ECONSTOR
Make Your Publications Visible.

A Service of

ZBW Leibniz-Informationszentrum Wirtschaft
Leibniz Information Centre for Economics

Bos, Charles S.; Janus, Pawel

Working Paper

A Quantile-based Realized Measure of Variation: New Tests for Outlying Observations in Financial Data

Tinbergen Institute Discussion Paper, No. 13-155/III

Provided in Cooperation with:
Tinbergen Institute, Amsterdam and Rotterdam

Suggested Citation: Bos, Charles S.; Janus, Pawel (2013) : A Quantile-based Realized Measure of Variation: New Tests for Outlying Observations in Financial Data, Tinbergen Institute Discussion Paper, No. 13-155/III, Tinbergen Institute, Amsterdam and Rotterdam

This Version is available at:
https://hdl.handle.net/10419/87335

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

WWW.ECONSTOR.EU

Mitglied der Leibniz-Gemeinschaft

TI 2013-155/III
Tinbergen Institute Discussion Paper

A Quantile-based Realized Measure of Variation: New Tests for Outlying Observations in Financial Data

Charles S. Bos
Pawel Janus

Faculty of Economics and Business Administration, VU University Amsterdam, Tinbergen Institute, and UBS, Statistical Risk Management.

Tinbergen Institute is the graduate school and research institute in economics of Erasmus University Rotterdam, the University of Amsterdam and VU University Amsterdam.

More TI discussion papers can be downloaded at http://www.tinbergen.nl

Tinbergen Institute has two locations:

Tinbergen Institute Amsterdam
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 1600

Tinbergen Institute Rotterdam
Burg. Oudlaan 50
3062 PA Rotterdam
The Netherlands
Tel.: +31(0)10 408 8900
Fax: +31(0)10 408 9031

Duisenberg school of finance is a collaboration of the Dutch financial sector and universities, with the ambition to support innovative research and offer top quality academic education in core areas of finance.

DSF research papers can be downloaded at: http://www.dsf.nl/

Duisenberg school of finance
Gustav Mahlerplein 117
1082 MS Amsterdam
The Netherlands
Tel.: +31(0)20 525 8579

A Quantile-based Realized Measure of Variation: New Tests for Outlying Observations in Financial Data

Charles S. Bos* and Paweł Janus†
Tinbergen Institute and Department of Econometrics & O.R.,
VU University Amsterdam
and
UBS, Statistical Risk Measurement

October 2013

Abstract

In this article we introduce a new class of test statistics designed to detect the occurrence of abnormal observations. It derives from the joint distribution of moment- and quantile-based estimators of power variation σ^r, under the assumption of a normal distribution for the underlying data.

Our novel tests can be applied to test for jumps and are found to be generally more powerful than widely used alternatives. An extensive empirical illustration for high-frequency equity data suggests that jumps can be more prevalent than inferred from existing tests on the second or third moment of the data.

Keywords: Finite activity jumps; higher order moments; order statistics; outliers; realized variation.
JEL classification codes: C10; C12; G12.

1 Introduction

It is widely acknowledged that financial returns are non-normal, as their distribution typically exhibits some degree of asymmetry and excess kurtosis. The non-Gaussian features of the return distribution can be explained by time-varying volatility and/or the presence of jumps, which are defined as outlying price increments; see Das and Sundaram (1999). There is considerable empirical evidence that volatility of assets is time-varying, but also that jumps are present in the underlying price process, see e.g. Carr and Wu (2003), Barndorff-Nielsen et al. (2006), Andersen et al. (2007), Lee and Mykland (2008), Jiang and Oomen (2008), Aït-Sahalia and Jacod (2009) and Bos, Janus, and Koopman (2012). For any particular series, there are however still two important issues: First, one has to decide whether outlying returns are indeed present in this series, and secondly the impact, if any, of these jumps on the variation of the price process must be assessed.

*Corresponding author: Charles Bos, Department of Econometrics & O.R., VU University Amsterdam, De Boelelaan 1105, NL-1081HV Amsterdam, The Netherlands. Tel +31 20 598 60 23, fax +31 20 598 60 20, email c.s.bos@vu.nl.

†The views expressed herein are those of the author and not necessarily those of UBS, which does not accept any responsibility for the contents and opinions expressed in this article.

Most of the existing tests designed to detect jumps decompose quadratic variation of the price process into a continuous part, the so-called integrated variance, and a part due to jumps. Such tests identify, through the second moment, the presence of jumps. They however do not quantify the impact of jumps on other measurements beyond integrated variance. The primary goal of this article is to provide a statistical method which can use arbitrary powers of return data, leading to formal inference on the impact of potential outlying observations also on other moments of the data. For instance, even though some jumps might not be detectable by tests built on squared returns, their possible presence still might affect higher moments.

The primary motivation for our study is the increasing use of higher moments of the data for a broad scope of financial applications; for a comprehensive review we refer to Jondeau, Poon, and Rockinger (2007). When going beyond variance as a measure of risk, the standard measures of skewness and kurtosis are mostly applied. Likewise, integrated quarticity, the variance of integrated variance, can be applied as an additional measure of risk and hence it is of considerable interest in financial applications (see Corsi, Mittnik, Pigorsch, and Pigorsch, 2008). Since it is widely agreed that unconditional non-Gaussianity of financial returns is either due to time-varying volatility and/or jumps, the standard moment-based measures quantify a total risk. However, it is of natural interest in financial economics to distinguish between the volatility and jump risk, as their implications differ from portfolio and risk management to option or bond pricing and hedging, see e.g. Bakshi, Cao, and Chen (2000) and Johannes (2004). To separate the jump risk, we consider two estimators of power variation jointly, where one estimator is robust to jumps and the other is not.

If the underlying return data, after adapting for volatility, is conditionally normally distributed (or following any other distribution known up to location and scale), the qth sample quantile equals the mean plus the standard deviation times the inverse of the appropriate standardized cumulative distribution function at probability q. It is possible to invert this relationship and solve for the standard deviation, or for the mean. Such quantile-based measures for the standard deviation for Gaussian data date back to Pearson (1920). The appealing feature of the quantile-based approach is its robustness to extreme observations, provided that the selected quantiles are sufficiently far away from the extreme tails. Christensen, Oomen, and Podolskij (2010) make use of this feature and propose to use sample quantiles to obtain a robust-to-jumps measure of integrated variance. Relative to the moment-based measure of standard deviation, efficiency of the quantile-based measure depends on the choice of sample quantiles. Mosteller (1946) was the first to suggest combining more quantiles to improve the efficiency of the quantile-based measure of standard deviation, such that it can approach the efficiency bound of the maximum likelihood estimator, while still being robust to extreme observations. Having the moment- and quantile-based measures of standard deviation, it is natural to investigate their joint distribution. The solution has been given recently by DasGupta and Haff (2006), who provided the joint asymptotic distribution of the inter-quartile range and standard deviation, and proposed a diagnostic for testing whether the underlying data is normally distributed. Their approach, however, is primarily based on the standard deviation (which is the first order power variation) and the inter-quantile range, while in this article we exploit higher orders of power variation as measured by moment- and quantile-based statistics. This will prove to provide additional power for detecting non-Gaussian observations.

Our methodological contribution can be summarized as follows. We introduce the

quantile-based measure of the rth order power variation of univariate data as an estimator of σ^r, with σ being the standard deviation and r a positive scalar, under the assumption of normality. We provide optimal quantiles for the estimator under the minimum asymptotic variance criterium, as well as weights to optimally combine multiple quantile-based sub-estimators with the aim to improve efficiency further. For our purpose, we consider the moment-based estimator of σ^r and, as a central contribution, we derive the joint limiting distribution of the two estimators. Contrary to the quantile-based estimator, the moment-based estimator is extremely sensitive to outliers. This proves useful in the construction of a Hausman (1978) type test designed to detect outlying observations, and hence to detect deviations from the null hypothesis of Gaussianity for the underlying data.

We use the quantile-based measure of variation for testing the occurrence of jumps, using a moment of arbitrary order r of returns to accentuate possible outlying observations, measuring their effect on the power variation of interest. This contrasts our setup with the one of Christensen et al. (2010), who use a sequence of similar quantile-based variation measures over short time intervals, and combine these into a daily integrated variation measure with the aim to test for jumps.

In principle, our testing procedure can be applied to any dataset with constant variance, where it is conjectured that the presence of a finite fraction of outliers leads to non-normality. In this article, we apply our new test to detect jumps in financial return data and we compare its performance to two widely applied alternatives: the tests of Barndorff-Nielsen et al. (2006) and of Jiang and Oomen (2008). Our test is less general, as it requires a normality assumption with constant variance within a time period. It can therefore be considered when financial data is sampled over a sufficiently short time horizon or on a suitably adapted time scale (cf. Christensen et al., 2010). In our application based on high-frequency data, we use a sampling scheme that, for each trading day, forms a series of intraday returns of constant variance. Our simulation results show that our test has good finite sample properties and that it can be more powerful than the alternatives, depending on the sample size and the order r. An extensive empirical illustration for fifteen NYSE equities suggests that jumps can occur more frequent than would be inferred from only the second moment of the data.

The structure of the article is as follows. In Section 2, we review the theory of sample quantiles and we present the quantile-based estimator of power variation. In Section 3, we provide the joint asymptotic distribution of the quantile- and moment-based measures of σ^r under normality. We also propose a new class of test statistics designed to detect abnormal observations. In Section 4, we study the behavior of the quantile-based estimator and resulting test statistics in a simulation study. In Section 5, we apply our theory to high frequency equity data, while Section 6 concludes. The proofs of the main results are given in Appendix A.1.

2 Measurements of variation

In this section, we first introduce both moment and maximum likelihood based measurements of variation. Afterwards, we briefly review the theory of sample quantiles, used for deriving the quantile-based estimator of σ^r. Derivation of the estimator with its asymptotic distribution is carried out assuming underlying observations are normally distributed.

2.1 Moment-based measurement of variation

First, we introduce the standard moment-based estimator of σ^r for the normal data. Let the observations $Y = \{Y_i\}_{i=1}^N$ be independently and identically normally distributed with mean μ and standard deviation σ. We use the central absolute moments of the data and define the moment-based estimator of rth order power variation as

$$\text{MPV}_N^r = \frac{1}{M^r}\frac{1}{N}\sum_{i=1}^{N}|Y_i - \mu_N|^r \qquad \text{with } \mu_N = \frac{1}{N}\sum_{i=1}^{N} Y_i \xrightarrow{p} \mu, \tag{1}$$

where $\xrightarrow{p}$ indicates convergence in probability.

$$\begin{aligned} M^r &= \int_{-\infty}^{\infty}|z|^r\phi(z)\mathrm{d}z = 2\int_0^{\infty} z^r\phi(z)\mathrm{d}z \overset{y=\frac{1}{2}z^2}{=} 2\int_0^{\infty}(2y)^{\frac{r-1}{2}}\frac{1}{\sqrt{2\pi}}e^{-y}\mathrm{d}y \\ &= \frac{2^{r/2}\Gamma\left((r+1)/2\right)}{\sqrt{\pi}} \end{aligned} \tag{2}$$

is the normalizing term, with $\phi(z)$ the density of a standard normal random variable z and $r \geq 0$. As $N \to 0$, it follows that

$$N^{1/2}\left(\text{MPV}_N^r - \sigma^r\right) \xrightarrow{d} \mathcal{N}\left(0,\, \sigma^{2r}\frac{M(r)}{(M^r)^2}\right),$$

where $\xrightarrow{d}$ denotes convergence in distribution, and

$$M(r) = M^{2r} - (M^r)^2. \tag{3}$$

2.2 Maximum likelihood measurement of variation

In Section 2.1 we specifically consider a moment-based estimator of σ^r for any positive r, as higher orders of r accentuate outliers and hence inflate their impact on a direct estimate of σ^r. Alternatively, one could have considered the maximum likelihood (ML) estimator of the variance,

$$\hat{\sigma}^2_{\text{ML}} = \frac{1}{N}\sum_{i=1}^{N}(Y_i - \mu_N)^2, \tag{4}$$

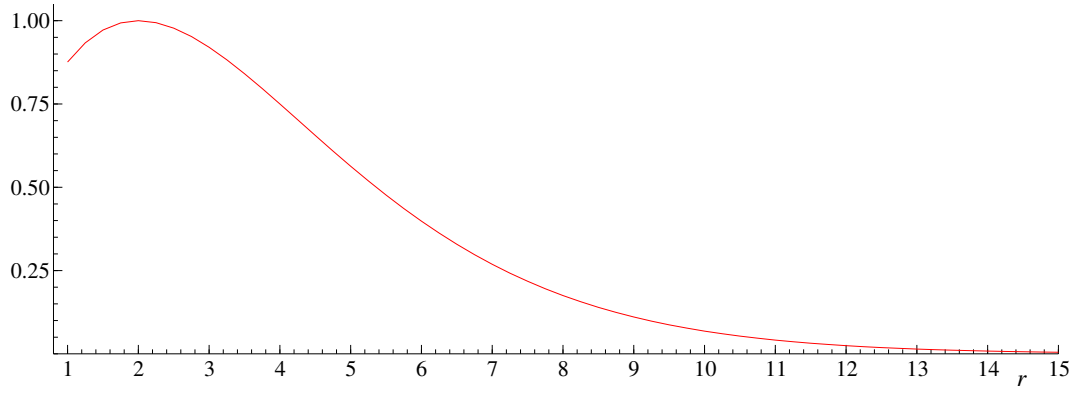
resulting in $\hat{\sigma}^r_{ML} = (\hat{\sigma}^2_{ML})^{r/2}$. As $N \to \infty$ we have that $N^{1/2}(\hat{\sigma}^r_{\text{ML}} - \sigma^r) \xrightarrow{d} \mathcal{N}(0, \frac{r^2}{2}\sigma^{2r})$. Note that $\hat{\sigma}^2_{\text{ML}} \equiv \text{MPV}_N^2$, implying that the second order moment-based estimator attains the efficiency bound of the ML estimator.

Figure 1 presents the efficiency of MPV_N^r relative to that of the corresponding ML estimator. While efficiency is still high for a small order of r, it declines dramatically once $r > 6$. As the MPV_N^r estimator is increasingly sensitive to outliers with increasing order of r, it may prove valuable in constructing a test statistic aimed at detecting potential outlying observations.

2.3 Quantiles and order statistics

Let $X = \{X_i\}_{i=1}^N$ denote an independently and identically distributed (*iid*) sequence of random variables from a univariate continuous distribution with cumulative distribution

Figure 1: Efficiency of MPV relative to ML



function (*cdf*) F and probability density function (*pdf*) f, with expectation μ and standard deviation σ. Denote the qth quantile by $Q(q) = F^{-1}(q)$ for $0 \leq q \leq 1$, implying that $q = \int_{-\infty}^{Q(q)} f(x)\mathrm{d}x$. For instance, for $q = 0.5$ we obtain the median of the variable of interest.

The simplest approach of computing the quantile $Q(q)$ empirically is by using the order statistics directly,

$$Q_N^o(q) = X_{([qN])}, \tag{5}$$

where $[\alpha]$ is the operator which rounds α to the nearest integer, and $X_{(k)}$ indicates the kth order statistic of X, such that $X_{(1)} \leq \ldots \leq X_{(k-1)} \leq X_{(k)} \leq X_{(k+1)} \leq \ldots \leq X_{(N)}$. If $q < \frac{1}{2N}$, then the order statistic $X_{(0)}$ would be needed; we define the 0th order statistic to be equal to $X_{(1)}$ for completing the above definition. Likewise, if an order statistic with $k > N$ is requested, then this will be replaced by $X_{(N)}$.

Apart from this most basic definition of an empirical quantile, Hyndman and Fan (1996) define a range of 9 different methods for calculating quantiles. For improved accuracy in cases of small samples, it is important to interpolate between order statistics, instead of rounding to the nearest one. Hence, in the remainder of this article we estimate the quantile as a weighted average between two adjacent order statistics,

$$Q_N(q) = wX_{(l)} + (1-w)X_{(l+1)}, \tag{6}$$

where the lower order statistic is chosen using $l = \lfloor (N+1)q \rfloor$, with $\lfloor \alpha \rfloor$ denoting the operator for rounding α down, and where the weight is $w = l + 1 - (N+1)q$. With this setup (which corresponds to method 6 of Hyndman and Fan (1996)), the kth order statistic corresponds to quantile $k/(N+1)$, effectively spreading out the order statistics uniformly over the quantiles $1/(N+1), \ldots, N/(N+1)$.

It is well known that the empirical quantile (irrespective of the choice of definition (5) or (6)) is asymptotically normally distributed with

$$N^{1/2}\left(Q_N(q) - Q(q)\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{q(1-q)}{[f(Q(q))]^2}\right), \quad \text{as} \quad N \to \infty, \tag{7}$$

provided $f(Q(q)) = f(F^{-1}(q)) > 0$. Actually, rather than considering the distribution of a single sample quantile, we are interested in the joint asymptotic distribution of a finite set of sample quantiles. This distribution is given in Theorem 2.1.

Theorem 2.1. *Define a p-dimensional vector* $\boldsymbol{q} = \{q_1, \ldots, q_p\}$ *with* $0 < q_1 < \ldots < q_p < 1$ *determining sample quantiles* $\boldsymbol{Q}_N(\boldsymbol{q}) = [Q_N(q_1), \ldots, Q_N(q_p)]'$ *and corresponding theoretical quantiles* $\boldsymbol{Q}(\boldsymbol{q}) = [Q(q_1), \ldots, Q(q_p)]'$*, of the random variable* X *defined before. The joint asymptotic distribution of multiple quantiles is given by the p-variate normal distribution,*

$$N^{1/2}\left(\boldsymbol{Q}_N(\boldsymbol{q}) - \boldsymbol{Q}(\boldsymbol{q})\right) \xrightarrow{d} \mathcal{N}\big(\mathbf{0}, \sigma^2\boldsymbol{\Sigma}(\boldsymbol{q})\big), \tag{8}$$

where the ij*th element of the* $p \times p$ *covariance matrix* $\boldsymbol{\Sigma}(\boldsymbol{q})$ *of standardized quantiles is given by*

$$\Sigma_{ij}(q_i, q_j) = \big(\boldsymbol{\Sigma}(\boldsymbol{q})\big)_{ij} = \big(\boldsymbol{\Sigma}(\boldsymbol{q})\big)_{ji} = \frac{q_i(1-q_j)}{\sigma^2 f(Q(q_i))f(Q(q_j))}, \qquad \textit{for} \quad i \leq j;\ i, j = 1, \ldots, p. \tag{9}$$

Proof. See Walker (1968).

