DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 4/10

Robust Estimation of Integrated Variance and Quarticity under Flat Price and No Trading Bias

by

Frowin C. Schulz

2nd version
August 7, 2010

DISKUSSIONSBEITRÄGE ZUR STATISTIK UND ÖKONOMETRIE

SEMINAR FÜR WIRTSCHAFTS- UND SOZIALSTATISTIK
UNIVERSITÄT ZU KÖLN

Albertus-Magnus-Platz, D-50923 Köln, Deutschland

DISCUSSION PAPERS IN STATISTICS AND ECONOMETRICS

SEMINAR OF ECONOMIC AND SOCIAL STATISTICS
UNIVERSITY OF COLOGNE

No. 4/10

Robust Estimation of Integrated Variance and Quarticity under Flat Price and No Trading Bias[1]

by

Frowin C. Schulz[2]

2^{nd} version
August 7, 2010

Abstract

This paper investigates a selection of methods disentangling contributions from price jumps to realized variance. Flat prices (consecutively sampled prices in calendar time with the same value) and no trading (no price observation at sampling points), both frequently occurring stylized facts in financial high-frequency datasets, can cause a considerable bias in each considered method. Hence, we propose an approach to robustify those methods so that they can provide undistorted statistical results based on intraday intervals not influenced by flat prices and no trading. The new approach is tested in realistic Monte Carlo experiments and shows to be extraordinary robust against varying levels of flat price and no trading bias. Additionally, we examine the new approach empirically with a dataset of electricity forward contracts traded on the Nord Pool Energy Exchange. We obtain coherent conclusions with respect to predefined qualitative indicators.

Keywords: Realized Variance, Zero-Returns, Price Jumps, Robust Estimation, High-Frequency Data, Electricity Forward Contract.

JEL: C12, C13, C14, G10.

---

[1]I would like to thank Annemie Denzer, Oliver Grothe, Stefan Klößner, Sébastien Laurent, Karl Mosler, Roberto Renó, Martin Ruppert and an anonymous referee for detailed suggestions, helpful comments and insightful discussions. Moreover, I am grateful to the conference participants at the Second Trondheim Winter Energy Workshop 2010 (Norway, March 4-5) and at the Fourth International Conference on Mathematical and Statistical Methods for Actuarial Sciences and Finance 2010 (Ravello, April 7-9) for discussions.

[2]University of Cologne, Research Training Group Risk Management, Albertus-Magnus-Platz, D-50923 Cologne, Germany, tel: +49 (0)221 470 6334, email: f.schulz@wiso.uni-koeln.de. Schulz gratefully acknowledges financial support provided by the German Research Foundation (DFG). The usual disclaimer applies.

1 Introduction

One of the risk measures, which repeatedly draws the attention in the field of science, is realized variance. It is a model-free ex-post measure for high-frequency data. This measure approximates the total quadratic variation of a financial asset over an interval $[0, t]$. The underlying theoretical price path for a financial asset is well described by a continuous-time stochastic volatility jump diffusion process. Empirical evidence for this is found in Eraker, Johannes and Polsen (2003) and Eraker (2004). To measure the contribution of a finite number of price jumps to realized variance, Barndorff-Nielsen and Shephard (2004a, 2006a) develop a test statistic. Therewith, an operator can decide, based on a statistical framework, whether price jumps occur within a trading day and how much they contribute to realized variance. This information is used to state more precisely conclusions about risk. In an extensive empirical application, Andersen, Bollerslev and Diebold (2007), Andersen, Bollerslev and Huang (2007) and Bollerslev, Kretschmer, Pigorsch and Tauchen (2009) show the economic value (i.e. improved forecasts for realized variance) of utilizing the separate information about the continuous and discontinuous component of realized variance in a time series model. Further nonparametric methods to separate the discontinuous component (also referred to as jump factor) from realized variance are elaborated by e.g. Aït-Sahalia and Jacod (2009), Andersen, Dobrev and Schaumburg (2008, 2009), Christensen, Oomen and Podolskij (2009), Corsi, Pirino and Renò (2009), Jiang and Oomen (2008), Lee and Mykland (2008) and Mancini (2009). Due to the fact that there exists a variety of methods, an operator may question which method shall be preferred in an empirical application. Hence, it is of interest to identify potential shortcomings of each method.

We recognize that most methods are based on the assumption of a continuous-time stochastic volatility jump diffusion price process, and are constructed for efficient prices sampled on equidistant discrete time grids due to feasibility.[1] Typical in empirical applications, however, is that the discretely sampled price process randomly switches from one sampling point to the next between two states. We term those states as either observable or latent. In the observable state we can assign an efficient price to a sampling point. The latent state enters in case of two frequently occurring empirical stylized facts. The first one is referred to as *flat prices*, i.e. consecutively sampled prices in calendar time with the same value, well addressed by Phillips and Yu (2008). According to Phillips and Yu (2008), flat prices can be understood as inefficient or noisy prices, since the occurrence of flat prices has zero probability if we assume a log-price process constituting a semimartingale. The second occurrence is *no trading*, i.e. no price observation over a certain period of time within a trading day, stressed by Corsi, Pirino and Renò (2009). That means, in the latent state we can either assign an inefficient price to a sampling point or no price at all. In either case, the price information does not represent the efficient price and is therefore latent. Consequently, the computation of continuously compounded interval returns, required to implement the jump detection methods, is only possible if two consecutive prices are in the observable state. Those interval returns are referred to as observable variation or return process fragments. However, if there is at least one latent price

[1] One possible sampling method is the previous tick method by Hansen and Lunde (2003, 2006).

at two consecutive sampling points, we cannot compute the corresponding interval return. In applications, we usually set this interval return to zero. Therefore, the latent return process fragment is likewise called zero-return.

