.98	0.98	0.98	0.97	0.97	3.22	3.33	2.52	2.96	3.22	2.62	2.47
120 sec	0.94	0.91	0.94	0.94	0.94	0.91	0.91	10.45	10.38	7.80	9.52	10.42	8.38	7.98
600 sec	0.82	0.78	0.84	0.81	0.82	0.78	0.78	32.41	33.49	25.47	30.26	32.20	27.88	26.69
Panel E: BM + 1 Volatility Jump														
30 sec	0.99	0.99	0.97	0.99	0.99	0.99	0.99	4.20	4.53	3.25	3.83	4.22	3.52	3.34
120 sec	0.97	0.96	0.94	0.95	0.97	0.96	0.96	15.08	16.19	11.67	13.76	15.15	12.91	12.34
600 sec	0.88	0.83	0.75	0.79	0.88	0.85	0.86	55.05	59.42	41.28	46.56	55.39	49.76	47.93
Panel F: BM + IID Noise														
30 sec	1.06	1.06	1.04	1.05	1.06	1.06	1.06	2.38	2.52	1.65	2.08	2.38	2.05	2.01
120 sec	1.00	1.00	0.99	0.99	1.00	1.00	1.00	6.43	7.01	4.86	5.74	6.44	5.36	5.09
600 sec	1.00	1.00	0.97	0.96	1.00	1.00	1.00	33.58	37.63	23.98	28.22	33.47	28.90	27.26
Panel G: BM + 1 Bounceback														
30 sec	1.02	1.01	1.01	1.01	1.01	1.01	1.01	1.76	1.77	1.30	1.46	1.62	1.36	1.33
120 sec	1.00	1.00	0.99	0.99	1.00	1.00	1.00	6.40	6.98	4.84	5.72	6.44	5.34	5.07
600 sec	1.00	1.00	0.98	0.96	1.00	1.00	1.00	34.07	38.05	24.26	28.54	34.06	29.32	27.77
Panel H: BM + Sparcity														
30 sec	0.96	0.93	1.05	0.98	0.97	0.97	0.99	1.90	2.46	2.01	1.64	1.91	1.59	1.41
120 sec	0.99	0.98	1.00	0.99	0.99	0.99	1.00	6.49	6.92	5.22	5.91	6.58	5.47	5.23
600 sec	1.00	1.00	0.98	0.96	1.00	1.00	1.00	33.98	37.90	24.58	28.57	33.96	29.30	27.79
Panel I: BM + Discrete Pricing														
30 sec	1.04	1.02	1.02	1.03	1.04	1.03	1.04	1.92	1.83	1.35	1.69	2.00	1.53	1.58
120 sec	1.00	1.00	0.99	0.99	1.00	1.00	1.00	6.42	7.00	4.86	5.73	6.42	5.35	5.08
600 sec	1.00	1.00	0.97	0.96	1.00	1.00	1.00	33.58	37.63	23.99	28.20	33.47	28.87	27.25

Table 4: **Relative bias and relative mean squared error (MSE) factors for pre-averaged estimators of** $\sqrt{TQ}/IV$. We report the relative bias (left half of the table) and MSE (right half of the table) factors for different estimators of IQ for the following models: BM (Panel A), BM + 1 Jump (Panel B), BM + 4 Jumps (Panel C), SV-U (Panel D), BM + 1 Volatility Jump (Panel E), BM + IID Noise (Panel F), BM + 1 Bounceback (Panel G), BM + Sparcity (Panel H), BM + Discrete Pricing (Panel I). The rows in each panel correspond to each of the following pre-averaging window sizes: 30, 120, 600 seconds. The table columns represent the different estimators, with the abbreviations defined in the text. The relative MSE is computed using baseline window size of 120-seconds as the sample mean of $195(\sqrt{\widehat{IQ}/\widehat{IV}} - \sqrt{IQ/IV})^2/(\sqrt{IQ/IV})^2$, where IV and IQ are the true simulated integrated variance and integrated quarticity for each day.