Note that the matrix $\boldsymbol{\Sigma}(\boldsymbol{q})$ refers to the covariance of quantiles of the standardized random variables $(X_i - \mu)/\sigma$, hence the adaptation for σ. For simplicity in notation, the explicit dependence of $\boldsymbol{\Sigma}(\boldsymbol{q})$ on the *pdf* and inverse *cdf* is omitted. In case the observations are assumed to be distributed according to the Gaussian density, (9) simplifies to

$$\Sigma_{ij}(q_i, q_j) = \frac{q_i(1-q_j)}{\phi(\Phi^{-1}(q_i))\phi(\Phi^{-1}(q_j))}, \qquad \text{for} \quad i \leq j;\ i, j = 1, \ldots, p, \tag{10}$$

where Φ^{-1} denotes the inverse *cdf* of a standard normal variable.

2.4 Quantile-based measurement of variation and its properties

Apart from the moment-based measures of data variation, many others have been developed. For an excellent history with a particular focus on early developments, see David (1998). For example, the inter-quantile range, i.e. the range between the quantiles of order q and $1-q$, is one of the oldest measures of dispersion of data. The inter-quantile range of any (fully specified, up to location and scale) symmetric distribution is a known multiple of the standard deviation. In the following, we will discuss results for the Gaussian distribution, as this is the distribution that will be of use in the empirical application in Section 5.

Assumption 1. *The observations* $Y = \{Y_i\}_{i=1}^N$ *are independently and identically normally distributed with mean* μ *and standard deviation* σ*.*

For such observations Y, the relationship between the sample quantile and the standard deviation is given by

$$Q(q) = \mu + \sigma\Phi^{-1}(q). \tag{11}$$

If we consider another quantile, e.g. the $(1-q)$th for symmetry, we can solve for the standard deviation even when the location μ is unknown, as in

$$\sigma = \frac{1}{c(q)}\Big[Q(1-q) - Q(q)\Big], \tag{12}$$

with

$$c(q) = \Phi^{-1}(1-q) - \Phi^{-1}(q) = 2\Phi^{-1}(1-q) \tag{13}$$

the scaling term corresponding to the standard normal distribution. Without loss of generality we assume that $q < \frac{1}{2}$. Using empirical quantiles, this delivers us our quantile-based first order power variation estimator,

$$\text{QPV}_N(q) = \frac{1}{c(q)}\Big[Q_N(1-q) - Q_N(q)\Big] = \frac{1}{c(q)} D' \boldsymbol{Q}_N(\vec{\boldsymbol{q}}), \tag{14}$$

where $\boldsymbol{Q}_N(\vec{\boldsymbol{q}})$ is the vector of quantile estimators evaluated at $\vec{\boldsymbol{q}} = [q, 1-q]'$, and $D = [-1, 1]'$. Such a quantile-based estimator of first order power variation for σ was first proposed by Pearson (1920). He proposed to use $q = 1/14 \approx 0.0714$, as a value of the quantile which would result in a relatively efficient estimator of σ. Benson (1949) examined the variance of (14) for different values of q and found that efficiency is greatest for $q = 0.0692$. Mosteller (1946) showed the asymptotic normality of the estimator (14) and he also suggested to combine multiple quantiles to improve the efficiency further. Ogawa (1951) further improved the quantile-based measure of σ by deriving a linear combination of a specified number of sub-estimators, based on symmetric quantiles as in (14), which are optimally weighted for maximum efficiency. Recently Vergote (2008, ch. 3) used similar ideas on the robustness of quantiles to estimate volatility in a Brownian bridge setting using quantiles and subestimators. All these authors stress the robustness to more extreme observations of the quantile-based approach which makes it appealing relative to the moment-based approach. Also, in practical situations, it may be the case that normality is a good approximation to a central region of empirical density, say except for 1% or 3% of observations in each tail. In such a case, the robustness property inherent to the quantile-based approach may prove extremely useful.

For our purpose, we more generally consider estimation of any positive power of σ, i.e. σ^r for $r \in \Re^+$, using the sample quantiles of the data. Along with the moment-based estimator of σ^r, the quantile-based estimator will be used in Section 3 to construct a new class of test statistics for detecting outlying (or abnormal) observations. For that purpose, we introduce an estimator that is built on the rth power of the estimator in (14).

Since the quantile-based estimator of first order power variation (14) is constructed through a differentiable function of two empirical quantiles, the limiting distribution can be derived using the general results of Section 2.3 and application of the delta method. We obtain

$$N^{1/2}\left(\text{QPV}_N(q) - \sigma\right) \xrightarrow{d} \mathcal{N}\left(0, \frac{\sigma^2}{c^2(q)} D'\Sigma(\vec{\boldsymbol{q}})D\right) = \mathcal{N}\left(0, \sigma^2\xi(q)\right), \tag{15}$$

where

$$\xi(q) = \frac{1}{c^2(q)} D'\Sigma(\vec{\boldsymbol{q}})D = \frac{q(1-2q)}{2\left[\Phi^{-1}(q)\phi(\Phi^{-1}(q))\right]^2}$$

is the variance of the standardized QPV estimator (i.e., for observations distributed with variance $\sigma^2 = 1$), and $\Sigma(\vec{\boldsymbol{q}})$ constructed as in (10), using the *cdf* and *pdf* of the Gaussian density. For estimators of higher orders of σ, the limiting distribution of $\text{QPV}_N^r(q) \equiv [\text{QPV}_N(q)]^r$ follows from further application of the delta method, and is

$$N^{1/2}\left(\text{QPV}_N^r(q) - \sigma^r\right) \xrightarrow{d} \mathcal{N}\left(0, r^2\sigma^{2r}\xi(q)\right), \tag{16}$$

implying that asymptotically the variance of $\text{QPV}_N^r(q)$ is a factor $r^2\sigma^{2(r-1)}$ higher than the variance of $\text{QPV}_N^1(q)$. It follows that the asymptotic variance of the $\text{QPV}_N^r(q)$ for any r in (16) is minimized for $q = 0.0692$. For instance, when $r = 2$ we obtain the asymptotic variance of $\text{QPV}_N^2(0.0692)$ equal to approximately $3.07\sigma^4$. The variance of the likelihood-based estimator of the sample variance equals $2\sigma^4$, thus $\text{QPV}_N^2(0.0692)$ has relative efficiency of around 0.65. With the optimal, minimum variance choice of q, the estimator retains robustness to almost 7% of outliers in each tail (at least asymptotically, as will be seen in Section 2.6).

The density in (16) gives the asymptotic result for the higher order quantile power variation. In small samples, the transformation of the (unbiased) QPV_N^1 estimator to the higher-order QPV_N^r may lead to a bias, as will be seen in a simulation exercise in Section 4 and especially in Table 2. For samples of larger sizes, this bias is found to disappear swiftly.

2.5 Improving efficiency by combining sub-estimators

As already suggested by Mosteller (1946), the efficiency of the quantile-based estimator can be improved by using more pairs of quantiles. We select a p-vector of quantiles $\boldsymbol{q} = [q_1, \ldots, q_p]'$ with $0 < q_1 < \ldots < q_p < \frac{1}{2}$. For each symmetric pair of quantiles, i.e. q_i and $1 - q_i$, we can compute the $\text{QPV}_N(q_i)$ as in (14) that can be combined into the estimator using an averaging scheme. The quantile-based measurement of power variation based on multiple pairs is given by

$$\text{QPV}_N(\boldsymbol{q}, \boldsymbol{\lambda}) = \sum_{i=1}^{p} \lambda_i \ \text{QPV}_N(q_i) = \boldsymbol{\lambda}' \, \text{QPV}_N(\boldsymbol{q}), \tag{17}$$

where $\lambda_i \in (0, 1]$ is a weight assigned to the ith sub-estimator, such that $\sum_{i=1}^{p} \lambda_i = 1$, and $\boldsymbol{\lambda} = [\lambda_1, \ldots, \lambda_p]'$ is the p-vector collecting the weights. The use of multiple quantiles leads to an improvement of the efficiency of the estimator. The choice of quantiles $\boldsymbol{q}$ and the corresponding vector of weights $\boldsymbol{\lambda}$ can be optimized such that a minimum variance estimator $\text{QPV}_N(\hat{\boldsymbol{q}}, \hat{\boldsymbol{\lambda}})$ is obtained. The quantile-based estimator using multiple pairs is constructed by using a function of multiple quantiles, thus similarly the limiting distribution is obtained by a direct application of the results from Section 2.3.

Proposition 1. *Suppose Assumption 1 is satisfied. Consider a p-vector of quantiles* $\boldsymbol{q} = [q_1, \ldots, q_p]'$ *with* $0 < q_1 < \ldots < q_p < \frac{1}{2}$. *Combine each quantile* q_i *with its symmetric complement* $1 - q_i$, *and collect these into the* $2p$*-vector* $\vec{\boldsymbol{q}} = [q_1, q_2, \ldots, q_p, 1 - q_p, \ldots, 1 - q_2, 1 - q_1]'$, *with quantiles in ascending order. It follows that the p-vector of* QPV *estimators can be constructed as*

$$\text{QPV}_N(\boldsymbol{q}) = C^{-1}(\boldsymbol{q}) D_p' \boldsymbol{Q}_N(\vec{\boldsymbol{q}})$$

with $C(\boldsymbol{q})$ *the* $p \times p$ *matrix with elements* $c(q_i)$ *from* (13) *on the diagonal,* $D_p = [-I_p, J_p]'$ *combining the unit matrix* I_p *and the exchange matrix* J_p, *to collect the elements of the* $2p$ *vector of quantile estimators* $\boldsymbol{Q}_N(\vec{\boldsymbol{q}})$. *This leads to*

$$N^{1/2} \left(\text{QPV}_N(\boldsymbol{q}, \boldsymbol{\lambda}) - \sigma \right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^2 \boldsymbol{\lambda}' \, \Xi(\boldsymbol{q}) \boldsymbol{\lambda}\right),$$

where

$$\Xi(\boldsymbol{q}) = C^{-1}(\boldsymbol{q}) D_p' \Sigma(\vec{\boldsymbol{q}}) D_p C^{-1}(\boldsymbol{q}) \tag{18}$$

is the $p \times p$ variance-covariance matrix of standardized QPV estimators, based on the $2p \times 2p$ asymptotic variance-covariance matrix of standardized sample quantiles $\Sigma(\vec{\boldsymbol{q}})$ as before.
Proof. This result follows straightforwardly from Theorem 2.1.

In the test statistic proposed in Section 3 the quantile-based estimator of the rth power variation is used, based on a weighted average of p single-quantile QPV estimators. For completeness, this estimator and its asymptotic density are given by

$$\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda}) \equiv \boldsymbol{\lambda}' \, \mathrm{QPV}_N^r(\boldsymbol{q}), \tag{19}$$

$$N^{1/2} \left(\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda}) - \sigma^r \right) \xrightarrow{d} \mathcal{N}\left(0, r^2 \sigma^{2r} \boldsymbol{\lambda}' \, \Xi(\boldsymbol{q}) \boldsymbol{\lambda}\right). \tag{20}$$

Table 1 reports optimal values of quantiles $\boldsymbol{q}$ and weights $\boldsymbol{\lambda}$ for the number of pairs up to $p = 15$. For instance, for $p = 2$ the optimal estimator combines two sub-estimators (14) constructed on symmetric pairs of quantiles $(0.0230; 0.9770)$ and $(0.1271, 0.8729)$, combined as in (17) with weights 0.4604 and 0.5396 respectively. We note that the optimal values of $\boldsymbol{q}$ and $\boldsymbol{\lambda}$ for $p = 3$ were first obtained by Ogawa (1951), and later extended by Eisenberger and Posner (1965) to include up to $p = 10$ pairs. Since the variance of $\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})$ is $r^2\sigma^{2(r-1)}$ larger than for $\mathrm{QPV}_N^1(\boldsymbol{q}, \boldsymbol{\lambda})$, the choice of the power r does not influence the optimal (in the minimum variance sense) values of $\boldsymbol{q}$ and $\boldsymbol{\lambda}$. It is seen from Table 1 that for the minimum variance estimator, the robustness property deteriorates as the optimal value of q_1 decreases to zero when the number of pairs p increases. This is intuitive, as there is more information about the scale in the tails than in the central part of the distribution. Hence, a gain in efficiency in this case leads to the drawback of a lower degree of robustness. Therefore we investigate the effect of small deviations from the optimal values of $\boldsymbol{q}$ on the efficiency of the estimators.

Panel *i*) of Figure 2 presents the efficiency of $\mathrm{QPV}_N(\boldsymbol{q}, \boldsymbol{\lambda})$ relative to the maximum likelihood-based estimator of σ for a number of pairs up to $p = 15$. It is seen that for one pair the efficiency of the QPV_N estimator is around 65% of that of the ML estimator. As we combine more sub-estimators, the value of the minimized variance approaches the lower bound of the likelihood-based estimator rapidly and consequently the efficiency approaches unity. Panel *ii*) plots the variance of $\mathrm{QPV}_N(\boldsymbol{q}, \boldsymbol{\lambda})$ as a function of q_1, for $p \in \{1, \ldots, 5\}$. The quantiles $q_2, \ldots, q_p$ and weights $\boldsymbol{\lambda}$ have been reoptimized for this non-optimal value of q_1. It is seen that for q_1 increasing, the variance curve is relatively flat near the maximum efficiency point. This is very attractive from a practical viewpoint, as it means that increasing q_1 by e.g. 2%-5% leads to only a moderate loss of efficiency, while augmenting robustness. This property also extends to higher quantiles. In panel *iii*) we present the efficiency surface of the estimator combining $p = 3$ sub-estimators as a function of q_1 and q_2. Similarly, the efficiency surface is reasonably flat near the maximum value as q_1 and/or q_2 increase. Panels *ii*) and *iii*) therefore nicely illustrate that gaining additional robustness against a larger fraction of outlying observations can be obtained in exchange for a relatively small drop in efficiency of the quantile-based estimator.

2.6 Finite sample corrections

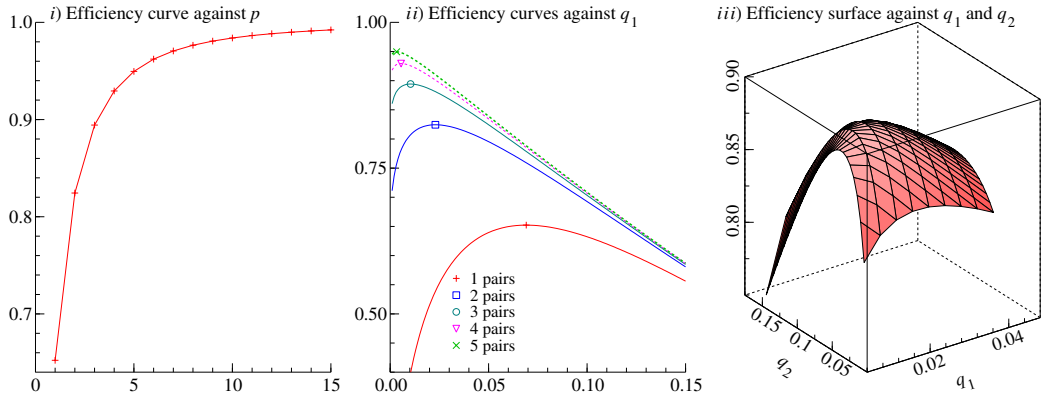
2.6.1 Adapting the bias in the scaling term $c(q)$

The quantile-based estimator (14) depends on the scaling term $c(q) = 2\Phi^{-1}(1 - q)$, to adapt the empirical distance between quantiles for the expected distance according

Table 1: Optimal values of $\boldsymbol{q}$ and $\boldsymbol{\lambda}$ for the minimum variance estimator of σ^r

	$p=1$	2	3	4	5	6	7	8	9	10	11	12	13	14	15
q_1	0.0691	0.0230	0.0104	0.0055	0.0033	0.0021	0.0014	0.0010	0.0008	0.0006	0.0004	0.0003	0.0003	0.0002	0.0002
q_2		0.1271	0.0548	0.0287	0.0169	0.0108	0.0073	0.0052	0.0038	0.0029	0.0022	0.0018	0.0014	0.0012	0.0010
q_3			0.1696	0.0851	0.0492	0.0311	0.0209	0.0148	0.0108	0.0081	0.0063	0.0050	0.0040	0.0032	0.0027
q_4				0.2017	0.1120	0.0695	0.0463	0.0324	0.0236	0.0177	0.0137	0.0108	0.0086	0.0070	0.0058
q_5					0.2269	0.1354	0.0886	0.0615	0.0445	0.0333	0.0256	0.0200	0.0160	0.0130	0.0107
q_6						0.2472	0.1558	0.1063	0.0762	0.0566	0.0433	0.0338	0.0270	0.0219	0.0180
q_7							0.2640	0.1737	0.1224	0.0901	0.0685	0.0533	0.0424	0.0343	0.0281
q_8								0.2782	0.1894	0.1372	0.1033	0.0799	0.0633	0.0510	0.0417
q_9									0.2903	0.2033	0.1507	0.1156	0.0909	0.0730	0.0595
q_{10}										0.3008	0.2158	0.1630	0.1271	0.1014	0.0824
q_{11}											0.3100	0.2270	0.1744	0.1379	0.1113
q_{12}												0.3182	0.2371	0.1849	0.1480
q_{13}													0.3255	0.2464	0.1945
q_{14}														0.3321	0.2547
q_{15}															0.3381
λ_1	1.0000	0.4604	0.2541	0.1566	0.1040	0.0730	0.0533	0.0403	0.0313	0.0248	0.0200	0.0164	0.0136	0.0115	0.0097
λ_2		0.5396	0.3979	0.2771	0.1973	0.1448	0.1094	0.0846	0.0669	0.0539	0.0440	0.0365	0.0306	0.0260	0.0222
λ_3			0.3480	0.3203	0.2559	0.2004	0.1578	0.1258	0.1016	0.0832	0.0690	0.0579	0.0490	0.0419	0.0361
λ_4				0.2460	0.2584	0.2263	0.1895	0.1574	0.1309	0.1094	0.0923	0.0784	0.0671	0.0578	0.0503
λ_5					0.1843	0.2116	0.1980	0.1746	0.1510	0.1297	0.1117	0.0963	0.0836	0.0728	0.0638
λ_6						0.1440	0.1762	0.1732	0.1590	0.1420	0.1255	0.1105	0.0972	0.0858	0.0759
λ_7							0.1158	0.1487	0.1521	0.1442	0.1322	0.1196	0.1074	0.0962	0.0862
λ_8								0.0954	0.1272	0.1343	0.1307	0.1228	0.1129	0.1033	0.0938
λ_9									0.0801	0.1101	0.1192	0.1186	0.1137	0.1065	0.0986
λ_{10}										0.0683	0.0963	0.1066	0.1080	0.1050	0.0998
λ_{11}											0.0591	0.0850	0.0957	0.0987	0.0974
λ_{12}												0.0516	0.0756	0.0865	0.0904
λ_{13}													0.0455	0.0676	0.0785
λ_{14}														0.0405	0.0609
λ_{15}															0.0363