Schulz and Mosler (2010) illustratively show that the existence of even a small percentage amount of zero-returns greatly distort the statistical conclusions by employing the method of Barndorff-Nielsen and Shephard (2004a, 2006a). Schulz and Mosler (2010) introduce a first ad-hoc approach to reduce the distortion. Alternatively, a price data manipulation method is employed by Barndorff-Nielsen and Shephard (2004b) in advance. This is a calibrated statistical model which simulates a price at a sampling point as soon as no trading, flat prices or even very small price movements occur. However, such a proceeding of stochastic or alternatively deterministic interpolation is quite critical as we add variation, which is actually latent, i.e. unknown to us. As such, there remains enough space putting further effort into analyzing the sensitivity of other methods with respect to zero-returns, and (if required) define a modified method with two properties. First, it should not require having to add variation for the latent fragments. Second, it should be robust against the distorting impact of the latent return process fragments on measuring price jumps within the observable return process fragments.

This paper contrasts four methods disentangling contributions from price jumps to realized variance. Employed comparable methods are by Barndorff-Nielsen and Shephard (2004a, 2006a), Corsi, Pirino and Renò (2009) and Andersen, Dobrev and Schaumburg (2009). For each method, zero-returns are a pivotal source of statistical distortion. Therefore, the first contribution of this paper is the introduction of a new approach, which robustifies each considered method to latent return process fragments so that they provide undistorted statistical conclusion for the observable variation. In this paper, our approach is referred to as **sustained integrated variance and quarticity estimation** *(SIVQE)*. Under ideal conditions of no zero-returns, we theoretically show that the asymptotic distribution of each method with *SIVQE* remains the same with respect to its original counterpart. Subsequently, we describe a return process with observable and latent fragments by a Bernoulli process. Here, we show that the original multipower variation based integrated variance and quarticity estimators by Barndorff-Nielsen and Shephard (2004a, 2006a) and Corsi, Pirino and Renò (2009) underestimate the actual quadratic variation of the observable return process fragments. However, implementing these methods with *SIVQE* does not produce an underestimation. Furthermore, we show that in case of no price jumps the difference between realized variance and the robustified multipower variation based integrated variance estimator of the observable variation converges in probability to zero for a decreasing sampling length.

The second contribution is that *SIVQE* is tested in Monte Carlo experiments under imperfect market conditions, i.e. market scenarios with different levels of flat price and no trading bias. To be more precise, we initially question to what extent the convergence criteria of the test statistic under the null hypothesis holds for return series with an increasing fraction of zero-returns. Astonishingly, even for more than 50% zero-returns, the convergence criteria is quite robust for each method. Beyond that, we investigate the accuracy of the (no) jump day detection rate for an increasing fraction of zero-returns with respect to the ideal case of no zero-returns, and find that the detection rates are also considerably robust. To analyze the overall performance of detecting days both with and without jumps across methods, we employ a nonparametric sensitivity index ($\mathcal{A}$) typically used in signal detection theory.[2] The results show that *SIVQE* in combination with the corresponding method definitely performs better across all zero-return levels than implementing the original methods. It even constitutes a better performer than the approach proposed by Schulz and Mosler (2010).

The final contribution of this study is to discuss the empirical relevance of the robust approach. The implemented time series is a high-frequency dataset of electricity forward contracts traded on the Nord Pool Energy Exchange. The traded contracts are of substantial economic relevance for the Nordic electricity market.[3] Besides, the empirical price process is characterized by flat prices and no trading. The focus is to analyze whether each original method yields the same amount of days with price jumps, and how trading days with detected price jumps can be characterized. The analysis shows that given a conventional level of significance, a very heterogeneous picture is produced with respect to the amount of days with jumps. Even the 5% most potential jump factors of each method greatly diverge in size and occurrence time. These potential jump days are typically characterized by below full-sample average trading activity and small amount of extreme price movements, indicating a large fraction of spurious price jumps. In light of the just mentioned, we analyze the empirical results for *SIVQE* and find more plausible conclusions by using again trading activity as a qualitative variable to discuss evidence of jump occurrences. Moreover, there are indications for preferring *SIVQE* in combination with the method of Corsi, Pirino and Renò (2009) as the qualitative indicators are strongest. Finally, we find a markedly increased intersection in occurrence time of the 5% most potential jump factors.

The remainder of the paper is organized as follows. In the next section, the concept of realized variance and relevant methods to separate the jump factor from realized variance are discussed. The proceeding of *SIVQE* is described in section 3. Moreover, evaluations from the Monte Carlo experiments with(out) the employed robust approach are discussed. Section 4 gives insights into the empirical analysis and section 5 concludes.