			Relative Bias								Relative MSE			
Estimator:	MPQ3	MPQ5	TRQ	TBQ	MedRQ	RMinRQ	RMedRQ	MPQ3	MPQ5	TRQ	TBQ	MedRQ	RMinRQ	RMedRQ
Panel A: BM														
30 sec	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.05	0.04	0.05	0.05	0.06	0.03	0.02
120 sec	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.18	0.15	0.18	0.18	0.22	0.11	0.10
600 sec	0.98	0.98	0.98	1.00	0.98	0.98	0.98	0.76	0.64	0.75	0.76	0.89	0.48	0.42
Panel B: BM + 1 Jump														
30 sec	1.03	1.01	1.00	1.00	1.00	1.00	1.00	0.43	0.09	0.05	0.05	0.06	0.03	0.04
120 sec	1.03	1.01	1.02	1.01	1.00	1.00	1.01	1.09	0.30	0.25	0.23	0.26	0.16	0.19
600 sec	1.01	0.99	1.03	1.02	0.99	1.00	1.00	1.60	0.79	1.12	1.01	1.12	0.64	0.70
Panel C: BM + 4 Jumps														
30 sec	1.04	1.01	1.02	1.01	1.01	1.01	1.01	0.59	0.12	0.12	0.07	0.09	0.05	0.09
120 sec	1.04	1.01	1.04	1.02	1.01	1.01	1.02	1.01	0.32	0.59	0.34	0.51	0.26	0.51
600 sec	1.01	0.99	1.04	1.02	1.01	1.00	1.01	1.24	0.75	1.24	1.16	1.56	0.74	0.92
Panel D: SV-U														
30 sec	0.99	0.99	0.99	0.99	0.99	0.99	0.99	0.27	0.26	0.22	0.25	0.28	0.20	0.19
120 sec	0.98	0.97	0.97	0.98	0.98	0.97	0.97	0.71	0.66	0.56	0.64	0.75	0.53	0.50
600 sec	0.94	0.93	0.94	0.95	0.94	0.94	0.94	1.81	1.74	1.55	1.44	1.94	1.45	1.36
Panel E: BM + 1 Volatility Jump														
30 sec	1.00	0.99	0.99	0.99	1.00	0.99	1.00	0.35	0.36	0.31	0.34	0.38	0.28	0.26
120 sec	0.98	0.98	0.97	0.98	0.98	0.98	0.98	1.14	1.21	1.17	1.13	1.20	0.97	0.92
600 sec	0.93	0.91	0.90	0.93	0.93	0.92	0.93	4.07	4.75	5.14	3.88	4.22	4.02	3.75
Panel F: BM + IID Noise														
30 sec	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.05	0.04	0.05	0.05	0.05	0.03	0.02
120 sec	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.18	0.15	0.17	0.18	0.22	0.11	0.10
600 sec	0.98	0.98	0.98	1.00	0.98	0.98	0.98	0.76	0.64	0.75	0.76	0.89	0.48	0.42
Panel G: BM + 1 Bounceback														
30 sec	1.00	1.00	1.00	1.00	1.00	1.00	1.00	0.06	0.04	0.05	0.05	0.05	0.03	0.03
120 sec	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.18	0.15	0.18	0.18	0.22	0.11	0.10
600 sec	0.98	0.98	0.98	1.00	0.98	0.98	0.98	0.76	0.64	0.75	0.76	0.90	0.48	0.42
Panel H: BM + Sparcity														
30 sec	1.02	1.01	1.03	1.02	1.03	1.02	1.02	0.13	0.07	0.27	0.17	0.20	0.10	0.11
120 sec	1.00	1.00	1.00	1.01	1.00	1.00	1.00	0.19	0.16	0.17	0.20	0.24	0.12	0.11
600 sec	0.98	0.98	0.98	1.00	0.98	0.98	0.99	0.76	0.64	0.75	0.77	0.90	0.48	0.42
Panel I: BM + Discrete Pricing														
30 sec	1.02	1.07	1.00	1.00	1.00	1.00	1.00	0.11	1.06	0.05	0.05	0.05	0.03	0.02
120 sec	1.00	1.00	0.99	1.00	1.00	1.00	1.00	0.18	0.15	0.18	0.18	0.22	0.11	0.10
600 sec	0.98	0.98	0.98	1.00	0.98	0.98	0.98	0.76	0.64	0.75	0.76	0.89	0.48	0.42