Figure 2: Efficiency of $\mathrm{QPV}_N(\boldsymbol{q}, \boldsymbol{\lambda})$ relative to the maximum likelihood estimator



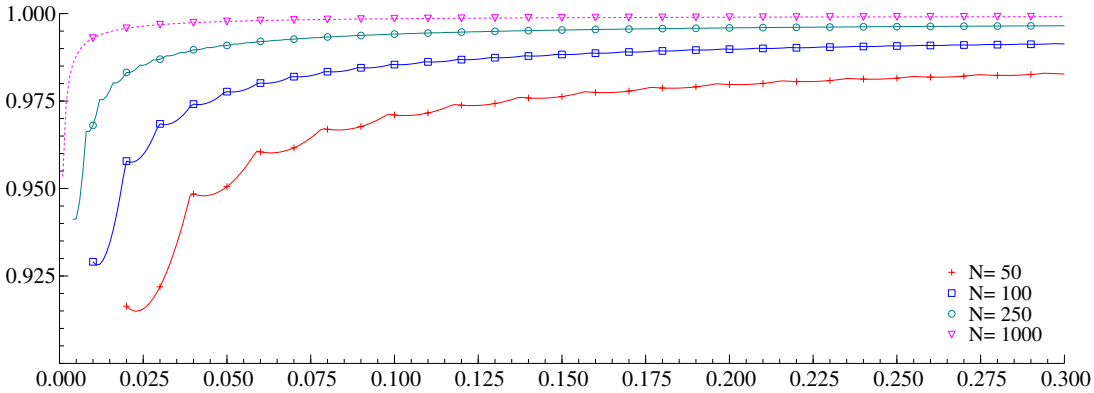
to the standard normal distribution. For small samples it is preferable to replace the asymptotic quantiles $\Phi^{-1}(q) \equiv Q^{\mathcal{N}}(q)$ by a term determined by the expected value of the quantile for the standard normal distribution. Using quantile definition (6), this would deliver

$$\overline{Q}_N^{\mathcal{N}}(q) = w\bar{Z}_{(l)} + (1-w)\bar{Z}_{(l+1)}, \tag{21}$$

with l and w as is (6), and where $\bar{Z}_{(i)}$ denotes the expected value of the ith order statistic of the independent standard normal variables $z_1, \ldots, z_N$. For a given finite N, $\bar{Z}_{(i)}$ can be obtained as

$$\bar{Z}_{(i)} = i\binom{N}{i}\int_{-\infty}^{\infty} z\big[1-\Phi(z)\big]^{N-i}\big[\Phi(z)\big]^{i-1}\phi(z)\mathrm{d}z, \tag{22}$$

Figure 3: Finite sample correction $\Phi^{-1}(q)/\overline{Q}_N^{\mathcal{N}}(q)$



When the number of available observations N is finite and relatively small, there are noticeable differences between the expected and limiting values of the quantiles used within the scaling term. This is illustrated in Figure 3 where the ratio between the asymptotic and final sample terms $\Phi^{-1}(q)/\overline{Q}_N^{\mathcal{N}}(q)$ is given for $N \in \{50, 100, 250, 1000\}$ and for values of $q \subset (0, 0.3]$. Note that the difference between the asymptotic and finite sample scaling term is especially important at the more extreme quantiles, and for sample sizes less than or equal to $N = 100$. For $N = 50$ it is also clearly visible how the linear interpolation between adjacent quantiles of (21) leads to additional movement in the finite sample correction that is needed. The use of the asymptotic scaling term in small samples therefore induces a substantial bias, even if Assumption 1 is satisfied. In the remainder of the paper, for sample sizes $N \leq 1000$ we apply a scaling term $c_N(q) = 2\overline{Q}_N^{\mathcal{N}}(q)$ based on the expected quantiles $\overline{Q}_N^{\mathcal{N}}(q)$ as defined in (21).

2.6.2 Adapting the bias in small sample higher power QPV^r estimators

By construction, the QPV_N estimator of (14) is unbiased, at least when the quantile estimator in the scaling term is adapted for the sample size, see Section 2.6.1. Higher order estimators QPV_N^r, with $r > 1$, however are only asymptotically unbiased, not in small samples.

The bias of the QPV_N^r estimator due to the non-linear transformation of the first-order estimator is

$$B_N^r(q) = \int_0^\infty v^r f_{\text{QPV}_N(q)}(v)dv - \sigma^r,$$

where $f_{\text{QPV}_N(q)}(v)$ is the (small sample) density of the first order QPV estimator.

Assuming that a truncated normal approximation to the small sample density of $\text{QPV}_N^1(q)$ holds, the bias can be quantified as

$$B_N^r(q) \approx \int_0^\infty v^r f_{\text{Tr}\mathcal{N}}(v; m, s^2)dv - \sigma^r,$$

where

$$m = \hat{\sigma}, \qquad\qquad s^2 = N^{-1/2}\hat{\sigma}^2\xi(q).$$

For $\hat{\sigma}$ we choose $\text{QPV}_N^1(\boldsymbol{q}, \boldsymbol{\lambda})$ (rather than e.g. MPV^r) as the most robust available estimate of the standard deviation before bias correction. Note that theoretically the normal approximation should be truncated at 0, as $\text{QPV}_N^1(q) > 0$ by definition. In practice it is found that the truncation is non-binding, as no appreciable mass of the normal approximating density is found below zero.

In the remainder, each $\text{QPV}_N^r(q)$ estimator will be adapted for this bias $B_N^r(q)$.

2.6.3 Further bias due to non-normality

The asymptotically unbiased estimator of σ constructed from a linear combination of sub-estimators as in (17) relies on the assumption of the normal distribution of underlying observations. However, if the assumption of normality does not hold as a fraction of outliers is found in the data, then the ordering of the normal order statistics is altered, which in turn can also lead to a bias. If we know the fraction of outliers, then we are able to bias-correct the estimator by suitably adopting the scaling term $c_N(q)$ further. We investigate this issue in the simulation study in Section 4, but will not adapt for this bias as in practice the fraction of outliers is not known. First, in the next section, we focus on the joint limiting distribution of the MPV and QPV estimators of σ^r.

3 Derivation of the outlier test

In Section 3.1, we first provide the joint asymptotic distribution of the quantile- and moment-based measurements of σ^r. This joint distribution of the two estimators facilitates testing for possible presence of outlying observations in the data. The testing procedure can be carried out through the three test statistics that we propose in Section 3.2.

3.1 Joint asymptotics of the two measurements of variation

Starting at the moment- and quantile-based measures of σ^r, it is natural to derive their joint distribution with the aim to develop formal test statistics. Under normality of the underlying observations, the resulting joint limiting distribution is shown to be bivariate normal. The joint distribution facilitates useful practical diagnostics based on these two estimators. DasGupta and Haff (2006) were the first to present the joint

asymptotic distribution of the inter-quartile range and standard deviation for normally, uniformly and exponentially distributed underlying observations. Their focus is on the correlation value between these two measures, both for fixed samples and asymptotically, as both inter-quartile range and standard deviation are two well accepted measurements of data dispersion. Based on the joint asymptotic distribution, DasGupta and Haff (2006, Corollary 3) also provide a rule of thumb for a diagnostic to see whether the underlying data are normally distributed. Here however we focus on different power variations of the data measured by our two, robust and non-robust, estimators.

Proposition 2. *Suppose Assumption 1 is satisfied. Let* $\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})$ *be the quantile-based estimator of* σ^r *as in* (19) *and let* MPV_N^r *be the moment-based estimator of* σ^r *as in* (1)*. Then as* $N \to \infty$*, we have*

$$N^{1/2}\left(\begin{pmatrix}\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})\\ \mathrm{MPV}_N^r\end{pmatrix} - \begin{pmatrix}\sigma^r\\ \sigma^r\end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(0, \sigma^{2r}\Omega(\boldsymbol{q}, \boldsymbol{\lambda}, r)\right), \tag{23}$$

$$\Omega(\boldsymbol{q}, \boldsymbol{\lambda}, r) = \boldsymbol{G}(r)'\,\boldsymbol{\Lambda}(\boldsymbol{\lambda})'\,\tilde{\Xi}(\boldsymbol{q}, r)\,\boldsymbol{\Lambda}(\boldsymbol{\lambda})\,\boldsymbol{G}(r), \tag{24}$$

where $\boldsymbol{\Lambda}(\boldsymbol{\lambda})$ *weights the variances of the* p *individual* $\mathrm{QPV}_N(q_i)$ *estimators, and* $\boldsymbol{G}(r)$ *adapts for the change to the* r*th moment* QPV_N *estimator and the* MPV_N *instead of the pure moment estimator* M_N*, as in*

$$\boldsymbol{\Lambda}(\boldsymbol{\lambda}) = \begin{bmatrix}\boldsymbol{\lambda} & \mathbf{0}_{p\times 1}\\ 0 & 1\end{bmatrix}', \qquad\qquad \boldsymbol{G}(r) = \begin{bmatrix}r & 0\\ 0 & \frac{1}{M^r}\end{bmatrix}.$$

The $(p+1)\times(p+1)$ *matrix* $\tilde{\Xi}(\boldsymbol{q}, r)$ *describes the covariance between the sample quantiles and the* r*th order moment of standardized data, with*

$$\tilde{\Xi}(\boldsymbol{q}, r) = \begin{bmatrix}\Xi(\boldsymbol{q}) & \xi_c(\boldsymbol{q}, r)\\ \xi_c(\boldsymbol{q}, r)' & M(r)\end{bmatrix}.$$

The scalars $M^r, M(r)$ *and* $p \times p$ *matrix* $\Xi(\boldsymbol{q})$ *have been defined before in* (2), (3)*, and* (18)*, respectively. Covariances between quantile estimators and moment estimators are collected in the* p*-vector* $\xi_c(\boldsymbol{q}, r)$*, with typical element*

$$\xi_c(q_i, r) = \frac{M_{1-q_i}^r - M_{q_i}^r + (1-2q_i)M^r}{c(q_i)\phi(\Phi^{-1}(q_i))}, \tag{25}$$

where

$$M_q^r \equiv \int_{\Phi^{-1}(q)}^{\infty} |z|^r \phi(z)\,dz.$$

Proof. See Appendix A.1.

The result in Proposition 2 allows us to construct test statistics and their limiting distributions under the null of normality.

3.2 Test statistics

The joint distribution theory developed for $\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})$ and MPV_N^r allows the construction of a formal test statistic designed to detect outlying observations. Our tests build on the insight that under Assumption 1 the underlying observations are normally

distributed, therefore both estimators measure the rth power variation of the data. In the presence of a fraction of outlying observations, the $\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})$ estimator retains robustness, provided a suitable choice of $\boldsymbol{q}$ is made such that quantiles are sufficiently far away from the extreme tails, while the MPV_N^r estimator measures a combination of the variation attributable to the normal and outlying observations. As a corollary to Proposition 2, we provide three versions of our test statistic in the spirit of Barndorff-Nielsen et al. (2006). Note that we indicate our test statistics by the initials BJ, in order to distinguish them from two related families of tests proposed by Barndorff-Nielsen et al. (2006, BNS) and Jiang and Oomen (2008, JO).

Corollary 3.1. *Under the setting of Proposition 2, we obtain the following test statistics:*

i) the linear test (linear BJ(p, r)):

$$\frac{N^{1/2}\Big(\mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda}) - \mathrm{MPV}_N^r\Big)}{\sqrt{\sigma^{2r}\Omega^{\mathrm{QM}}}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{26}$$

ii) the logarithmic test (log BJ(p, r)):

$$\frac{N^{1/2}\Big(ln\ \mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda}) - ln\ \mathrm{MPV}_N^r\Big)}{\sqrt{\Omega^{\mathrm{QM}}}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{27}$$

iii) the ratio test (ratio BJ(p, r)):

$$\frac{N^{1/2}\Big(1 - \big[\,\mathrm{MPV}_N^r \,/\, \mathrm{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})\big]\Big)}{\sqrt{\Omega^{\mathrm{QM}}}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{28}$$

where $\Omega^{\mathrm{QM}} = \boldsymbol{\psi}'\,\Omega(\boldsymbol{q}, \boldsymbol{\lambda}, r)\,\boldsymbol{\psi}$, $\boldsymbol{\psi} = [1\ \ -1]'$ *and* $\Omega(\boldsymbol{q}, \boldsymbol{\lambda}, r)$ *defined in* (24)*, Proposition 2.*

Note that for application of the family of BJ tests, both the moment r and the number of pairs of quantiles p has to be pre-specified. Given p, we use the optimal set of quantiles $\boldsymbol{q}$ with corresponding weights $\boldsymbol{\lambda}$ from Table 1.

The linear BJ test depends on the unknown quantity σ^{2r}, therefore a feasible test needs to rely on an estimate of this quantity. Note that the possible gain in power by using a higher order r might be offset by a deterioration in the asymptotic variance of the test statistic, due to the estimation of σ^{2r}. To ensure that the linear test retains power when outliers are present in the data under the alternative, we estimate σ^{2r} with the robust quantile-based estimator, i.e. $\widehat{\sigma}^{2r} = \mathrm{QPV}_N^{2r}(\boldsymbol{q}, \boldsymbol{\lambda})$.

The test statistics given in Corollary 3.1 effectively test for the validity of the normality assumption, with the power of the test concentrated in the direction of testing for outlying observations, in any moment of the data. In the financial context, these statistics allow testing for jumps as shown in Sections 4 and 5.

Alternatively, in this situation the jump tests of Barndorff-Nielsen et al. (2006) and Jiang and Oomen (2008) could be used. These alternative BNS and JO tests are reviewed in Appendix A.2. The tests of BNS relate realized variance and realized bipower variation, where the former measures the quadratic variation (integrated variance plus

jump component), while the latter is robust to jumps and hence measures the integrated variance. As such, the BNS tests exploit the second moment of the data. The JO tests relate swap variance to realized variance. Jiang and Oomen (2008) show that their tests primarily use the third moment of financial returns. Both the simulation study and the empirical exercise will compare the performance of our tests relative to the two sets of alternatives.

4 Investigating the properties of the tests in a simulated setting

In this part, we study the finite sample properties of the test statistics given in Corollary 3.1. We first describe the design of our simulation study. Then we discuss the performance of the quantile-based estimator in terms of the bias in small samples. We finally discuss the behavior of the testing procedure and relate it to results for the alternative approaches.

4.1 Simulation design

Our empirical illustration in Section 5 studies financial data to test for the presence of jumps. Therefore, the simulation study is set up around the jump diffusion model (Merton, 1976). The log price process X_t is assumed to follow a Brownian semi-martingale plus jumps defined as

$$\mathrm{d}X_t = \mu_t \mathrm{d}t + \sigma_t \mathrm{d}W_t + J_t \mathrm{d}C_t, \qquad t \in [0,1], \tag{29}$$

where μ_t denotes the drift coefficient, $\sigma_t > 0$ specifies the spot volatility, W_t is a standard Brownian motion, J_t is a random-sized jump with mean $\mu_J(t)$ and variance $\sigma_J^2(t)$ and C_t is a process that counts the finite number of jumps in the price path up to time t. Price observations are assumed to be available at normalized equidistant times $0 = t_0 < t_1 < \ldots < t_N = 1$, where the interval $[0,1]$ denotes a trading day. Since our theory applies only when the variance is constant, we assume $\sigma_t = \sigma = 1$ which corresponds to $\sigma\sqrt{252} \approx 16\%$ annualized volatility and is a realistic level of financial volatility. The constant volatility model with jumps was proposed by Merton (1976) who stressed the necessity to distinguish between normal and abnormal price changes. The constant variance assumption is not necessarily restrictive and can be considered when data is sampled over a short horizon or on a suitably deformed time scale.

To simulate the price process in (29) we use an Euler discretization scheme. Without loss of generality, we assume $\mu_t = 0$. We generate one day of data at the highest frequency considered, setting $\Delta t = \frac{1}{N_{\max}}$. From each generated price path we sample price observations at different frequencies as determined by the choice of N. We consider $N \in \{50, 250, 1000, 5000\}$, to have situations of relatively low and high frequency. For each simulation, we use 100 000 replications to lower simulation variance. To keep output manageable, we first report results for $p \in \{1,2,5\}$ and $r \in \{1,2,3,4,6\}$ in tabular format, followed by a selected set of graphs concerning size and power of the BJ tests for a range of values of $r \in [0.1, 0.2, \ldots, 8]$. The nominal size is set at 5%.

To examine the bias of the quantile-based measure and to study the power of our tests, we simulate the price process in (29) by adding random-sized jumps. The number of jumps J is set to be either one or three per day. The arrival times of the jumps are uniformly distributed over the day, and their sizes are drawn from a normal distribution

with mean $\mu_J(t) = 0$ and standard deviation chosen as a multiple κ of the daily standard deviation σ, in particular $\sigma_J(t) = \sigma_J = \kappa\sigma$ with $\kappa \in \{0, 0.25, 0.50, 1.00\}$. The choice of κ determines the relative jump contribution (RJC) to the price variation, measured as

$$RJC(J,\kappa) = \mathrm{E}\left[100 \times \frac{\sum_{i=1}^{J} J_i^2}{\int_0^1 \sigma_t^2 dt + \sum_{i=1}^{J} J_i^2}\right] = 100 \int_0^\infty \frac{\kappa^2\sigma^2\xi}{\sigma^2 + \kappa^2\sigma^2\xi} \frac{2^{-J/2}}{\Gamma(J/2)} \xi^{J/2-1} \exp(-\xi/2) d\xi$$

Here this contribution varies from 5.3% (one jump with $\kappa = 0.25$) to 65.6% (three jumps with $\kappa = 1.00$). The ability to detect jumps in asset prices naturally deteriorates with lower sampling frequency. This effect will be apparent in our simulation results as well, especially for combinations of both low κ and N.