2 Basics of Quantifying Price Variability

2.1 Concept of Realized Variance

The logarithmic price is denoted $X(t)$. The price expansion is assumed to be well described by the following continuous-time stochastic volatility jump diffusion process

$$dX(t) = \mu(t)dt + \sigma(t)dW(t) + \kappa(t)dq(t) , \qquad t \in [0,1] , \tag{1}$$

where $\mu(t)$ is the drift term, $\sigma(t)$ is a strictly positive stochastic càdlàg process and $W(t)$ is a

[2] $\mathcal{A}$ incorporates the probabilities of (not) correctly detecting days with and without jumps. $\mathcal{A} \in [0,1]$, where 1 (0) is the best (worst) possible outcome. See Zhang and Mueller (2005) for further details.

[3] The reader is referred to Schulz and Mosler (2010) for further details.

standard Brownian motion. $\kappa(t)$ is the size of a discrete jump in time t in the log price process and $q(t)$ is a counting process with finite activity and (possibly) time-varying intensity $\lambda(t)$. The associated realized price variability over a predetermined period of time, here $[t\text{-}h, t]$ with $0 < h \leq t \leq 1$, is defined as follows (Andersen, Bollerslev and Diebold, 2002):

$$NV_t \equiv \underbrace{\int_{t-h}^{t} \sigma^2(s)ds}_{\text{continuous variation}} \quad + \quad \underbrace{\sum_{q_{t-h}<s\leq q_t} \kappa^2(s)}_{\text{jump factor}} , \tag{2}$$

where $\sigma^2(s)$ is the instantaneous return variation, $\kappa^2(s)$ is the squared size of a discrete jump in time t. Typically, h is set to one, representing one trading day. The ex-post variability measure in equation (2) is called **notional variance** and is composed of two parts. The first part, denoted as continuous variation or **integrated variance** (IV_t), is the quadratic variation of the Brownian motion in equation (1) over $[t\text{-}h, t]$. Correspondingly, the **jump factor** represents the quadratic variation of the Poisson process.

In order to evaluate the time variable exposure for discretely sampled prices, Andersen and Bollerslev (1998) motivate a model-free ex-post measure for high-frequency data, called **realized variance**. For $h = 1$,

$$RV_t \equiv \sum_{j=1}^{M} r_j^2 , \qquad \text{with} \quad r_j := r_{j,t,M} := X\left(\frac{j\,t}{M}\right) - X\left(\frac{(j-1)t}{M}\right) , \tag{3}$$

where $M \in \mathbb{N}^+$ determines the interval length for intraday returns r_j. Realized variance is under the maintained assumptions a consistent nonparametric estimator for the notional variance. In this respect, realized variance converges for $M \to \infty$ in probability to the notional variance:

$$RV_t \quad \xrightarrow{p} \quad \int_{t-1}^{t} \sigma^2(s)ds + \sum_{q_{t-1}<s\leq q_t} \kappa^2(s) .$$

If we further assume a mean of zero for the underlying return process, realized variance is also an unbiased estimator of the ex-ante expected variance, the key interest of practitioners in financial markets. Even if we relax the assumption to a stochastically evolving mean return process over the predetermined interval, the statement remains approximately true.[4]

2.2 Methods Separating Jump Factor from Realized Variance

In the literature, there exist several methods to measure contributions from price jumps to realized variance. Considered methods in this study are by Barndorff-Nielsen and Shephard (2004a, 2006a) (henceforth BNS), Corsi, Pirino and Renò (2009) (henceforth CPR) and Andersen, Dobrev and Schaumburg (2009) (henceforth ADS). In principle, each implemented method proceeds in a similar fashion.

First, it matters to establish a consistent estimator for integrated variance ($\widehat{IV}_{t,\nu}$) which is robust against a finite number of jumps over a finite period of time. An intuitive jump measure is then simply the difference between RV_t and $\widehat{IV}_{t,\nu}$. This difference is meant to converge for

[4]See Andersen, Bollerslev and Diebold (2002) for a detailed discussion.

$M \to \infty$ in probability to the jump factor in equation (2). However, Andersen, Bollerslev and Diebold (2007) point out that due to measurement errors in practice, this jump measure likely yields inapt and even incorrect outputs. In order to handle such finite sample problems, BNS (2004a, 2005, 2006a) propose a statistic, testing for the following null hypothesis:

H_0: *No jumps are present in the underlying price process versus* H_1: $\neg\, H_0$.

Here, the test statistic of interest is based on a relative jump measure and is defined as:

$$Z_{t,\nu} = \frac{(RV_t - \widehat{IV}_{t,\nu})/RV_t}{\sqrt{\frac{\vartheta_\nu}{M} \max\left\{1, \frac{\widehat{IQ}_{t,\nu}}{(\widehat{IV}_{t,\nu})^2}\right\}}} \quad \xrightarrow{\text{d}} \quad N(0,1)\;, \qquad \nu = 1,2,3,4\;. \tag{4}$$

where ν is an index for the implemented method. $\widehat{IQ}_{t,\nu}$ factored with ϑ_ν/M denotes a consistent estimator for the asymptotic variance of $RV_t - \widehat{IV}_{t,\nu}$. For $M \to \infty$,