4.2 Remaining bias of the estimator

We start by studying the small sample performance of the quantile-based estimator of powers of volatility. Note that the estimate of QPV_N^r has been adapted for the scale and moment biases as described in Sections 2.6.1 and 2.6.2. Furthermore, as was discussed in Section 2.6.3, the quantile-based estimator can also experience a bias when jumps are added to the price process because the ordering of the normal order statistics (as under the null) is distorted. As a result, empirical quantiles used to construct the sub-estimator (14) might be biased. In general, this bias depends on the fraction of jumps and their variance.

Table 2 presents the simulation results for this setup. For the smallest sample size considered, $N = 50$, the estimator is indeed unbiased when $r = 1$, as the average estimator QPV_N differed no more than 0.02% from the true standard deviation when the data generating process (DGP) contained no jumps ($\kappa = 0$). For higher orders, a remaining bias of a maximum of 4.4% is found, for this small sample. Without jumps, the size of the bias quickly drops, to less than 0.5% for $N = 5000$ observations. If the bias corrections of Section 2.6 had not been used, biases of between 30% ($N = 50$) and 0.74% ($N = 5000$) would have been found.

Including one or more jumps in the DGP may lead to a more substantial bias, as the first quantile q_1 may be too extreme. In particular, in case we generate three jumps, then for all quantile measures we observe a substantial bias increasing rapidly with larger size of jumps κ, when $N = 50$. Since the number of jumps in general is not known a-priori, and the jumps are randomly-sized, the changes in ordering are stochastic and no obvious bias correction is feasible.

When we increase the sample size to $N = 250$, the resulting bias diminishes, apart from the estimator combining $p = 5$ pairs which makes use of yet too extreme quantiles to retain robustness in this sample size. In general, the gain in efficiency by combining more pairs of quantiles can only be obtained at the cost of lowering robustness as more and more information is extracted from the tails. In such a case, one could always increase the value of q_1 as discussed in Section 2.4 and observe the behavior of the measure.

Increasing the sample size further to $N = 1000$, the estimator using $p = 5$ pairs retains robustness for the one jump case, while it still makes use of the $3rd$ and $998th$ order statistics, which may be affected by a jump if three jumps are randomly placed onto the price process. Consequently, we may still observe considerable bias in this scenario. For $N = 5000$, the bias is negligible in this setup, with no estimates deviating more than 2.05% from the true value of the volatility.

Table 2: Remaining bias of $\text{QPV}_N^r(\boldsymbol{q}, \boldsymbol{\lambda})$, after using bias correction techniques

		$J=1$ jump					$J=3$ jumps				
κ	p	$r=1$	2	3	4	6	1	2	3	4	6
$N=50$											
κ=0	p=1	-0.026	-0.076	-0.238	-0.723	-4.405	-0.026	-0.076	-0.238	-0.723	-4.405
	p=2	-0.006	-0.026	-0.115	-0.454	-3.425	-0.006	-0.026	-0.115	-0.454	-3.425
	p=5	0.006	-0.004	-0.072	-0.360	-3.009	0.006	-0.004	-0.072	-0.360	-3.009
κ=0.25	p=1	1.883	3.809	5.690	7.293	7.633	*5.868*	*12.186*	*18.897*	*25.761*	*37.366*
	p=2	*3.437*	*7.560*	*12.612*	*18.891*	*37.203*	*9.860*	*22.365*	*38.774*	*61.143*	*141.108*
	p=5	*3.128*	*6.858*	*11.459*	*17.303*	*35.904*	*9.097*	*20.538*	*35.565*	*56.356*	*135.835*
κ=0.5	p=1	3.375	6.895	10.475	13.867	17.828	*11.584*	*24.935*	*40.295*	*57.739*	*97.005*
	p=2	*12.289*	*31.844*	*65.382*	*127.569*	*519.352*	*33.087*	*94.942*	*222.016*	*508.403*	*3071.751*
	p=5	*10.349*	*27.104*	*57.258*	*117.590*	*567.371*	*29.102*	*83.186*	*197.104*	*468.375*	*3275.207*
κ=1	p=1	4.433	9.094	13.902	18.601	25.244	*18.309*	*41.969*	*73.584*	*117.440*	*275.875*
	p=2	*34.614*	*118.181*	*347.651*	*1054.375*	*12272.202*	*87.959*	*376.655*	*1462.972*	*6035.390*	*132629.020*
	p=5	*27.798*	*99.095*	*315.480*	*1067.006*	*16071.945*	*75.811*	*324.552*	*1310.521*	*5829.070*	*157956.115*
$N=250$											
κ=0	p=1	-0.001	-0.006	-0.019	-0.048	-0.224	-0.001	-0.006	-0.019	-0.048	-0.224
	p=2	0.004	0.003	-0.004	-0.024	-0.160	0.004	0.003	-0.004	-0.024	-0.160
	p=5	0.001	0.001	0.001	-0.002	-0.046	0.001	0.001	0.001	-0.002	-0.046
κ=0.25	p=1	0.608	1.217	1.824	2.420	3.505	1.843	3.721	5.633	7.568	11.433
	p=2	0.801	1.612	2.432	3.253	4.845	2.485	5.065	7.746	10.528	16.366
	p=5	*2.878*	*7.302*	*14.786*	*28.639*	*115.746*	*7.860*	*21.260*	*47.111*	*102.803*	*569.914*
κ=0.5	p=1	0.752	1.507	2.263	3.007	4.396	2.298	4.652	7.058	9.507	14.460
	p=2	1.030	2.075	3.136	4.206	6.316	3.274	6.712	10.324	14.119	22.266
	p=5	*8.130*	*27.603*	*81.890*	*252.485*	*3048.078*	*20.488*	*83.188*	*312.498*	*1262.151*	*26929.130*
κ=1	p=1	0.831	1.666	2.501	3.328	4.882	2.545	5.156	7.831	10.562	16.115
	p=2	1.156	2.331	3.526	4.733	7.132	3.735	7.680	11.852	16.264	25.842
	p=5	*19.441*	*100.033*	*504.737*	*2838.857*	*123243.323*	*46.305*	*324.273*	*2330.675*	*18684.406*	*1599218.153*
$N=1000$											
κ=0	p=1	-0.013	-0.025	-0.038	-0.051	-0.084	-0.013	-0.025	-0.038	-0.051	-0.084
	p=2	-0.008	-0.023	-0.045	-0.075	-0.165	-0.008	-0.023	-0.045	-0.075	-0.165
	p=5	-0.008	-0.024	-0.051	-0.087	-0.192	-0.008	-0.024	-0.051	-0.087	-0.192
κ=0.25	p=1	0.171	0.342	0.514	0.685	1.022	0.540	1.083	1.629	2.177	3.277
	p=2	0.229	0.453	0.669	0.879	1.272	0.713	1.426	2.138	2.850	4.268
	p=5	0.381	0.760	1.139	1.518	2.278	*1.479*	*3.154*	*5.133*	*7.615*	*15.904*
κ=0.5	p=1	0.190	0.380	0.570	0.761	1.136	0.596	1.197	1.800	2.407	3.626
	p=2	0.258	0.510	0.755	0.994	1.446	0.800	1.601	2.404	3.208	4.816
	p=5	0.441	0.882	1.325	1.771	2.672	*2.383*	*5.932*	*12.247*	*25.738*	*153.475*
κ=1	p=1	0.199	0.399	0.600	0.800	1.195	0.626	1.255	1.889	2.526	3.807
	p=2	0.272	0.538	0.798	1.051	1.533	0.845	1.692	2.542	3.394	5.100
	p=5	0.472	0.945	1.421	1.901	2.875	*4.123*	*14.275*	*48.173*	*187.776*	*4307.690*
$N=5000$											
κ=0	p=1	0.030	0.060	0.090	0.120	0.179	0.030	0.060	0.090	0.120	0.179
	p=2	0.042	0.084	0.126	0.168	0.252	0.042	0.084	0.126	0.168	0.252
	p=5	0.070	0.140	0.209	0.279	0.419	0.070	0.140	0.209	0.279	0.419
κ=0.25	p=1	0.070	0.141	0.211	0.281	0.422	0.152	0.304	0.456	0.608	0.913
	p=2	0.095	0.189	0.284	0.379	0.568	0.201	0.402	0.603	0.805	1.210
	p=5	0.149	0.299	0.449	0.600	0.904	0.312	0.626	0.944	1.264	1.914
κ=0.5	p=1	0.072	0.144	0.216	0.288	0.432	0.157	0.314	0.471	0.628	0.943
	p=2	0.097	0.194	0.292	0.389	0.583	0.208	0.417	0.626	0.835	1.255
	p=5	0.154	0.308	0.463	0.619	0.932	0.326	0.655	0.987	1.323	2.004
κ=1	p=1	0.073	0.146	0.219	0.292	0.437	0.159	0.319	0.479	0.638	0.959
	p=2	0.098	0.197	0.295	0.394	0.591	0.212	0.424	0.637	0.850	1.278
	p=5	0.156	0.313	0.470	0.628	0.947	0.333	0.670	1.009	1.352	2.049

Note: The table reports the percentage bias of the estimator of σ^r averaged over number of repetitions. In italics, the biases for cases where the quantile q_1, and hence the QPV estimator, may be contaminated by a jump, i.e. when $\kappa > 0$ and $q_1 < \frac{1+J}{N+1}$.

4.3 Size of the test: Distribution under the null

To study the size of our test statistics, we simulate the price process as discussed in Section 4.1 under the null of no jumps, i.e. with $J = 0$ (or $\kappa = 0$, which of course results in the same data generating process). Table 3 reports the sample statistics: standard deviation, skewness, kurtosis and empirical size of the proposed test statistics. We report analogous results for the BNS tests, and skip results for the JO tests as those are qualitatively similar.

The results in Table 3 show that in a small sample of $N = 50$ the densities of the BJ test statistics with higher order r generally are asymmetric and have heavier tails than the standard normal density as judged by values of skewness and kurtosis respectively. However, the size properties of the tests using $p = 1$ set of quantiles are still decent, which is remarkable given the fact that the sample size is extremely limited. A detailed analysis of the simulation results shows that the more extreme skewness and kurtosis of the linear and ratio tests for $r = 6$ is caused by a single extreme BJ statistic, for a generated data series which indeed happens to contain an increment of size 4.4σ. The BJ test based on the logarithm is less influenced by this single extreme event, as taking logarithms results in a BJ test which is far less extreme. Indeed, even though skewness and kurtosis of the linear and ratio tests seem to indicate a larger difference with the Gaussian density, their size properties at $N = 50$ are better than the counterpart of the BJ test based on the logarithmic transformation. Below we will return to the relative size of the tests, in the discussion of Figure 4, as long as only a single set of quantiles is used ($p = 1$). Increasing p to either 2 or 5 sets of quantiles leads to the most extreme quantile being too small to be estimated well from such a small sample, and the size of the tests deteriorates.

As we increase the sample size and look at the simulation results based on $N = 250$, we observe a noticeable improvement in performance. The values of variance, skewness and kurtosis are more consistent with the standard normal distribution. The size distortions are far less severe, as the empirical sizes match the nominal size of 5% closely. At this sampling frequency the linear and ratio BJ tests still performs worse than the log BJ test in terms of skewness and kurtosis, though the power of the latter test is considerably worse, especially for higher order r. When we increase the sample size to $N = 1000$ and $N = 5000$, then the empirical sizes are consistent with the nominal size of 5%.[1]

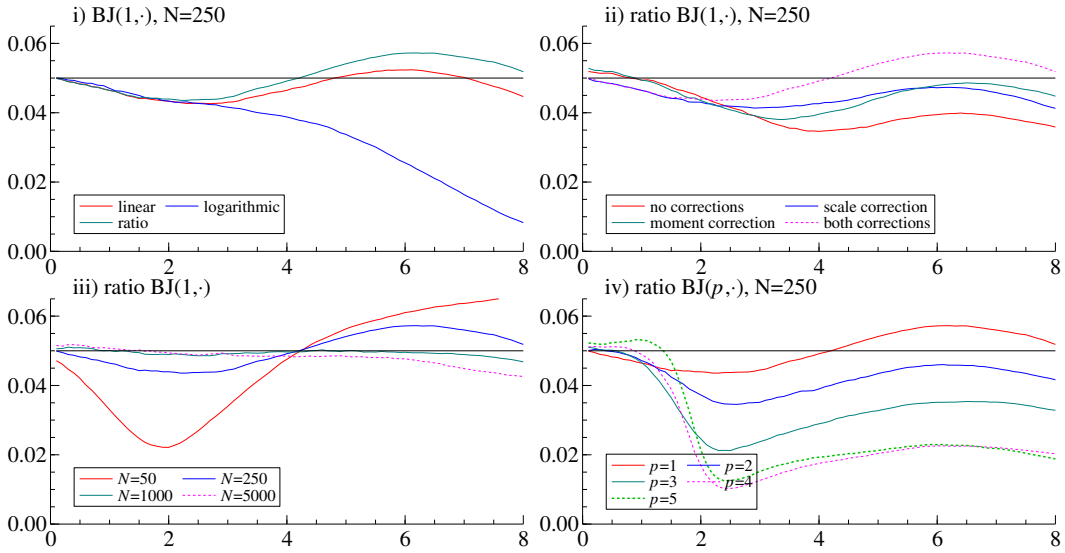
Although the analysis of the size of the alternative tests on the basis of Table 3 is detailed, for overview Figure 4 is presented. Conceptually it could be easier to think of the BJ tests to test for certain integer orders r of the volatility σ, but the theory was developed to also test non-integer orders r. Hence, in the first panel of Figure 4, a comparison is made between the sizes of the three versions of the BJ test, for a range of values of $r \in \{0.1, 0.2, \ldots, 8\}$, with $N = 250$ observations, applying the tests using only

[1]The theory developed in this article is based on the constant volatility assumption, as the test effectively relates two measures of unconditional variation under normality. If this assumption is not met, i.e. if there is remaining deterministic or stochastic time-varying volatility, then the density of returns will have heavier tails than the normal density. Our test on such data will detect those tails, and hence it becomes oversized. A simulation study (results not reported here) indeed indicates that the size distortion is minor for approximately constant volatility processes. The size however increases up to 25% when volatility process has common GARCH persistence, or if there is a periodic variance pattern in returns such as for intraday data. However, as explained in the empirical section of this article, we propose to devolatilize the return data by suitably sampling the observations, before applying the jumps test.

Table 3: Distribution under the null of BJ and BNS statistics

		BJ(·, 1)			BJ(·, 2)			BJ(·, 3)			BJ(·, 4)			BJ(·, 6)			BNS			
		lin	log	ratio	lin	log	ratio	lin	log	ratio	lin	log	ratio	lin	log	ratio	lin	log	ratio	adj.ratio
$N = 50$																				
mean	BJ(1,·)	0.068	0.104	0.068	0.033	0.084	0.034	−0.023	0.060	−0.024	−0.079	0.056	−0.087	−0.175	0.075	−0.224	−0.158	−0.092	−0.029	−0.002
	BJ(2,·)	0.137	0.166	0.137	0.164	0.192	0.166	0.118	0.150	0.122	0.076	0.132	0.083	0.022	0.132	0.035				
	BJ(5,·)	0.210	0.236	0.210	0.404	0.423	0.407	0.232	0.261	0.240	0.133	0.197	0.146	0.043	0.172	0.063				
std. dev.	BJ(1,·)	0.921	0.929	0.921	0.829	0.829	0.841	0.839	0.789	0.876	0.916	0.768	1.000	1.153	0.667	1.472	1.205	1.136	1.094	0.986
	BJ(2,·)	0.970	0.980	0.970	0.715	0.732	0.722	0.531	0.547	0.545	0.508	0.530	0.527	0.446	0.472	0.477				
	BJ(5,·)	1.040	1.051	1.040	0.701	0.719	0.705	0.525	0.543	0.533	0.580	0.598	0.589	0.525	0.519	0.533				
skew.	BJ(1,·)	−0.128	0.103	−0.128	−0.657	−0.240	−0.657	−1.761	−0.781	−1.761	−3.195	−1.005	−3.195	−11.324	−0.997	−11.324	−1.221	−0.788	−0.427	−0.272
	BJ(2,·)	−0.065	0.115	−0.065	−0.212	−0.017	−0.222	−0.675	−0.331	−0.672	−0.826	−0.125	−0.762	−2.742	0.121	−1.758				
	BJ(5,·)	−0.074	0.070	−0.074	−0.028	0.052	−0.053	−0.520	−0.235	−0.519	−1.041	−0.147	−0.887	−3.703	0.012	−2.091				
kurt.	BJ(1,·)	3.008	3.008	3.008	4.004	3.235	4.004	9.422	4.143	9.422	26.453	4.541	26.453	402.249	4.286	402.249	5.801	4.072	3.229	2.978
	BJ(2,·)	3.015	3.037	3.015	3.114	3.010	3.105	4.125	3.324	3.984	5.432	3.206	4.521	29.140	3.323	10.951				
	BJ(5,·)	3.031	3.039	3.031	3.164	3.107	3.103	4.857	3.835	4.463	8.221	3.754	5.927	48.255	3.639	13.244				
size	BJ(1,·)	0.033	0.036	0.033	0.021	0.019	0.022	0.031	0.018	0.034	0.040	0.018	0.048	0.044	0.009	0.061	0.082	0.074	0.069	0.044
	BJ(2,·)	0.045	0.049	0.045	0.007	0.009	0.007	0.002	0.001	0.002	0.003	0.000	0.003	0.005	0.000	0.005				
	BJ(5,·)	0.064	0.069	0.064	0.014	0.018	0.014	0.002	0.002	0.002	0.006	0.004	0.005	0.009	0.001	0.008				
$N = 250$																				
mean	BJ(1,·)	0.029	0.047	0.029	0.003	0.033	0.003	−0.032	0.022	−0.032	−0.062	0.027	−0.063	−0.106	0.077	−0.111	−0.065	−0.039	−0.014	−0.005
	BJ(2,·)	0.060	0.073	0.060	0.061	0.079	0.061	0.028	0.063	0.028	0.000	0.065	0.000	−0.037	0.113	−0.038				
	BJ(5,·)	0.097	0.107	0.097	0.181	0.189	0.181	0.102	0.121	0.103	0.060	0.104	0.064	0.029	0.143	0.041				
std. dev.	BJ(1,·)	0.982	0.984	0.982	0.970	0.968	0.973	0.979	0.964	0.988	0.999	0.956	1.017	1.054	0.884	1.105	1.043	1.031	1.023	0.989
	BJ(2,·)	0.986	0.988	0.986	0.938	0.941	0.940	0.932	0.925	0.938	0.950	0.919	0.962	0.993	0.841	1.028				
	BJ(5,·)	1.008	1.010	1.008	0.837	0.841	0.838	0.800	0.802	0.801	0.827	0.822	0.826	0.819	0.748	0.798				
skew.	BJ(1,·)	−0.062	0.049	−0.062	−0.308	−0.111	−0.308	−0.674	−0.310	−0.674	−1.094	−0.426	−1.094	−2.778	−0.552	−2.778	−0.520	−0.361	−0.209	−0.155
	BJ(2,·)	−0.018	0.063	−0.018	−0.216	−0.095	−0.216	−0.662	−0.402	−0.663	−1.117	−0.546	−1.120	−2.990	−0.724	−3.004				
	BJ(5,·)	−0.015	0.047	−0.015	−0.056	−0.006	−0.058	−0.312	−0.169	−0.306	−0.587	−0.210	−0.541	−2.088	−0.379	−1.600				
kurt.	BJ(1,·)	3.015	3.016	3.015	3.215	3.058	3.215	3.899	3.195	3.899	5.397	3.328	5.397	21.079	3.595	21.079	3.519	3.243	3.068	3.005
	BJ(2,·)	3.000	3.011	3.000	3.153	3.076	3.152	3.987	3.399	3.980	5.834	3.731	5.826	24.417	4.207	24.556				
	BJ(5,·)	3.020	3.025	3.020	3.016	3.004	3.013	3.334	3.135	3.275	4.046	3.222	3.756	14.626	3.543	9.069				
size	BJ(1,·)	0.046	0.047	0.046	0.043	0.043	0.044	0.043	0.042	0.044	0.046	0.039	0.049	0.052	0.026	0.057	0.057	0.055	0.054	0.046
	BJ(2,·)	0.047	0.048	0.047	0.037	0.038	0.037	0.034	0.034	0.035	0.038	0.033	0.039	0.043	0.022	0.046				
	BJ(5,·)	0.053	0.054	0.053	0.021	0.023	0.021	0.015	0.016	0.015	0.021	0.019	0.019	0.025	0.011	0.023				
$N = 1000$																				
mean	BJ(1,·)	0.011	0.021	0.011	−0.003	0.013	−0.003	−0.020	0.008	−0.020	−0.035	0.012	−0.035	−0.056	0.047	−0.057	−0.028	−0.016	−0.003	0.001
	BJ(2,·)	0.029	0.036	0.029	0.024	0.034	0.024	0.005	0.025	0.005	−0.009	0.027	−0.009	−0.028	0.062	−0.028				
	BJ(5,·)	0.046	0.051	0.046	0.076	0.080	0.076	0.035	0.048	0.036	0.015	0.042	0.015	−0.002	0.076	−0.001				
std. dev.	BJ(1,·)	1.000	1.000	1.000	0.997	0.997	0.998	0.998	0.994	1.000	1.001	0.991	1.006	1.013	0.958	1.025	1.009	1.006	1.004	0.990
	BJ(2,·)	1.002	1.003	1.002	0.992	0.992	0.993	0.987	0.985	0.989	0.989	0.980	0.993	1.001	0.943	1.011				
	BJ(5,·)	0.998	0.998	0.998	0.942	0.943	0.943	0.945	0.944	0.946	0.957	0.951	0.958	0.969	0.916	0.973				
skew.	BJ(1,·)	−0.034	0.022	−0.034	−0.145	−0.048	−0.145	−0.305	−0.133	−0.305	−0.489	−0.196	−0.489	−1.227	−0.366	−1.227	−0.250	−0.176	−0.102	−0.078
	BJ(2,·)	−0.012	0.029	−0.012	−0.117	−0.054	−0.117	−0.322	−0.197	−0.322	−0.526	−0.282	−0.526	−1.424	−0.511	−1.425				
	BJ(5,·)	−0.006	0.024	−0.006	−0.099	−0.069	−0.099	−0.275	−0.195	−0.275	−0.453	−0.258	−0.451	−1.491	−0.562	−1.484				
kurt.	BJ(1,·)	2.986	2.985	2.986	3.026	2.990	3.026	3.154	3.011	3.154	3.435	3.049	3.435	6.762	3.384	6.762	3.106	3.044	3.002	2.982
	BJ(2,·)	3.004	3.004	3.004	3.072	3.051	3.072	3.251	3.117	3.249	3.645	3.220	3.641	8.989	3.782	8.988				
	BJ(5,·)	3.002	3.004	3.002	3.070	3.059	3.070	3.221	3.132	3.212	3.636	3.287	3.611	10.182	4.155	10.119				
size	BJ(1,·)	0.050	0.050	0.050	0.049	0.049	0.049	0.049	0.048	0.049	0.049	0.047	0.050	0.048	0.040	0.050	0.052	0.051	0.051	0.048
	BJ(2,·)	0.050	0.050	0.050	0.048	0.048	0.048	0.046	0.046	0.046	0.045	0.045	0.046	0.045	0.037	0.046				
	BJ(5,·)	0.050	0.050	0.050	0.039	0.039	0.039	0.038	0.038	0.038	0.040	0.039	0.040	0.038	0.032	0.039				
$N = 5000$																				
mean	BJ(1,·)	0.043	0.047	0.043	0.044	0.051	0.044	0.030	0.042	0.030	0.015	0.036	0.015	−0.005	0.043	−0.005	−0.011	−0.006	−0.000	0.001
	BJ(2,·)	0.092	0.095	0.092	0.119	0.124	0.119	0.088	0.097	0.088	0.057	0.074	0.057	0.023	0.066	0.023				
	BJ(5,·)	0.187	0.190	0.187	0.368	0.371	0.368	0.212	0.218	0.212	0.125	0.139	0.125	0.059	0.098	0.060				
std. dev.	BJ(1,·)	1.001	1.002	1.001	0.996	0.996	0.996	0.994	0.995	0.995	0.995	0.994	0.996	0.997	0.986	1.000	1.003	1.002	1.002	0.996
	BJ(2,·)	1.002	1.003	1.002	0.998	0.999	0.998	0.996	0.997	0.997	0.996	0.996	0.997	0.996	0.985	0.998				
	BJ(5,·)	1.005	1.006	1.005	0.991	0.993	0.991	0.986	0.988	0.987	0.987	0.989	0.988	0.988	0.978	0.988				
skew.	BJ(1,·)	−0.009	0.016	−0.009	−0.064	−0.020	−0.064	−0.149	−0.072	−0.149	−0.236	−0.107	−0.236	−0.540	−0.219	−0.540	−0.125	−0.092	−0.059	−0.049
	BJ(2,·)	0.013	0.031	0.013	−0.053	−0.025	−0.053	−0.148	−0.095	−0.148	−0.236	−0.135	−0.236	−0.583	−0.287	−0.583				
	BJ(5,·)	0.014	0.027	0.014	−0.050	−0.036	−0.050	−0.143	−0.107	−0.143	−0.225	−0.145	−0.225	−0.625	−0.343	−0.624				
kurt.	BJ(1,·)	3.028	3.029	3.028	3.036	3.029	3.036	3.055	3.023	3.055	3.109	3.025	3.109	3.703	3.164	3.703	3.023	3.010	3.000	2.994
	BJ(2,·)	2.959	2.960	2.959	2.988	2.984	2.988	3.027	3.002	3.026	3.092	3.019	3.091	3.861	3.263	3.860				
	BJ(5,·)	2.991	2.992	2.991	3.022	3.019	3.022	3.045	3.025	3.044	3.111	3.046	3.107	4.083	3.409	4.076				
size	BJ(1,·)	0.051	0.051	0.051	0.049	0.050	0.049	0.049	0.049	0.049	0.048	0.048	0.049	0.047	0.047	0.048	0.050	0.050	0.050	0.049
	BJ(2,·)	0.051	0.052	0.051	0.050	0.051	0.050	0.049	0.049	0.049	0.048	0.049	0.048	0.047	0.046	0.047				
	BJ(5,·)	0.055	0.056	0.055	0.063	0.064	0.063	0.050	0.051	0.050	0.047	0.049	0.047	0.044	0.044	0.044				

Figure 4: Size characteristics of BJ tests



($p=$)1 set of quantiles. For values of $r < 3$, the three tests have similar size, close to the nominal level of α=5%. For $r > 3$, the logarithmic BJ test becomes size distorted, while both the linear and ratio tests retain a size close to the nominal level.

Of these latter two tests, the ratio test had the advantage that no estimate of σ^{2r} is needed (see (28)), and hence can be expected to be more robust in real-life situations. Therefore this ratio BJ test is the test of choice in the remainder of the article. Panel ii) of Figure 4 investigates the effect of correcting either the scale coefficient (see Section 2.6.1), or the effect of taking higher powers (Section 2.6.2), or both, for the small sample size. Using no corrections at all leads to a (ratio) BJ test which is more and more undersized for higher powers of r. Using both corrections indeed improves the size characteristics of this test.

In Panel iii) the reaction of the size of the test to the sample size is depicted. For $N = 50$, the ratio BJ test can become rather severely undersized, especially around $r = 2$, or oversized for values of $r > 5$. For such a small sample size, estimating the $q = 0.0691$ quantile may be too extreme. For sample sizes $N \geq 250$, the size properties quickly improve.

The last panel displays a similar message as panel iii), as it compares using multiple sets of quantiles for estimating the QPV(p, r) statistics underlying the BJ test. Even though the variance of the estimators might improve, the size of the resulting test statistic becomes worse as more and more extreme quantiles of the data are taken into account.

4.4 Power of the test: Detection of jumps

To examine the power of our tests, we simulate the price process by adding random jumps as discussed in Section 4.1. The construction of the test would suggest that the power of the test will increase with the order r, as the moment-based measure becomes more heavily influenced by outliers. To study this effect, Figure 5 presents the power of the tests as a function of the order r, for a range of settings.

Panel i) of Figure 5 compares again the power of the three different BJ test specifications, when $J = 1$ jump of random size with standard deviation $\kappa = 1$ times the daily variation of the process is added to the DGP, using $N = 250$ observations. Only for $r > 3$, the logarithmic test seems to lose some power relative to the linear and ratio BJ specifications. The power of the BJ tests at $r = 2$ is fractionally higher than the BNS ratio test, as indicated by the symbol. Likewise, at $r = 3$ the BJ tests display slightly higher power than the JO test.[2]

Continuing as before with the ratio test, panel ii) displays the power as a function of the jump size. For jumps with a standard deviation of only $\kappa = 0.25$ of the daily variation, the power levels off to around 40%; a higher jumpsize of $\kappa = 1$ is detected in up to 80% of the simulated data series.[3] In all these situations, the ratio BJ test displays slightly higher power than the BNS and JO tests, as before. When $J = 3$ jumps are introduced to this DGP, the difference in power between BJ and BNS/JO tests increases, as will be shown in Table 4 below.

When comparing the effect of sample size on power, in Panel iii), a single jump is detected more easily with larger sample sizes, as expected. In the case of $N = 50$, the JO test attains slightly more power than the BJ(1,3) test, as does the BNS test vs the BJ(1,2) for $N = 5000$.

In Panel iv) the effect of the number of pairs p of quantiles on the power of the test is depicted, when $N = 250$. For this number of observations, and a single jump, only the BJ-tests with $p \leq 3$ are consistent. For higher number of quantiles the most extreme quantile used in the QPV estimator may already be contaminated by the jump (see also the discussion in Section 4.2). Indeed, with $p \geq 4$, the power signature of the BJ test is no longer monotonically increasing from the nominal size of $\alpha = 5\%$ at $r \approx 0$ to higher levels of power, as the power now depends on the relative effect of the jump on both the (contaminated) QPV and MPV estimators.

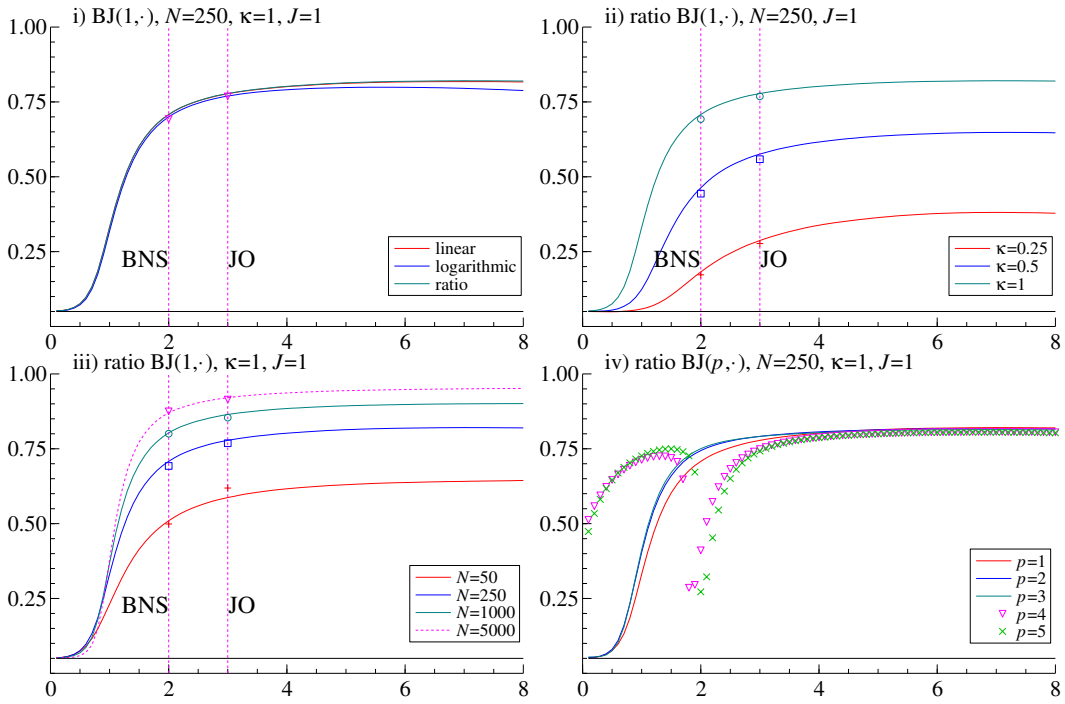
Alternatively, Table 4 presents the rejection rates of our ratio test in numerical format, with analogous results reported for the (adjusted) BNS and JO ratio tests. Note that again cases where the BJ test itself may be contaminated by a jump are printed in italics, i.e. in cases where the most extreme quantile $q_1 < (1 + J)/(N + 1)$. The simulation results can be summarized as follows. The use of the BJ tests built on higher order of r does improve the power to detect jumps considerably, also for relatively large sample sizes. This is fully in line with expectations. Indeed, apart for the exception of the JO test at $N = 50$, our BJ$(\cdot, 6)$ tests are uniformly more powerful than any other test considered here, regardless of the variance of jumps and of the number of jumps.

We find, as expected, that the BJ$(\cdot, 2)$ test performs similarly to the BNS test as both use the second moment of the data. However, unlike the BJ$(\cdot, 2)$ test, the BNS test can (substantially) lose power when jumps occur at adjacent locations, as a key requirement for the robustness of both bipower and quadpower measures is violated in such a case, see (36) and (37) in the appendix respectively. The BJ$(\cdot, 3)$ test is found to perform similarly in the single jump case to the JO test. As the JO test primarily uses the third order of the data (see Jiang and Oomen, 2008), this result is expected. However, unlike the BJ$(\cdot, 3)$ test, the JO test can (substantially) lose power when multiple jumps occur

[2]For the JO ratio test (see Appendix A.2, Equation (40)), the variance is estimated using multi-power variation of order $v = 6$ in (42).

[3]As the jumps are random, there is considerable probability that a jump of relatively small size is drawn which is not distinguishable from a normal increment. Those situations are not detectable with any test, hence a power of 100% should not be expected.

Figure 5: Distribution under the alternative



with different signs as the cubed jumps then offset each other.

If we decrease the sample size to $N = 1000$, then the advantage of the BJ$(\cdot, 6)$ test over the BJ$(\cdot, 4)$ version decreases slightly, but still both versions are more powerful than the BNS or JO test, and the difference in power can be substantial. For instance, with three jumps with variance determined by $\kappa = 0.25$, the BJ$(2, 4)$ has a power of 92.48% and the BJ$(2, 6)$ has a power of 93.84%, rejection rates considerably higher than the 78.13% of the JO test and 72.48% of the adjusted BNS. If we increase the variance of jumps to e.g. $\kappa = 1.00$, then the power of all tests is close to unity.

For the sample size $N = 250$, a sample size for which the size distortions are modest, the BJ tests still outperform the JO and BNS tests, at least as $r \geq 3$. Notice that in this case the BJ tests the restriction that $q_1 < (J+1)/(N+1)$ may become binding for e.g. $p = 3$: If $p = 3$, then $q_1 = 0.0104$ and the BJ test loses robustness if more than 1 jump is found in one tail of the distribution of returns. When $N = 50$, only a single set of quantiles should be used for detecting 1 jump. If the three jumps of the right-hand panel of Table 4 would occur in the same tail, the BJ test loses its ability to detect these deviations from the null. Hence the reported powers in the table should be interpreted with care in this case.