$$\widehat{IQ}_{t,\nu} \quad \xrightarrow{p} \quad IQ_t \equiv \int_{t-1}^{t} \sigma^4(s)ds\;, \tag{5}$$

that is $\widehat{IQ}_{t,\nu}$ is converging in probability to its theoretical counterpart IQ_t, called **integrated quarticity**. The asymptotic variance of $RV_t - \widehat{IV}_{t,\nu}$ depends on the efficiency of $\widehat{IV}_{t,\nu}$. Huang and Tauchen (2005) show in their simulation study that $Z_{t,\nu}$ has good power, using the specification of BNS (2004a, 2006a) for $\widehat{IV}_{t,\nu}$ and $\widehat{IQ}_{t,\nu}$, shortly presented below. This result motivates the implementation of $Z_{t,\nu}$ for each method ν. By implementing $Z_{t,\nu}$, we can make inference about the significance of the difference between RV_t and $\widehat{IV}_{t,\nu}$, and therewith statistically separate contributions from continuous variation and price jumps to realized variance on a daily basis. Jumps are detected, if the test statistic $Z_{t,\nu}$ is greater than a predetermined quantile function $(\Phi^{-1}_{1-\alpha})$ say for $\alpha \le 5\%$. The jump factor and integrated variance amount to:

$$J_{t,\nu} \equiv [RV_t - \widehat{IV}_{t,\nu}]\mathbb{1}_{\{Z_{t,\nu} > \Phi^{-1}_{1-\alpha}\}}\;, \qquad \text{and} \qquad C_{t,\nu} \equiv RV_t - J_{t,\nu}\;.$$

where $\mathbb{1}$ is an indicator function, equaling one if $Z_{t,\nu} > \Phi^{-1}_{1-\alpha}$, and zero else. After having presented key insights concerning the general proceeding of each method, we will briefly present the respective specifications for $\widehat{IV}_{t,\nu}$, $\widehat{IQ}_{t,\nu}$ and ϑ_ν, besides further details if required.

<u>*Method: BNS*</u>

The first considered estimator for integrated variance is theoretically derived by BNS (2004a). This estimator is called **bipower variation**, defined as:

$$BP_{t,i} \equiv \varphi_1 \left(\frac{M}{M-1-i}\right) \sum_{j=2+i}^{M} |r_{j-(1+i)}||r_j|\;, \qquad i \ge 0\;, \tag{6}$$

where $\varphi_1 = \pi/2$. The robustness property of bipower variation against a finite number of price jumps is due to the following fact. Asymptotically, for $M \to \infty$, there is maximally one jump in the infinitesimal small adjacent interval returns r_j and r_{j-1}. Furthermore, it is imperative that if there is a jump in r_{j-1} its impact will vanish as it is multiplied by a subsequent return r_j of order $1/\sqrt{M}$. Obviously, for $M \to \infty$, r_j gets extremely small reducing the impact of the jump in r_{j-1} to a negligible amount. An empirical application to a selection of financial time series by Andersen, Bollerslev and Diebold (2007) suggests to choose $i = 1$. This means that bipower variation sums up the cross-products of absolute interval returns with one lag. Directly adjacent absolute interval returns (i.e. $i = 0$) are not chosen for computation due to their potential serial correlation, which in turn might bias bipower variation. Analytical evidence of this issue is provided by Huang and Tauchen (2005), who assume a noisy price process. A consistent estimator for IQ_t, termed **tripower quarticity** and employed by Andersen, Bollerslev and Diebold (2007), formulates as follows:

$$TriP_{t,i} \equiv \Delta_i \sum_{j=1+2(1+i)}^{M} |r_{j-2(1+i)}|^{4/3} |r_{j-(1+i)}|^{4/3} |r_j|^{4/3} , \qquad i \geq 0 , \tag{7}$$

where $\Delta_i = M\left(2^{2/3} \cdot \Gamma(\frac{7}{6}) \cdot \Gamma(\frac{1}{2})^{-1}\right)^{-3} \left(\frac{M}{M-2(1+i)}\right)$. The adjustment parameter ϑ_1 for the asymptotic variance in equation (4) equals to $\varphi_1^2 + 2\varphi_1 - 5$. Despite the theoretical appeal of this method, it is influenced by nonnegligible issues in finite samples.[5] These drawbacks are mainly due to the concept of bipower variation, and provoked by microstructure noise and a finite choice of M. The following methods explicitly name and seize up these pitfalls, and suggest potential alternatives.

Method: CPR

CPR (2009) propose an estimator for IV_t and IQ_t by forming a combination of the multipower variation concept, introduced in several papers by BNS,[6] and the threshold approach by Mancini (2009). Broadly speaking, the general idea is to initially trim or correct the return series with a threshold function before computing bipower variation and tripower quarticity. They motivate their approach with a drawback of bipower variation stressed by ADS (2009). Asymptotically, the concept of bipower variation works fine. However, for finite M the impact of a jump in r_{j-1} does not completely vanish, causing a positive distortion of the bipower variation measure. Moreover, a potential appearance of jumps in two adjacent returns cannot be excluded. In order to circumvent such a finite sample issue, CPR (2009) propose to filter out large jump occurrences in the interval return series with a threshold function before computing bipower variation. The threshold function is defined as:

$$\theta := \theta_\tau^\delta = c_\Theta^2 \cdot \Theta_\tau^\delta ,$$

composed of a fixed scaling factor c_Θ, typically set to three,[7] and a local Kernel smoothed and jump controlled variance estimator Θ_τ^δ with the following specification:

$$\Theta_\tau^\delta = \frac{\sum_{j=-L,\, j\neq -1,0,1}^{L} K\left(\frac{j}{L}\right) r_j^2 \, \mathbb{1}_{\{r_j^2 \leq c_\Theta^2 \cdot \Theta_j^{\delta-1}\}}}{\sum_{j=-L,\, j\neq -1,0,1}^{L} K\left(\frac{j}{L}\right) \mathbb{1}_{\{r_j^2 \leq c_\Theta^2 \cdot \Theta_j^{\delta-1}\}}} , \qquad \delta = 1, 2, 3, \ldots , \tag{8}$$

[5] The drawbacks in practical applications are explicitly discussed by CPR (2009), ADS (2009) and Schulz and Mosler (2010).

[6] BNS (2006b) review their recent contributions to multipower variation and reference corresponding papers.

[7] See CPR (2009) or Mancini and Renò (2008); the choice of the scaling factor is the most critical point in the method of CPR (2009) as it is a predetermined exogenous variable.

where $\tau = \{1, 2, 3, ..., T \cdot M\}$, T denotes the total amount of trading days in the sample, and L is the bandwidth parameter determining the window around τ to estimate the local variance. $\mathbb{1}$ is an indicator function, equaling one if $r_j^2 \leq c_\Theta^2 \cdot \Theta_j^{\delta-1}$, and zero else. For K, a Gaussian kernel with the form $K\left(\frac{j}{L}\right) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\left(\frac{j}{L}\right)\frac{1}{2}\right\}$ is proposed by CPR (2009). Furthermore, they set $L = 25$ and point out that the choice of L is not crucial.

The final threshold function is computed iteratively. Squared returns smaller than or equal to the threshold are kept in the series, whereas all others are set to zero. Under the null, this series is biased as it is possible that normal *iid* returns are greater than the threshold. Due to this fact, CPR (2009) propose not to set $r_j = 0$ if $r_j^2 > \theta$ but to replace r_j with its conditional expected value under the null, $E\left(|r_j|^\lambda \middle| r_j^2 > \theta\right)$. The final series of absolute interval returns raised to the λ-th power now can be defined as:

$$\Psi_\lambda(r_j, \theta) = \begin{cases} |r_j|^\lambda & \text{if } r_j^2 \leq \theta \\ \frac{1}{2M(-c_\Theta)\sqrt{\pi}} \left(\frac{2}{c_\Theta^2}\theta\right)^{\frac{\lambda}{2}} \Gamma\left(\frac{\lambda+1}{2}, \frac{c_\Theta^2}{2}\right) & \text{if } r_j^2 > \theta \end{cases} \quad ; \quad \lambda = 1, 4/3\,.$$

Generally, **threshold bipower variation** and **tripower quarticity** is formulated as:

$$\begin{aligned} TBP_{t,i} &\equiv \varphi_2 \left(\frac{M}{M-1-i}\right) \sum_{j=2+i}^{M} \Psi_1(r_{j-(1+i)}, \theta)\, \Psi_1(r_j, \theta)\,, \qquad i \geq 0\,, \\ TTriP_{t,i} &\equiv \Delta_i \sum_{j=1+2(1+i)}^{M} \Psi_{4/3}(r_{j-2(1+i)}, \theta)\, \Psi_{4/3}(r_{j-(1+i)}, \theta)\, \Psi_{4/3}(r_j, \theta)\,, \qquad i \geq 0\,. \end{aligned}$$

In empirical application, it is also advisable to use staggered returns, following the discussion above. An important final note to the method of CPR (2009) is that the asymptotic theory does not change from multipower variation to threshold multipower variation. Therefore, φ_2 and ϑ_2 equal to φ_1 and ϑ_1.

Method: ADS-Min and ADS-Med

Even two jump robust estimators for integrated variance, called $\boldsymbol{MinRV_t}$ and $\boldsymbol{MedRV_t}$, are proposed by ADS (2009):

$$\begin{aligned} MinRV_t &\equiv \varphi_3 \left(\frac{M}{M-1}\right) \sum_{j=2}^{M} \min\left(|r_{j-1}|, |r_j|\right)^2\,, \\ MedRV_t &\equiv \varphi_4 \left(\frac{M}{M-2}\right) \sum_{j=3}^{M} \text{med}\left(|r_{j-2}|, |r_{j-1}|, |r_j|\right)^2\,, \end{aligned}$$

where $\varphi_3 = \frac{\pi}{\pi-2}$, $\varphi_4 = \frac{\pi}{6-4\sqrt{3}+\pi}$ and med $\hat{=}$ median. An appealing theoretical property of both estimators in finite sample applications is that the bias as in bipower variation emanating due to large jumps is here less intense. Moreover, according to ADS (2009) $MedRV_t$ is meant to be less exposed to zero-returns than any presented jump robust estimator. The corresponding estimators for integrated quarticity are:

$$\begin{aligned} MinRQ_t &\equiv M\, \frac{\pi}{3\pi-8} \left(\frac{M}{M-1}\right) \sum_{j=2}^{M} \min\left(|r_{j-1}|, |r_j|\right)^4\,, \\ MedRQ_t &\equiv M\, \frac{3\pi}{9\pi+72-52\sqrt{3}} \left(\frac{M}{M-2}\right) \sum_{j=3}^{M} \text{med}\left(|r_{j-2}|, |r_{j-1}|, |r_j|\right)^4\,. \end{aligned}$$

At this point we supplement two propositions to formulate a test statistic for jumps.

Proposition 1: *$MinRQ_t$ and $MedRQ_t$ are jump robust and consistent estimators for integrated quarticity. For $M \to \infty$, it is*

$$MinRQ_t \quad \xrightarrow{p} \quad \int_{t-1}^{t} \sigma^4(s)ds \ , \quad and \quad MedRQ_t \quad \xrightarrow{p} \quad \int_{t-1}^{t} \sigma^4(s)ds \ .$$

Referring to ADS (2009), the asymptotic theory required to prove *Proposition 1* is entirely analogous and results similar to their Propositions 1-3. Based on Proposition 3 of ADS (2009) and *Proposition 1*, we can straightforwardly derive the asymptotic distribution of the difference between RV_t and $MinRV_t$ ($MedRV_t$):

Proposition 2: *Given the joint asymptotic results for $MinRV_t$ and $MedRV_t$ derived by ADS (2009), the asymptotic distribution of '$(RV_t - MinRV_t)/RV_t$' and '$(RV_t - MedRV_t)/RV_t$' is:*

$$Z_{t,3} = \frac{(RV_t - MinRV_t)/RV_t}{\sqrt{\frac{1.81}{M} \max\left\{1, \frac{MinRQ_t}{(MinRV_t)^2}\right\}}} \quad \xrightarrow{d} \quad N(0,1) \ ,$$

$$Z_{t,4} = \frac{(RV_t - MedRV_t)/RV_t}{\sqrt{\frac{0.96}{M} \max\left\{1, \frac{MedRQ_t}{(MedRV_t)^2}\right\}}} \quad \xrightarrow{d} \quad N(0,1) \ .$$

Proof. We derive the asymptotic variance of $RV_t - MinRV_t$, using the joint asymptotic distribution of RV_t and $MinRV_t$ (ADS, 2009). After that, we deduce the distribution of $RV_t - MinRV_t$. Finally, we formulate the test statistic as in *Proposition 2*. The proof for $MedRV_t$ proceeds analogously, utilizing the joint asymptotic distribution of RV_t and $MedRV_t$ (ADS, 2009).

□

3 Dealing with Flat Prices and No Trading

3.1 Sustained Integrated Variance and Quarticity Estimation

In Section 1 and 2 we stressed that each considered method assumes a continuous price process as in equation (1), and is constructed for discretely sampled efficient prices. This is a crucial assumption for each integrated variance and quarticity estimator as it rules out the presence of zero-returns. The question we are now dealing with is how the estimators can be robustified if the discretely sampled price process consists of randomly occurring observable and latent states. In the observable state we can sample an efficient price for a time grid. However, in the latent state we either sample an inefficient price (flat price) or no price at all (no trading). Latent states translate into zero-returns and cause a bias in each integrated variance and quarticity estimator, but not in RV_t, which is shown below. This bias implicates the distortion in the test statistics. For such conditions we suggest **sustained integrated variance and quarticity estimation** *(SIVQE)* which robustifies each method with respect to the impact of flat prices and no trading on detecting price jumps. Note, if *SIVQE* is applied, the notation of the method, and integrated variance and quarticity estimator will start with a *'S'*. The proceeding of the new approach is explicitly illustrated for the respective integrated variance estimators $BP_{t,i}$, $TBP_{t,i}$, $MinRV_t$ and $MedRV_t$. In order to set up the idea of *SIVQE*, let us resume some of the assumptions and theoretical results of each method under ideal conditions of no zero-returns.

Under the null hypothesis, assume that $r_j \overset{\text{iid}}{\sim} N(0,\sigma^2)$; it follows $E(|r_j|) = \sqrt{\frac{2}{\pi}}\sigma$. Beyond that, we assume that the second and fourth moments of $|r_j|$ exist. We further know from BNS (2004a, p.10), CPR (2009, pp.4-8), and ADS (2009, p.7) that each increment of the integrated variance estimator, produced by the respective method, delivers an unbiased estimate of the corresponding underlying spot variance. The same applies to realized variance, where the spot variance is σ_j^2 with its unbiased estimate r_j^2. It is also known that the difference between the sum of spot variances of the RV_t concept and the respective method converges for $M \to \infty$ in probability limit to zero. These results are summarized in the following:

$\widehat{IV}_t$	Unbiased spot variance estimate	Convergence result for $M \to \infty$
$BP_{t,i}$,	$\varphi_1 \, \lvert r_{j-(1+i)}\rvert \, \lvert r_j \rvert$,	$\sum_{j=1}^{M} \sigma_j^2 - \sum_{j=2+i}^{M} \sigma_{j-(1+i)} \, \sigma_j \xrightarrow{p} 0$,
$TBP_{t,i}$,	$\varphi_2 \, \Psi_1(r_{j-(1+i)},\theta) \, \Psi_1(r_j,\theta)$,	$\sum_{j=1}^{M} \sigma_j^2 - \sum_{j=2+i}^{M} \sigma_{j-(1+i),\theta} \, \sigma_{j,\theta} \xrightarrow{p} 0$,
$MinRV_t$,	$\varphi_3 \, \min(\lvert r_{j-1}\rvert, \lvert r_j\rvert)^2$,	$\sum_{j=1}^{M} \sigma_j^2 - \sum_{j=2}^{M} \min(\sigma_{j-1}, \sigma_j)^2 \xrightarrow{p} 0$,
$MedRV_t$,	$\varphi_4 \, \text{med}(\lvert r_{j-2}\rvert, \lvert r_{j-1}\rvert, \lvert r_j\rvert)^2$,	$\sum_{j=1}^{M} \sigma_j^2 - \sum_{j=3}^{M} \text{med}(\sigma_{j-2}, \sigma_{j-1}, \sigma_j)^2 \xrightarrow{p} 0$.

In section 2, we specified that each integrated variance estimator approximates the continuous variation over $[t-1,t]$ (under the null hypothesis, this is also true for RV_t). Of importance for *SIVQE* is to discuss whether each unbiased spot variance estimate can be referred to an interval j. If this is not directly possible, we will elaborate a proxy. Starting with RV_t, it is obvious that r_j^2 is an unbiased spot variance estimate for interval j. For $BP_{t,i}$, the increment $\varphi_1 \, |r_{j-(1+i)}| \, |r_j|$, is not the most favorable choice as the interval return j occurs in the subsequent increment as well. As such, the subsequent increment $\varphi_1 \, |r_j| \, |r_{j+(1+i)}|$ is likewise unfavorable. Therefore, we propose to rewrite the increment so that it is approximately equal to r_j^2. One natural proxy is to take half of each cross-product including r_j, i.e. $\frac{1}{2}\,\varphi_1 |r_{j-1}| \, |r_j| + \frac{1}{2}\,\varphi_1 |r_j| \, |r_{j+1}|$, which averages the influence of $|r_{j-1}|$ and $|r_{j+1}|$. Analogously, we proceed for the unbiased spot variance estimates of the remaining methods. For each method, we propose the following adapted increments, which are likewise approximately equal to r_j^2:

$$\left.\begin{array}{ll} BP_{t,i}: & \frac{1}{2}\,\varphi_1 |r_{j-1}| \, |r_j| + \frac{1}{2}\,\varphi_1 |r_j| \, |r_{j+1}| \\ TBP_{t,i}: & \frac{1}{2}\,\varphi_2 \, \Psi_1(r_{j-1},\theta) \, \Psi_1(r_j,\theta) + \frac{1}{2}\,\varphi_2 \, \Psi_1(r_j,\theta) \, \Psi_1(r_{j+1},\theta) \\ MinRV_t: & \frac{1}{2}\,\varphi_3 \, \min(|r_{j-1}|,|r_j|)^2 + \frac{1}{2}\,\varphi_3 \, \min(|r_j|,|r_{j+1}|)^2 \\ MedRV_t: & \frac{1}{3}\,\varphi_4 \, \text{med}(|r_{j-2}|,|r_{j-1}|,|r_j|)^2 + \frac{1}{3}\,\varphi_4 \, \text{med}(|r_{j-1}|,|r_j|,|r_{j+1}|)^2 \\ & \quad + \frac{1}{3}\,\varphi_4 \, \text{med}(|r_j|,|r_{j+1}|,|r_{j+2}|)^2 \end{array}\right\} \overset{\text{approx.}}{\sim} r_j^2 . \quad (9)$$

The outlined adaption in equation (9) for interval j can be done for all intervals M within $[t-1,t]$. Up to now, except rewriting and linking unbiased spot variance estimates to a certain interval j, the integrated variance estimators has been left unchanged. That means, the sum over $[t-1,t]$ of the adapted increments remains unchanged. But why is the outlined consideration in equation (9) supportive to define *SIVQE*? This abstraction is the basis to understand the nature of the bias and to locate where the robust approach should intervene, if we do not have ideal conditions of no zero-returns. In order to understand the whole purpose, we come to the issue emanating from zero-returns. We illustrate with two examples the origin of the bias and introduce the main idea of how *SIVQE* solves for the bias.

Example 1:

Imagine the following first exemplary case, which is solely one or one of many potential instances of flat prices and no trading within a trading day. In this example, we only choose an extract of a trading day with five consecutive intraday return intervals to keep the illustration simple and intuitive. That means, we only focus on the unbiased spot variance estimates resulting from this part of the series. Obviously, more complex zero-return patterns are possible:

$$\dots\,, \quad |r_{j-2}| = 0\,, \quad |r_{j-1}| = 0\,, \quad |r_j| > 0\,, \quad |r_{j+1}| = 0\,, \quad |r_{j+2}| = 0\,, \quad \dots \quad .$$

The resulting sum of spot variance estimates for realized variance and the other methods in the original setting are as follows (for $i = 0$):

$$\begin{array}{lll} RV_t: & 0+0+r_j^2+0+0 & = \; r_j^2\,, \\ BP_{t,0}\,, TBP_{t,0}\,, MinRV_t: & 0+0+0+0 & = \; 0\,, \\ MedRV_t: & 0+0+0 & = \; 0\,. \end{array}$$