5 Testing for jumps in equity price data

5.1 Describing the data

The data we analyze includes fifteen components of the Dow Jones Industrial Average index. The data are consolidated trades extracted from the Trade and Quote (TAQ)

Table 4: Power of ratio tests

		$J=1$ jump								$J=3$ jumps							
		BJ(.,1)	BJ(.,2)	BJ(.,3)	BJ(.,4)	BJ(.,6)	BNS	adj.BNS	JO	BJ(.,1)	BJ(.,2)	BJ(.,3)	BJ(.,4)	BJ(.,6)	BNS	adj.BNS	JO
$N=50$																	
κ=0.25	BJ(1,.)	0.0369	0.0491	0.0876	0.1133	0.1352	0.0937	0.0643	0.1611	*0.0438*	*0.0970*	*0.1708*	*0.2131*	*0.2473*	0.1372	0.1006	0.2402
	BJ(2,.)	*0.0771*	*0.0124*	*0.0048*	*0.0192*	*0.0359*				*0.1315*	*0.0269*	*0.0100*	*0.0447*	*0.0770*			
	BJ(5,.)	*0.1003*	*0.0269*	*0.0171*	*0.0380*	*0.0471*				*0.1660*	*0.0556*	*0.0416*	*0.0854*	*0.0991*			
κ=0.5	BJ(1,.)	0.0685	0.2167	0.3022	0.3393	0.3681	0.2299	0.1918	0.3569	*0.1289*	*0.4598*	*0.6002*	*0.6497*	*0.6797*	0.4588	0.4078	0.5846
	BJ(2,.)	*0.2461*	*0.0418*	*0.0690*	*0.1723*	*0.2254*				*0.4941*	*0.1727*	*0.1433*	*0.3105*	*0.3761*			
	BJ(5,.)	*0.2666*	*0.0611*	*0.1658*	*0.2330*	*0.2493*				*0.5500*	*0.2269*	*0.3070*	*0.4164*	*0.4163*			
κ=1	BJ(1,.)	0.2314	0.5097	0.5874	0.6165	0.6373	0.4989	0.4664	0.6189	*0.4820*	*0.8480*	*0.9085*	*0.9244*	*0.9328*	0.8322	0.8043	0.8560
	BJ(2,.)	*0.5291*	*0.0691*	*0.3358*	*0.4767*	*0.5271*				*0.8528*	*0.4045*	*0.4264*	*0.5989*	*0.6367*			
	BJ(5,.)	*0.5413*	*0.0831*	*0.4691*	*0.5349*	*0.5483*				*0.8879*	*0.4780*	*0.5988*	*0.6880*	*0.6480*			
$N=250$																	
κ=0.25	BJ(1,.)	0.0587	0.1823	0.2874	0.3381	0.3773	0.1727	0.1609	0.2771	0.0992	0.4489	0.6245	0.6880	0.7294	0.4059	0.3918	0.5161
	BJ(2,.)	0.0627	0.2251	0.3105	0.3444	0.3708				0.1047	0.5220	0.6607	0.7011	0.7264			
	BJ(5,.)	*0.2121*	*0.1074*	*0.2110*	*0.2938*	*0.3342*				*0.4634*	*0.2923*	*0.4782*	*0.6431*	*0.6912*			
κ=0.5	BJ(1,.)	0.1239	0.4633	0.5758	0.6163	0.6448	0.4438	0.4335	0.5588	0.3816	0.8574	0.9235	0.9403	0.9501	0.8313	0.8259	0.8467
	BJ(2,.)	0.1572	0.5192	0.5969	0.6227	0.6420				0.4311	0.8909	0.9335	0.9437	0.9497			
	BJ(5,.)	*0.4943*	*0.2307*	*0.5141*	*0.5879*	*0.6181*				*0.8445*	*0.6327*	*0.8305*	*0.9295*	*0.9418*			
κ=1	BJ(1,.)	0.3313	0.7080	0.7782	0.8022	0.8189	0.6926	0.6861	0.7688	0.7958	0.9777	0.9891	0.9917	0.9932	0.9711	0.9701	0.9597
	BJ(2,.)	0.4043	0.7439	0.7915	0.8066	0.8166				0.8408	0.9833	0.9906	0.9922	0.9930			
	BJ(5,.)	*0.7261*	*0.2725*	*0.7428*	*0.7870*	*0.8047*				*0.9710*	*0.8379*	*0.9394*	*0.9898*	*0.9919*			
$N=1000$																	
κ=0.25	BJ(1,.)	0.0707	0.3386	0.4998	0.5619	0.6080	0.3326	0.3277	0.4679	0.1635	0.7436	0.8800	0.9157	0.9364	0.7285	0.7248	0.7813
	BJ(2,.)	0.0818	0.4176	0.5373	0.5784	0.6109				0.2147	0.8219	0.9034	0.9248	0.9384			
	BJ(5,.)	0.0886	0.4945	0.5610	0.5868	0.6097				*0.2029*	*0.8628*	*0.9159*	*0.9285*	*0.9382*			
κ=0.5	BJ(1,.)	0.1512	0.6183	0.7317	0.7711	0.7980	0.6132	0.6101	0.7117	0.5065	0.9548	0.9811	0.9876	0.9908	0.9501	0.9493	0.9484
	BJ(2,.)	0.2075	0.6779	0.7553	0.7804	0.7993				0.6283	0.9708	0.9857	0.9891	0.9912			
	BJ(5,.)	0.2539	0.7293	0.7715	0.7853	0.7989				*0.6131*	*0.9774*	*0.9877*	*0.9899*	*0.9913*			
κ=1	BJ(1,.)	0.3680	0.8033	0.8647	0.8846	0.8978	0.8004	0.7989	0.8542	0.8620	0.9939	0.9976	0.9983	0.9987	0.9932	0.9931	0.9879
	BJ(2,.)	0.4651	0.8364	0.8772	0.8890	0.8987				0.9171	0.9960	0.9980	0.9984	0.9987			
	BJ(5,.)	0.5290	0.8640	0.8841	0.8918	0.8986				*0.8887*	*0.9968*	*0.9982*	*0.9985*	*0.9987*			
$N=5000$																	
κ=0.25	BJ(1,.)	0.0736	0.5155	0.6907	0.7501	0.7931	0.5323	0.5307	0.6684	0.2054	0.9123	0.9727	0.9845	0.9906	0.9166	0.9161	0.9311
	BJ(2,.)	0.0886	0.5882	0.7197	0.7628	0.7968				0.2952	0.9455	0.9797	0.9866	0.9908			
	BJ(5,.)	0.1066	0.6700	0.7440	0.7720	0.7981				0.3619	0.9694	0.9841	0.9881	0.9911			
κ=0.5	BJ(1,.)	0.1587	0.7419	0.8435	0.8735	0.8951	0.7533	0.7525	0.8314	0.5692	0.9871	0.9965	0.9979	0.9987	0.9878	0.9877	0.9864
	BJ(2,.)	0.2244	0.7871	0.8579	0.8801	0.8970				0.7069	0.9923	0.9972	0.9983	0.9988			
	BJ(5,.)	0.2898	0.8320	0.8704	0.8843	0.8978				0.7880	0.9958	0.9979	0.9984	0.9988			
κ=1	BJ(1,.)	0.3791	0.8698	0.9208	0.9365	0.9480	0.8750	0.8745	0.9144	0.8881	0.9982	0.9996	0.9998	0.9999	0.9984	0.9984	0.9972
	BJ(2,.)	0.4839	0.8932	0.9284	0.9407	0.9488				0.9400	0.9991	0.9997	0.9998	0.9999			
	BJ(5,.)	0.5615	0.9153	0.9354	0.9426	0.9493				0.9623	0.9995	0.9998	0.9998	0.9999			

Note: The table power of the ratio BJ test as the fraction of rejections over $M = 100.000$ simulated series of length N of the DGP, including either $J = 1$ or $J = 3$ jumps of a random size with standard deviation of $\kappa \times \sigma$, where σ is the daily standard deviation of the process. In italics, the power for cases where the quantile q_1 (and hence the BJ test) may be contaminated by a jump, i.e. when $q_1 < \frac{1+J}{N+1}$.

database through the Wharton Research Data Services (WRDS) system. The sample period spans almost five years, from January 3, 2006 to December 31, 2010, with a total of $T = 1259$ trading days for all equities.

Before we construct measures and compute tests, the data needs to be cleaned. The importance of tick-by-tick data cleaning is highlighted by Brownlees and Gallo (2006), Hansen and Lunde (2006) and Barndorff-Nielsen, Hansen, Lunde, and Shephard (2009) who provide a guideline on cleaning procedures based on the TAQ qualifiers that are included in the files and described in the TAQ User's Guide available at the WRDS website. Following Barndorff-Nielsen et al. (2009) we consider five steps: **P1**. delete entries with a time stamp outside the 9:30am to 4:00pm Eastern Time window; **P2**. delete entries with transaction prices equal to zero; **P3**. retain entries only from the single (primary) exchange which is the NYSE (including NYSE Direct+) in our application; **T1**. delete incorrect trades as indicated by the correction indicator; **T2**. delete entries when the sale condition is not regular. Table 5 provides statistics regarding the pre-processing of the data, reporting the number of trades per year before/after cleaning, and the average number of trades per day including minimum and maximum, over the years in the sample. The cleaning procedures lead to a substantial data reduction. For all equities the largest deletion of the raw data is due to step **P3**, which is implemented also to reduce the impact of time-delays in the reporting of trades updates, see Barndorff-Nielsen et al. (2009) for further discussion. The largest number of transactions for all equities was recorded over 2007 and 2008. The average number of clean transactions for all equities per day amounts to around 12 000.

Intraday returns can be constructed using different types of sampling schemes, see Zhou (1992), Hansen and Lunde (2006), Oomen (2006) and Griffin and Oomen (2008). Most common is the *calendar time sampling* scheme, where data is sampled at fixed time intervals. This scheme has a shortcoming that information can be insufficient in highly volatile intraday time intervals and redundant at other moments of the day. Alternatively, one can apply a *transaction time sampling* scheme, such that the observations are sampled every kth transaction. This sampling scheme adapts naturally to the trading activity, and thus estimates have lower variability relative to estimates based on data sampled in calendar time; see Oomen (2006). In our application however, we use a sampling scheme that forms a series of de-volatilized intraday returns, through a sampling Oomen (2006) refers to as *business time sampling* (BT). This scheme dates back to Zhou (1992) who proposes to sample data each time some pre-specified amount of variation has realized. The resulting series is therefore equally informative. This scheme requires an estimate of intraday cumulative volatility. Since volatility is latent, we follow the idea of Fukasawa (2010) by applying cumulative volume as a proxy for cumulative latent volatility.[4] Thus, our sampling times $t_{0,N}, t_{1,N}, \ldots, t_{N,N}$ are such that amount of realized volume in each interval $(t_{i-1,N}, t_{i,N}]$ equals $(\textit{total volume})/N$, and we set $N \in \{250, 500\}$. Note that in this sampling scheme, we skip the first observation and volume of each day, as this observation contains all kind of overnight effects. On average, this corresponds to sampling every 48th and 24th transaction. Relative to calendar time sampling, we sample more frequently at the start and at the end of the trading day, while less fre-

[4] Alternatively, we also experimented using estimators of integrated variance robust to microstructure noise and jumps to obtain an estimate of cumulative variance. This delivered very similar results that do not change our conclusions. However, since the high-frequency based estimators depend on tuning parameters, and in general are considerably more complicated to obtain than cumulative volume, we report results based on the latter proxy only.

Table 5: Equity data: summary statistics on the number of consolidated trades

		2006			2007			2008			2009			2010	
Symbol	Raw	Clean	Avg	Raw	Clean	Avg	Raw	Clean	Avg	Raw	Clean	Avg	Raw	Clean	Avg
AA	2938194	1399359	5575	10484729	2730384	10878	19052862	3002379	11867	25252398	2428419	9637	19525993	1457355	5783
		(3095; 15103)			(3566; 35077)			(3521; 25190)			(3844; 31749)			(1251; 20235)	
AXP	2131247	1149936	4581	6960172	2466466	9827	18694570	3171851	12537	22100137	3507182	13917	13600280	2188410	8684
		(1849; 19541)			(3744; 24526)			(4199; 39268)			(2340; 43138)			(2525; 27365)	
BA	2332928	1297892	5171	6033728	2326756	9270	10433805	2581156	10202	9177067	2030485	8057	7490214	1763060	6996
		(1837; 10114)			(2827; 23193)			(3293; 24507)			(1613; 19235)			(1728; 19253)	
CAT	2917482	1358294	5412	6273666	2387487	9512	12101199	2488655	9837	16313099	2490239	9882	11081987	2218368	8803
		(2584; 11976)			(3027; 39266)			(1953; 21282)			(2280; 30249)			(2003; 26790)	
GE	7308481	1682671	6704	17890755	4073410	16229	53097581	6299728	24900	59023304	4494888	17837	33400067	2591851	10285
		(2384; 25629)			(7257; 42630)			(10238; 78211)			(2751; 58635)			(2390; 42812)	
HD	4908720	1473846	5872	12053593	2940748	11716	22558723	3210311	12689	19088604	2638725	10471	15549317	2159806	8571
		(2274; 21766)			(4413; 27376)			(3416; 30495)			(2672; 21788)			(2095; 33095)	
IBM	2811212	1533389	6109	7417523	3129273	12467	12485302	3202219	12657	11580628	2555039	10139	8888025	2398707	9519
		(2483; 15266)			(3981; 31309)			(3874; 30967)			(4036; 24791)			(2074; 27314)	
JPM	4127533	1484066	5913	14296841	3937709	15688	54047407	7229963	28577	68011683	6825146	27084	40436261	4016108	15937
		(2414; 12739)			(5414; 44706)			(6771; 89980)			(4561; 67384)			(4218; 47420)	
KO	2643426	1179259	4698	7068542	2313047	9215	15265008	3066471	12120	12916579	2472864	9813	11271635	1818047	7214
		(2246; 8131)			(2822; 20877)			(3010; 35184)			(3192; 26250)			(2632; 25352)	
MCD	2543832	1169975	4661	6436406	2028862	8083	12970858	2680197	10594	12452641	2298559	9121	8227844	1675680	6650
		(1830; 9538)			(3266; 20252)			(2942; 33469)			(2187; 23720)			(2104; 22560)	
PFE	9148652	1627340	6483	14623994	3235968	12892	27636290	3477087	13743	32199521	3298404	13089	25908356	1763808	6999
		(3022; 12164)			(4444; 21838)			(5410; 41934)			(2844; 45629)			(2446; 24496)	
PG	3064038	1624277	6471	8667109	2863358	11408	19369125	3658744	14461	16275148	2803493	11125	13201658	1994077	7913
		(3404; 16053)			(4046; 26284)			(6299; 37735)			(2499; 24520)			(2461; 27580)	
T	5183297	1663205	6626	13641001	3439970	13705	29034642	4198031	16593	25862175	3076086	12207	20289457	1790558	7105
		(1807; 18575)			(5348; 27409)			(5282; 45161)			(2334; 28679)			(2380; 22337)	
WMT	5264919	1614599	6433	12456917	3290958	13111	26919968	4266108	16862	22534757	3449871	13690	14660113	2383525	9458
		(3037; 10200)			(4236; 27298)			(4728; 39818)			(2897; 44283)			(3418; 24949)	
XOM	7873641	2946807	11740	20516820	6059941	24143	42366522	6983246	27602	33762439	5306792	21059	27371541	3830038	15199
		(5738; 35560)			(6030; 51552)			(10220; 70619)			(5818; 54789)			(3349; 57633)	

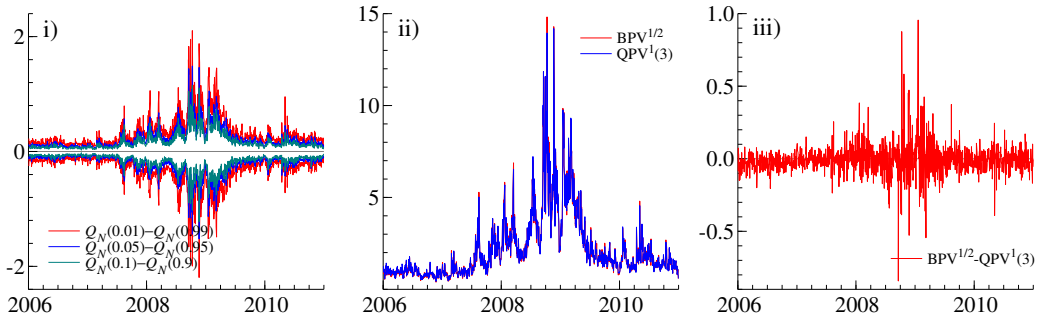
Note: The table reports the yearly number of trades before and after cleaning, and the average daily activity together with the minimum and maximum of number of trades between parentheses. The sample period is January 3 2006 to December 31 2010. The stocks are AA- Alcoa; AXP- American Express; BA- Boeing Company; CAT- Caterpillar; GE- General Electric; HD- Home Depot; IBM- Int. Business Machines; JPM- JPMorgan Chase & Co; KO- Coca-Cola; MCD- McDonald's; PFE- Pfizer; PG- Procter & Gamble; T- AT&T; WMT- Wal-Mart Stores; XOM- Exxon Mobil.

quently in the middle of the day. Our sample sizes also yield a reasonable compromise between potential microstructure effects and problems due to too small sample size. Actually, since the data is sampled sparsely, microstructure noise and price discreteness (as judged by first order serial correlation, volatility signature plots and histogram of price increments) are not expected to affect our results greatly.

5.2 Empirical results

We start by illustrating the time series of computed empirical quantiles. Panel *i*) of Figure 6 plots symmetric pairs of quantiles of intraday BT returns for the JPM equity. We observe a considerably larger spread of empirical quantiles around 2008 and 2009, the period corresponding to the subprime crisis. Panels *ii*) and *iii*) plot the volatility estimates of σ_t resulting from the bipower and quantile-based variation measures for all days in series, and their difference. The two measures correspond closely, although a more detailed inspection indicates that for several days the bipower variation deviates from the quantile-based measure. We observe that for these days the requirement that large price increments are preceded and succeeded by very small returns is not satisfied. This may in turn lead to loss of power of the BNS tests for these days.

Figure 6: Empirical quantiles and variation measures



We compute the ratio tests using intraday data from the trading days. As such, each day and each equity is treated on its own. Table 6 reports the fraction of days where one or more significant jumps are detected using the 5% significance level. The results can be summarized as follows. Using only $N = 250$ returns per day the rejections of the null hypothesis of no jumps are less frequent than for the sample size $N = 500$, as was also expected from the Monte Carlo results. At $N = 250$, the fraction of days where the second moment is affected by the presence of jumps is around 20-50%, and it increases to 50-80% for the third, fourth and sixth moment as judged by the BJ test. At $N = 500$, we find that the fraction of days with the second moment significantly affected by jumps is roughly 47-75%, and it increases to 75-90% for the higher moment BJ tests. It is seen that the three cases of $p \in \{1, 2, 3\}$ in general lead to similar results. We do not report results when $p > 3$ sub-estimators are used in the quantile-based measure, as the optimal quantile q_1 may be too extreme in this case, cf. Table 1. Indeed, when $N = 250$, the BJ test with $p = 3$ already is only robust up to a single outlier in a tail, hence for such small sample sizes it would be advisable to use fewer sub-estimators in the construction of the BJ test statistic.

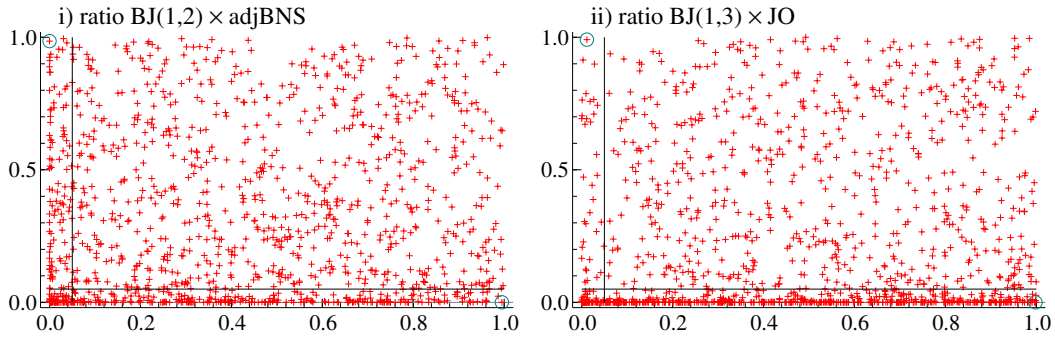
Table 6: Testing results for equity data

		$N = 250$								$N = 500$							
Symbol	BJ(p, $\cdot$)	BJ($\cdot$,1)	BJ($\cdot$,2)	BJ($\cdot$,3)	BJ($\cdot$,4)	BJ($\cdot$,6)	BNS	adj.BNS	JO	BJ($\cdot$,1)	BJ($\cdot$,2)	BJ($\cdot$,3)	BJ($\cdot$,4)	BJ($\cdot$,6)	BNS	adj.BNS	JO
AA	BJ(1,$\cdot$)	0.4138	0.4559	0.6616	0.7331	0.7633	0.2867	0.2836	0.1326	0.6728	0.6926	0.8276	0.8809	0.9039	0.6434	0.6426	0.1787
	BJ(2,$\cdot$)	0.5909	0.4631	0.6529	0.7117	0.6934				0.7975	0.6791	0.8610	0.8991	0.9023			
	BJ(3,$\cdot$)	0.7228	0.4273	0.6608	0.7315	0.6815				0.8848	0.6728	0.8554	0.8975	0.8952			
AXP	BJ(1,$\cdot$)	0.3948	0.3956	0.6529	0.7419	0.7681	0.2288	0.2264	0.1724	0.6473	0.6037	0.8523	0.9118	0.9301	0.4480	0.4480	0.1954
	BJ(2,$\cdot$)	0.6068	0.4146	0.6894	0.7601	0.7387				0.8451	0.5854	0.8856	0.9404	0.9357			
	BJ(3,$\cdot$)	0.7935	0.3193	0.6561	0.7355	0.6990				0.9245	0.5679	0.9118	0.9476	0.9309			
BA	BJ(1,$\cdot$)	0.3415	0.4392	0.7204	0.7951	0.8213	0.1994	0.1970	0.1851	0.6029	0.6037	0.9023	0.9484	0.9587	0.4122	0.4122	0.1859
	BJ(2,$\cdot$)	0.6259	0.3741	0.7299	0.7983	0.7935				0.8539	0.5886	0.9341	0.9635	0.9571			
	BJ(3,$\cdot$)	0.8165	0.2732	0.7220	0.7943	0.7554				0.9468	0.5330	0.9444	0.9658	0.9508			
CAT	BJ(1,$\cdot$)	0.3844	0.3074	0.5886	0.6751	0.7061	0.1875	0.1851	0.1350	0.6815	0.4790	0.7983	0.8753	0.8967	0.3757	0.3741	0.1565
	BJ(2,$\cdot$)	0.6346	0.2804	0.6021	0.6736	0.6473				0.8650	0.4662	0.8586	0.9190	0.9055			
	BJ(3,$\cdot$)	0.7800	0.2113	0.5902	0.6839	0.6235				0.9325	0.4639	0.8729	0.9150	0.8912			
GE	BJ(1,$\cdot$)	0.4400	0.4496	0.5330	0.5766	0.5909	0.3479	0.3384	0.1009	0.7164	0.7101	0.7554	0.7625	0.7641	0.6775	0.6767	0.1517
	BJ(2,$\cdot$)	0.5592	0.5171	0.5528	0.5536	0.5155				0.7307	0.7633	0.7919	0.7951	0.7760			
	BJ(3,$\cdot$)	0.5941	0.5203	0.5774	0.5751	0.5115				0.8110	0.7244	0.7562	0.7665	0.7331			
HD	BJ(1,$\cdot$)	0.3312	0.4567	0.6902	0.7665	0.7935	0.2145	0.2121	0.1612	0.6203	0.6561	0.8221	0.8888	0.9222	0.5004	0.4996	0.1755
	BJ(2,$\cdot$)	0.5711	0.4170	0.6712	0.7419	0.7379				0.7736	0.6743	0.8570	0.9142	0.9166			
	BJ(3,$\cdot$)	0.6966	0.3868	0.6680	0.7538	0.7141				0.8308	0.6743	0.8602	0.9095	0.9007			
IBM	BJ(1,$\cdot$)	0.2629	0.3090	0.5679	0.6680	0.7029	0.1525	0.1469	0.1533	0.5520	0.4948	0.8102	0.8880	0.9150	0.3709	0.3701	0.1827
	BJ(2,$\cdot$)	0.4821	0.2867	0.5774	0.6505	0.6489				0.8181	0.4790	0.8300	0.8944	0.8944			
	BJ(3,$\cdot$)	0.6751	0.2017	0.5536	0.6561	0.6172				0.8991	0.4543	0.8531	0.9134	0.8975			
JPM	BJ(1,$\cdot$)	0.2939	0.2836	0.4583	0.5314	0.5592	0.1557	0.1517	0.1048	0.5481	0.4940	0.6878	0.7760	0.8149	0.3606	0.3590	0.1366
	BJ(2,$\cdot$)	0.4265	0.3082	0.4901	0.5576	0.5385				0.7228	0.4829	0.7403	0.8173	0.8133			
	BJ(3,$\cdot$)	0.5870	0.2526	0.4535	0.5242	0.4837				0.8070	0.5028	0.7569	0.8149	0.7990			
KO	BJ(1,$\cdot$)	0.4083	0.4639	0.6640	0.7331	0.7720	0.2176	0.2153	0.1311	0.6195	0.6521	0.8253	0.9047	0.9301	0.4782	0.4782	0.1501
	BJ(2,$\cdot$)	0.5830	0.4392	0.6751	0.7458	0.7276				0.7768	0.6926	0.8594	0.9198	0.9222			
	BJ(3,$\cdot$)	0.7379	0.3884	0.6815	0.7538	0.7101				0.8443	0.6728	0.8880	0.9365	0.9245			
MCD	BJ(1,$\cdot$)	0.4432	0.4091	0.6624	0.7609	0.7768	0.2438	0.2407	0.1819	0.6704	0.6243	0.8697	0.9230	0.9365	0.5250	0.5250	0.2113
	BJ(2,$\cdot$)	0.6315	0.4241	0.6910	0.7593	0.7474				0.7998	0.6704	0.8896	0.9357	0.9373			
	BJ(3,$\cdot$)	0.7887	0.3336	0.7037	0.7736	0.7307				0.9047	0.6386	0.9047	0.9484	0.9349			
PFE	BJ(1,$\cdot$)	0.4901	0.5608	0.6481	0.6648	0.6831	0.3852	0.3828	0.1501	0.7705	0.7816	0.7911	0.8014	0.8110	0.7887	0.7887	0.1628
	BJ(2,$\cdot$)	0.5973	0.6068	0.6585	0.6807	0.6458				0.7562	0.8451	0.8435	0.8411	0.8229			
	BJ(3,$\cdot$)	0.6481	0.6370	0.6672	0.6855	0.6330				0.8658	0.7784	0.7847	0.7975	0.7705			
PG	BJ(1,$\cdot$)	0.3550	0.4392	0.6632	0.7514	0.7784	0.2272	0.2240	0.1739	0.5925	0.6195	0.8268	0.8872	0.9150	0.4543	0.4535	0.2065
	BJ(2,$\cdot$)	0.5504	0.4432	0.6966	0.7538	0.7395				0.7879	0.6775	0.8586	0.9095	0.9095			
	BJ(3,$\cdot$)	0.7268	0.3527	0.6831	0.7395	0.7021				0.8618	0.6434	0.8832	0.9261	0.9087			
T	BJ(1,$\cdot$)	0.4281	0.5068	0.6958	0.7625	0.7879	0.3002	0.2986	0.1922	0.7077	0.7117	0.8268	0.8769	0.9079	0.6211	0.6211	0.2343
	BJ(2,$\cdot$)	0.5703	0.5401	0.6910	0.7355	0.7284				0.7371	0.7458	0.8824	0.9134	0.9071			
	BJ(3,$\cdot$)	0.6910	0.4917	0.7006	0.7538	0.7061				0.8491	0.7141	0.8562	0.8975	0.8991			
WMT	BJ(1,$\cdot$)	0.2994	0.3511	0.5822	0.6680	0.7037	0.1795	0.1755	0.1295	0.5624	0.6322	0.8038	0.8769	0.9047	0.4512	0.4512	0.1525
	BJ(2,$\cdot$)	0.4527	0.4114	0.5973	0.6608	0.6521				0.7554	0.6132	0.8118	0.8785	0.8872			
	BJ(3,$\cdot$)	0.6021	0.3384	0.6013	0.6743	0.6211				0.8348	0.6052	0.8324	0.8904	0.8904			
XOM	BJ(1,$\cdot$)	0.1859	0.1922	0.3550	0.4265	0.4512	0.0985	0.0953	0.0953	0.4639	0.3376	0.5330	0.6227	0.6632	0.2788	0.2772	0.1104
	BJ(2,$\cdot$)	0.3550	0.2025	0.3487	0.4091	0.3987				0.6100	0.3820	0.5663	0.6450	0.6346			
	BJ(3,$\cdot$)	0.4392	0.1716	0.3407	0.4075	0.3717				0.7125	0.4067	0.6116	0.6545	0.6307			

Note: Table reports the fraction of days where one or more jumps are found, at a significance level of 5%.

Contrasting the BJ$(\cdot, 2)$ and BNS test statistics reveals that the BJ test finds one or more significant jumps on around 20% more days, when $N = 250$. For $N = 500$, this difference becomes considerably smaller. Surprisingly, the JO test rejects the null less frequently than the BJ$(\cdot, 3)$ and BNS tests, which suggests that both positive and negative jumps occur on the same day, offsetting each other. Indeed, a more detailed inspection of the data suggests that abnormally large price increments appear in both directions. We find that the fraction of days with identified jumps is slightly higher than reported by Gilder (2009) in his extensive study (based on the approach proposed by Andersen, Bollerslev, and Dobrev (2007)), also using US equities but over the period 2002-2006. As our sample period spans the financial crisis of 2007-2010, the jump process can be expected to be more active than in the earlier period as studied by Gilder (2009).

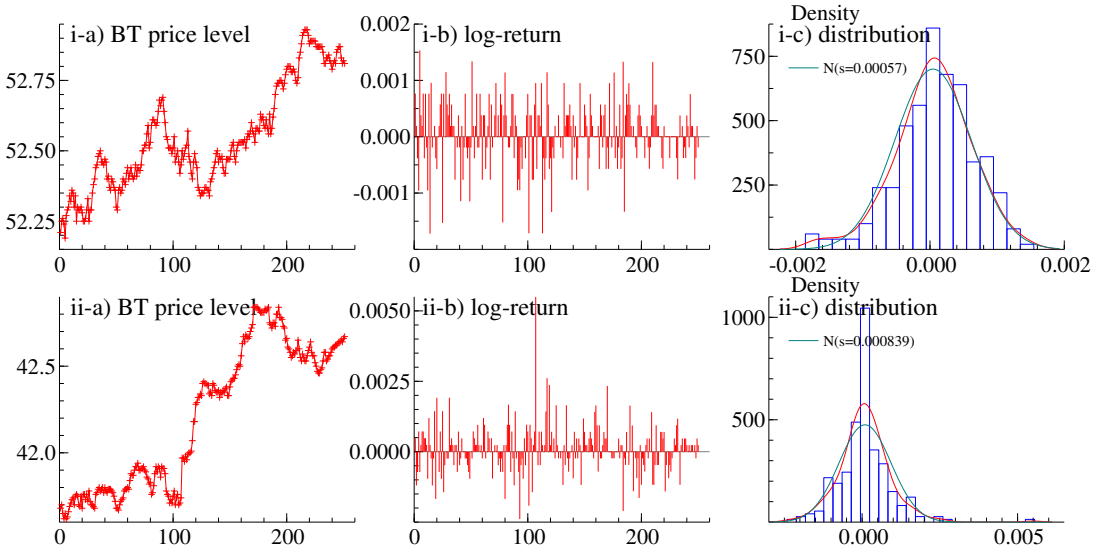
Figure 7: Scatter plots of p-values for the JPM equity



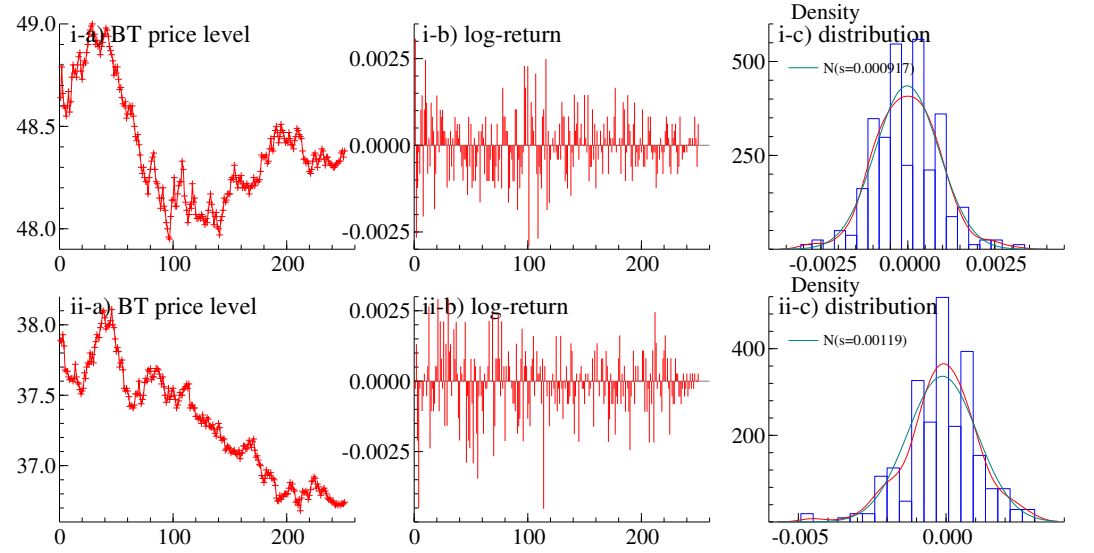
We study a few cases in more detail, in a similar way as Jiang and Oomen (2008). Figure 7 presents the scatter plot of the p-values of the test statistics considered for the JPM equity. Dotted lines present the nominal size of 5%. The lower left quadrant indicates instances where both the BJ and the adjusted BNS respectively the JO detect presence of jumps. The bottom right quadrant indicates when the BJ detects outliers while the alternative does not, while the top left quadrant shows the reverse. There is a substantial number of instances where the testing results overlap, but there are also days when they lead to opposite decisions. It is of interest to have a closer look at these cases. Specifically, we look at the most extreme instances as measured by the distance between the p-values of the two competing tests, marked by circles in top-left and bottom-right corners of the panels of Figure 7.

Table 7 reports p-values for the (adjusted) ratio tests for these most extreme instances. We focus on the second moment of the data and contrast BJ and BNS tests. Let us consider the April 25, 2007 case. We observe that for $r = 2$ and $p \in \{1, 2\}$ the BJ ratio tests detect one or more outliers, while (adj)BNS but also JO do not. Only when $p = 3$ (and using the quantiles prescribed by the asymptotic minimum variance criterion used throughout the article), the BJ test does not detect significant jumps. The smallest quantile used in this case is $q_1 = 0.069$, implying a robustness of the BJ test to at most a single outlier. The top panel of Figure 8a zooms in on this date, April 25, 2007. We observe that there are indeed several large negative returns, often preceded and/or followed by other relatively large returns. In such a situation, the bi- and quad-power variation can lose robustness, leading to a lower value of the BNS test statistic. This can explain the advantage of the BJ test over the BNS test.

Figure 8: Case studies for the JPM equity



(a) The figure displays price levels in business time, log-returns, and histogram/density of log-returns for the JPM equity; top panel 2007-04-25, bottom panel 2006-6-29.



(b) The figure displays price levels in business time, log-returns, and histogram/density of log-returns for the JPM equity; top panel 2007-03-30, bottom panel 2010-06-07.

Table 7: Ratio tests for the selected case studies

	BJ(·, 1)	BJ(·, 2)	BJ(·, 3)	BJ(·, 4)	BJ(·, 6)	BNS	adjBNS	JO
2007-04-25								
$p = 1$	0.0731	0.0002	0.0000	0.0000	0.0000	0.9950	0.9950	0.6098
$p = 2$	0.0009	0.0189	0.4565	0.9995	0.7821			
$p = 3$	0.0774	0.8508	0.2686	0.1926	0.3524			
2006-06-29								
$p = 1$	0.0019	0.9853	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
$p = 2$	0.0179	0.0077	0.0000	0.0000	0.0000			
$p = 3$	0.0634	0.0000	0.0000	0.0000	0.0000			
2007-03-30								
$p = 1$	0.0229	0.0012	0.0001	0.0000	0.0000	0.1397	0.1447	0.9982
$p = 2$	0.2114	0.3220	0.8750	0.4038	0.2710			
$p = 3$	0.9233	0.6750	0.2191	0.1348	0.2177			
2010-06-07								
$p = 1$	0.0099	0.1759	0.9911	0.4191	0.1223	0.7910	0.7917	0.0116
$p = 2$	0.2919	0.2612	0.0069	0.0013	0.0006			
$p = 3$	0.0261	0.6417	0.1512	0.0489	0.0432			

Note: The table reports p-values for the indicated tests.

For the June 29, 2006 case, the BJ ratio test with $p = 1$ does not identify jumps for $r = 2$, whereas other choices of p (and r) do seem to clearly find jump(s), as do the (adj)BNS and JO tests. Looking at the bottom panel Figure 8a, focusing on this day, we notice one single very large outlier (which is easily detectable by BNS or JO tests), but also a large spike in the histogram indicating many zero-returns. For many periods, the price remains flat, resulting in the normality assumption of the non-jump returns underlying the BJ test being broken. Using a single quantile in this case doesn't detect the outlier, combining multiple quantiles does detect the jump correctly.

We now focus on the case when the test is built on the third moment and we contrast the BJ and JO tests. On March 30, 2007, the ratio BJ(1,3) detects jumps, whereas the JO test sees no evidence at all. BJ tests using higher number of quantiles agree with the JO test, that no jumps are apparent on this date. The top panel of Figure 8b zooms in on the data. We observe that, apart from possibly the opening return of 0.3%, little is going on. The BJ(1,3) test seems to be confused by the relative lack of returns close to zero, hence the assumption of normality of the returns seems to be broken. Using the BJ(2,3) test robustifies against this effect.

On June 7, 2010 the JO test indicates the presence of jumps while the BJ$(1, 3)$ does not. The bottom panel of Figure 8b displays the returns for this date. There are two large negative returns, which show up in the left tail of the histogram in panel *ii-c*) of Figure 8b. On this particular day, the distribution of the returns (even excluding these two outliers) seems far from normal, leading to a bias on the QPV estimator. Notice that indeed that BJ tests do not agree among themselves, nor does the JO test agree with BNS, with a large variation in p-values for the tests. This is an indication that on such a day the JO/BNS and BJ tests should be used with care.

6 Conclusions

This article introduces the quantile-based measure of power variation of returns, which is robust against outlying (or abnormal) observations. We derive the joint asymptotic distribution of the quantile and moment-based measures of σ^r, for any positive value of r, under the assumption of normally distributed data. With the limiting distribution, we propose a new class of test statistics designed to detect outliers that lead to deviations

from normality. The test statistics exploit different powers r, and hence facilitate judging the impact of outliers on different moments of the data. The tests are easy to implement and can be applied in principle to any homoskedastic dataset where it is conjectured that only the presence of a finite number of outliers leads to non-normality.

We apply our theory to test for the presence of jumps in equity data over a sequence of trading days. In order to do so, we use the cumulative volume of the trades as a proxy for cumulative volatility, allowing us to devolatilize the returns to a homoskedastic series. Throughout the article, we compare the performance of our tests to widely applied alternatives: the tests of Barndorff-Nielsen et al. (2006) and of Jiang and Oomen (2008). The simulation results show that for higher orders of r our tests are more powerful than the alternative testing approaches. The empirical illustration shows that jumps can occur more frequent than judged by the second or third moment of the data only.

Our novel test has the additional advantage over existing tests that it can be applied using a range of powers of r, and for different numbers p of pairs of quantiles. If several combinations of p and r arrive at the same conclusion of (non-)rejection of the null hypothesis, then this could be considered a stronger signal than the (non-)rejection of only a single (existing or novel) test. The possibility of combining $\mathrm{BJ}(p, r)$ tests could be explored in future work, to derive the joint distribution of a collection of BJ tests, and to see if the test thus could be robustified against minor deviations from the underlying Gaussianity assumption.

In the empirical study, we alleviate the presence of microstructure noise in the high-frequency data by sampling the data more sparsely. On the one hand, the lower frequency decreases the ability to detect jumps. On the other hand, problems related to price discreteness and other microstructure effects can also adversely effect the quantile-based and moment-based estimators of σ^r, and these problems are evaded here by sparse sampling. Application of the same techniques on ultra high-frequency data, possibly using noise-reduction techniques as the pre-averaging of price increments proposed by Jacod et al. (2009), requires an extension of the asymptotic theory presented here, to allow for the dependence in data introduced by the noise reduction techniques. We leave this option for future research.

A Appendix

A.1 Proof of Proposition 2

In this proposition, it is the asymptotic covariance between the sample quantile power variation and sample absolute moment (or their powers) that is given. This asymptotic covariance term directly follows from the following theorem:

Theorem A.1. *Let $Y = \{Y_i\}_{i=1}^N$ be a series of* iid *random normal variables with mean μ, standard deviation σ, probability density function $f(y) = \frac{1}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)$ and inverse cumulative distribution function $F^{-1}(q) = \mu + \sigma\Phi^{-1}(q)$. Let $\boldsymbol{Q}_N(\boldsymbol{q})$ denote a p-vector of sample quantiles of Y and let M_N^r denote the estimator of the rth absolute sample moment Y, i.e. $M_N^r = N^{-1}\sum_{i=1}^N |Y_i - \mu|^r$. The joint asymptotic distribution of $\boldsymbol{Q}_N(\boldsymbol{q})$ and M_N^r is given by the bivariate normal distribution*

$$N^{1/2}\left(\begin{pmatrix}\boldsymbol{Q}_N(\boldsymbol{q})\\ M_N^r\end{pmatrix} - \begin{pmatrix}\boldsymbol{Q}(\boldsymbol{q})\\ \sigma^r M^r\end{pmatrix}\right) \xrightarrow{d} \mathcal{N}\left(0, \boldsymbol{S}(\sigma,r,p)'\tilde{\Sigma}(\boldsymbol{q},r)\boldsymbol{S}(\sigma,r,p)\right), \tag{30}$$

where

$$\tilde{\Sigma}(\boldsymbol{q},r) = \begin{bmatrix}\Sigma(\boldsymbol{q}) & \boldsymbol{s}(\boldsymbol{q},r)\\ \boldsymbol{s}(\boldsymbol{q},r)' & (M^{2r} - (M^r)^2)\end{bmatrix}, \qquad \boldsymbol{S}(\sigma,r,p) = \begin{bmatrix}\sigma I_p & 0\\ 0 & \sigma^r\end{bmatrix},$$

with $M_q^r = \int_{\Phi^{-1}(q)}^{\infty} |z|^r\phi(z)dz$, $M^r = M_0^r = \int_{-\infty}^{\infty} |z|^r\phi(z)dz$, and $\boldsymbol{s}(\boldsymbol{q},r)$ has typical element $s(q,r) = \frac{(M_q^r - (1-q)M^r)}{\phi(\Phi^{-1}(q))}$. Here $\phi(z)$ and $\Phi(z)$ are the probability density and cumulative distribution function of a standard normally distributed random variable z.

Proof. Our proof extends the results of Lin et al. (1980). Notice that the empirical quantile can be written as (see Bahadur, 1966)

$$Q_N(q) = Q(q) + \frac{1 - F_N(Q(q)) - (1-q)}{f(Q(q))} + R_{N,q},$$

with $F_N(y)$ the empirical distribution function of observations y. Here $R_{N,q}$ is a remainder term of order $O(N^{-3/4}\log N)$. Hence, the asymptotic behavior of a single $Q_N(q)$ can be related to the asymptotic behavior of the empirical density function, which is the average of a set of indicator functions as $1 - F_N(Q(q)) = \frac{1}{N}\sum I_{\{Y_i > Q(q)\}}$.

To look at the covariance between the indicator and summands of the absolute moment functions, notice that

$$\begin{pmatrix}I_{\{Y_i > Q(q)\}}\\ \left|\frac{Y_i-\mu}{\sigma}\right|^r\end{pmatrix} - \begin{pmatrix}1-q\\ M^r\end{pmatrix} \sim \text{IID}\left(0, I(q,r)\right),$$

with

$$I(q,r) = \begin{pmatrix}q(1-q) & M_q^r - (1-q)M^r\\ M_q^r - (1-q)M^r & M^{2r} - (M^r)^2\end{pmatrix}.$$

The covariances in this last equation are found from

$$\begin{aligned}
\operatorname{cov}(I_{\{Y_i>Q(q)\}}, \left|\frac{Y_i-\mu}{\sigma}\right|^r) &= \operatorname{cov}(I_{\{z>\Phi^{-1}(q)\}}, |z|^r) = \mathrm{E}\left(I_{\{z>\Phi^{-1}(q)\}} - (1-q)\right)(|z|^r - M^r) \\
&= \int (I_{\{z>\Phi^{-1}(q)\}} - (1-q))(|z|^r - M^r)\phi(z)dz \\
&= \int I_{\{z>\Phi^{-1}(q)\}}(|z|^r - M^r)\phi(z)dz - (1-q)\int (|z|^r - M^r)\phi(z)dz \\
&= \int I_{\{z>\Phi^{-1}(q)\}}|z|^r\phi(z)dz - M^r\int I_{\{z>\Phi^{-1}(q)\}}\phi(z)dz - (1-q)\times 0 \\
&= \int_{\Phi^{-1}(q)}^{\infty} |z|^r\phi(z)dz - M^r\int_{\Phi^{-1}(q)}^{\infty}\phi(z)dz \\
&= M_q^r - (1-q)M^r.
\end{aligned}$$

Using the central limit theorem, this implies that

$$N^{-\frac{1}{2}}\sum\left(\begin{pmatrix} I_{\{Y_i>Q(q)\}} \\ \left|\frac{Y_i-\mu}{\sigma}\right|^r \end{pmatrix} - \begin{pmatrix} 1-q \\ M^r \end{pmatrix}\right) = N^{\frac{1}{2}}\left(\begin{pmatrix} 1-F_N(Q(q)) \\ \frac{1}{N}\sum|z_i|^r \end{pmatrix} - \begin{pmatrix} 1-q \\ M^r \end{pmatrix}\right) \overset{d}{\sim} \mathcal{N}\left(0, I(q,r)\right).$$

After premultiplying by $\begin{bmatrix} \frac{\sigma}{\phi(\Phi^{-1}(q))} & 0 \\ 0 & \sigma^r \end{bmatrix}$, one finds that

$$N^{\frac{1}{2}}\begin{pmatrix} Q_N(q) - Q(q) - R_{N,q} \\ \frac{1}{N}\sum|Y_i-\mu|^r - \sigma^r M^r \end{pmatrix} \overset{d}{\sim} \mathcal{N}\left(0, S(\sigma,r,p)'\tilde{\Sigma}(q,r)S(\sigma,r,p)\right).$$

As $R_{N,q}$ goes to zero at a quicker rate than the remaining terms (Ghosh, 1971), asymptotically it has no influence and one arrives at the asymptotic density given in Theorem A.1, both for the case of a univariate q and also (with the necessary generalization in notation) for the p-variate vector of quantiles $\boldsymbol{q}$. □

The joint distribution of the QPV estimator and the moment-based estimator in Proposition 2 now follows directly from the representation of the estimators as functions of the vector of quantiles and moments in Theorem A.1, recognizing that indeed typical element $\xi_c(q,r)$ in (25) is related to the above results by

$$\xi_c(q,r) \equiv c^{-1}(q)D_1 s(\vec{q},r) = \frac{s(1-q,r) - s(q,r)}{c(q)}.$$

A.2 Review of alternative jump test statistics

This part reviews the tests of Barndorff-Nielsen et al. (2006) and Jiang and Oomen (2008). The presented test statistics are derived under the null of no jump in the price process. We first fix some notation. Let $\{P_i\}_{i=0}^N$ denote the asset prices and let $Y_i = \ln P_i$ denote the log price. Define $R_i = P_i/P_{i-1} - 1$ and $r_i = Y_i - Y_{i-1}$ as the simple and log return respectively. Then,

BNS tests:

i) the linear test:

$$\frac{N^{1/2}\Big(BPV_N - RV_N\Big)}{\sqrt{\vartheta\cdot QV_N}} \xrightarrow{d} \mathcal{N}\big(0,\,1\big), \tag{31}$$

ii) the logarithmic test:

$$\frac{N^{1/2} \cdot BPV_N \Big(\ln BPV_N - \ln RV_N \Big)}{\sqrt{\vartheta \cdot QV_N}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{32}$$

iii-a) the ratio test:

$$\frac{N^{1/2} \cdot BPV_N \Big(\big[BPV_N / RV_N \big] - 1 \Big)}{\sqrt{\vartheta \cdot QV_N}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{33}$$

iii-b) the adjusted ratio test:

$$\frac{N^{1/2} \Big(\big[BPV_N / RV_N \big] - 1 \Big)}{\sqrt{\vartheta \cdot \max\big(1, QV_N / (BPV_N)^2\big)}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{34}$$

where

$$RV_N = \sum_{i=1}^{N} r_i^2, \tag{35}$$

$$BPV_N = \mu_1^{-2} \frac{N}{N-1} \sum_{i=1}^{N-1} |r_i| |r_{i+1}|, \tag{36}$$

$$QV_N = \mu_1^{-4} N \sum_{i=0}^{N-4} \prod_{k=1}^{4} |r_{i+k}|, \tag{37}$$

and $\vartheta = (\pi^2/4) + \pi - 5$, $\mu_k = \mathbf{E}[|z|^k]$, $z \sim \mathcal{N}(0,1)$. The adjusted ratio test adapts for deviations from a intraday homoskedastic variance.

JO tests:

i) the linear test:

$$\frac{N \Big(SwV_N - RV_N \Big)}{\sqrt{\Omega_N^{SwV}(v)}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{38}$$

ii) the logarithmic test:

$$\frac{N \cdot BPV_N \Big(\ln SwV_N - \ln RV_N \Big)}{\sqrt{\Omega_N^{SwV}(v)}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{39}$$

iii) the ratio test:

$$\frac{N \cdot BPV_N \Big(1 - \big[RV_N / SwV_N \big] \Big)}{\sqrt{\Omega_N^{SwV}(v)}} \xrightarrow{d} \mathcal{N}\big(0,\, 1\big), \tag{40}$$

where

$$SwV_N = 2\sum_{i=1}^{N}(R_i - r_i), \tag{41}$$

$$\Omega_N^{SwV}(v) = \frac{\mu_6}{9}\frac{N^3\mu_{6/v}^{-v}}{N-v+1}\sum_{i=0}^{N-v}\prod_{k=1}^{v}|r_{i+k}|^{6/v}, \tag{42}$$

with either $v = 4$ or $v = 6$ in the variance estimator (42). Throughout the article, the variance estimator $\Omega_N^{SwV}(6)$ is used.

Acknowledgments

The authors wish to thank Peter Boswijk, Peter Hansen, and participants of the Amsterdam Econometric Seminars and Workshop Series for helpful comments and suggestions.

References

Aït-Sahalia, Y. and J. Jacod (2009). Testing for jumps in a discretely observed process. *Annals of Statistics 37*(1), 184–222.

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*(1), 125–180.

Bahadur, R. R. (1966). A note on quantiles in large samples. *Annals of Mathematical Statistics 37*(3), 570–580.

Bakshi, G., C. Cao, and Z. Chen (2000). Pricing and hedging long-term options. *Journal of Econometrics 94*(1-2), 277–318.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2009). Realised kernels in practice: Trades and quotes. *Econometrics Journal 12*(3), C1–C32.

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

Benson, F. (1949). A note on the estimation of mean and standard deviation from quantiles. *Journal of the Royal Statistical Society, Series B 11*(1), 91–100.

Bos, C. S., P. Janus, and S. J. Koopman (2012). Spot variance path estimation and its application to high frequency jump testing. *Journal of Financial Econometrics 10*(2), 354–389.

Brownlees, C. T. and G. M. Gallo (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics and Data Analysis 51*(4), 2232–2245.

Carr, P. and L. Wu (2003). What type of process underlies options? A simple robust test. *Journal of Finance 58*(6), 2581–2610.

Christensen, K., R. C. A. Oomen, and M. Podolskij (2010). Realised quantile-based estimation of the integrated variance. *Journal of Econometrics 159*(1), 74–98.

Corsi, F., S. Mittnik, C. Pigorsch, and U. Pigorsch (2008). The volatility of realized volatility. *Econometric Reviews 27*(1-3), 46–78.

Das, S. R. and R. K. Sundaram (1999). Of smiles and smirks: A term structure perspective. *The Journal of Financial and Quantitative Analysis 34*(2), 211–239.

DasGupta, A. and L. R. Haff (2006). Asymptotic values and expansions for the correlation between different measures of spread. *Journal of Statistical Planning and Inference 136*(7), 2197–2212.

David, H. A. (1998). Early sample measures of variability. *Statistical Science 13*(4), 368–377.

Eisenberger, I. and E. C. Posner (1965). Systematic statistics used for data compression in space telemetry. *Journal of the American Statistical Association 60*(309), 97–133.

Fukasawa, M. (2010). Realized volatility with stochastic sampling. *Stochastic Processes and their Applications 120*(6), 829–852.

Ghosh, J. K. (1971). A new proof of the Bahadur representation of quantiles and an application. *Annals of Mathematical Statistics 42*(6), 1957–1961.

Gilder, D. (2009). An empirical investigation of intraday jumps and cojumps in US equities. Working paper ID1343779, SSRN.

Griffin, J. E. and R. C. A. Oomen (2008). Sampling returns for realized variance calculations: Tick time or transaction time? *Econometric Reviews 27*, 230–253.

Hansen, P. R. and A. Lunde (2006). Realized variance and market microstructure noise. *Journal of Business and Economic Statistics 24*(2), 127–161.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica 46*(6), 1251–1271.

Hyndman, R. J. and Y. Fan (1996). Sample quantiles in statistical packages. *The American Statistician 50*(4), 361–365.

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

Jiang, G. J. and R. C. A. Oomen (2008). Testing for jumps when asset prices are observed with noise – A "swap variance" approach. *Journal of Econometrics 144*(2), 352–370.

Johannes, M. (2004). The statistical and economic role of jumps in continuous-time interest rate models. *Journal of Finance 59*(1), 227–260.

Jondeau, E., S.-H. Poon, and M. Rockinger (2007). *Financial modeling under non-Gaussian distribution*. Springer: Springer Finance.

Lee, S. S. and P. A. Mykland (2008). Jumps in financial markets: A new nonparametric test and jump dynamics. *Review of Financial Studies 21*(6), 2535–2563.

Lin, P.-E., K.-T. Wu, and I. A. Ahmad (1980). Asymptotic joint distribution of sample quantiles and sample mean with applications. *Communications in Statistics: Theory and Methods 9*(1), 51–60.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics 3*(1-2), 125–144.

Mosteller, F. (1946). On some useful "inefficient" statistics. *Annals of Mathematical Statistics 17*(4), 377–408.

Ogawa, J. (1951). Contributions to the theory of systematic statistics, I. *Osaka Mathematical Journal 3*(2), 175–213.

Oomen, R. C. A. (2006). Properties of realized variance under alternative sampling schemes. *Journal of Business and Economic Statistics 24*(2), 219–237.

Pearson, K. (1920). On the probable errors of frequency constants, III. *Biometrika 13*(1), 113–132.

Vergote, O. (2008). *Financial Transaction Data and Volatility*. Ph. D. thesis, Faculteit Economie en Bedrijfswetenschappen, Katholieke Universiteit Leuven.

Walker, A. M. (1968). A note on the asymptotic distribution of sample quantiles. *Journal of the Royal Statistical Society, Series B 30*(3), 570–575.

Zhou, B. (1992). Forecasting foreign exchange rates subject to de-volatilization. Working paper 3510, MIT Sloan School.