Despite the fact that there is observable variation in interval j, $BP_{t,0}$, $TBP_{t,0}$, $MinRV_t$ and $MedRV_t$ fail to capture it, whereas realized variance does not. This is why we mentioned above that realized variance is not biased or is robust as the unbiased spot variance estimate r_j^2 of the observable return process fragment is not influenced by the previous or following latent return process fragment and completely captures the observable variation. In the example, this is not the case for the other methods as they require at least two consecutive return intervals with observable variation. To solve for this shortcoming we could simply draw a random variable or any feasible value for $|r_{j-2}|$, $|r_{j-1}|$, $|r_{j+1}|$ and $|r_{j+2}|$ to bridge the gap. However, we want to avoid these circumstances, i.e. define an approach insuring that the estimates reflect only but completely the observable variation like RV_t, despite the alignment of returns as in the example. Beyond that, the approach has to be defined in such a manner that no additional variation is externally added, else RV_t would increase as well.

Bearing the conditions in mind, we propose to exploit the relationship of equation (9). This means to find for r_j^2, the only spot variance estimate of realized variance greater than zero, a corresponding estimate produced by $BP_{t,0}$, $TBP_{t,0}$, $MinRV_t$ and $MedRV_t$. For each integrated variance estimator we suggest,

$$\left.\begin{array}{ll} BP_{t,0}: & \frac{1}{2}\,\varphi_1\,\wp_{j-1,1}\,|r_j| \;+\; \frac{1}{2}\,\varphi_1\,|r_j|\,\wp_{j+1,1} \\ TBP_{t,0}: & \frac{1}{2}\,\varphi_2\,\wp_{j-1,2}\,\Psi_1(r_j,\theta) \;+\; \frac{1}{2}\,\varphi_2\,\Psi_1(r_j,\theta)\,\wp_{j+1,2} \\ MinRV_t: & \frac{1}{2}\,\varphi_3\,\min\left(\wp_{j-1,3},|r_j|\right)^2 \;+\; \frac{1}{2}\,\varphi_3\,\min\left(|r_j|,\wp_{j+1,3}\right)^2 \\ MedRV_t: & \frac{1}{3}\,\varphi_4\,\mathrm{med}\left(\wp_{j-2,4},\wp_{j-1,4},|r_j|\right)^2 \;+\; \frac{1}{3}\,\varphi_4\,\mathrm{med}\left(\wp_{j-1,4},|r_j|,\wp_{j+1,4}\right)^2 \\ & \;+\; \frac{1}{3}\,\varphi_4\,\mathrm{med}\left(|r_j|,\wp_{j+1,4},\wp_{j+2,4}\right)^2 \end{array}\right\} \overset{\text{approx.}}{\sim} \; r_j^2\,,$$

where $\wp_{\cdot,\nu}$ is understood as a required sustainer that the variation, attributed by r_j, does not completely vanish while estimating the spot variance for interval j.

□

Reasonable estimators for $\wp_{j,\nu}$ will be discussed below. For illustration purposes, we will go through another pattern of zero-returns to understand where *SIVQE* intervenes.

Example 2:

In this example, we imagine to observe a short extract of a trading day with six consecutive interval returns:

$$\dots\,, \quad |r_{j-2}| = 0\,, \quad |r_{j-1}| = 0\,, \quad |r_j| > 0\,, \quad |r_{j+1}| > 0\,, \quad |r_{j+2}| = 0\,, \quad |r_{j+3}| = 0\,, \quad \dots \quad .$$

Skipping the part of computing the spot variance estimates with the original proceeding, we exploit once more the relationship of equation (9). Thus, we define a corresponding estimate of $BP_{t,0}$, $TBP_{t,0}$, $MinRV_t$ and $MedRV_t$ for the spot variance estimates r_j^2 and r_{j+1}^2. To get the intuition, we will only present the estimates for $BP_{t,0}$ and $MedRV_t$, as $TBP_{t,0}$ and $MinRV_t$ are from a conceptual point of view close to $BP_{t,0}$. The corresponding sum of adapted spot variance estimates for interval j and $j+1$ of $BP_{t,0}$ are

$$\begin{aligned}
&\underbrace{\tfrac{1}{2}\,\varphi_1\,\wp_{j-1,1}|r_j| \;+\; \tfrac{1}{2}\,\varphi_1\,|r_j||r_{j+1}|}_{\overset{\text{approx.}}{\sim}\; r_j^2} \;+\; \underbrace{\tfrac{1}{2}\,\varphi_1\,|r_j||r_{j+1}| \;+\; \tfrac{1}{2}\,\varphi_1\,|r_{j+1}|\wp_{j+2,1}}_{\overset{\text{approx.}}{\sim}\; r_{j+1}^2} \\
=\;& \underbrace{\tfrac{1}{2}\,\varphi_1\,\wp_{j-1,1}|r_j| \;+\; \varphi_1\,|r_j||r_{j+1}| \;+\; \tfrac{1}{2}\,\varphi_1\,|r_{j+1}|\wp_{j+2,1}}_{\overset{\text{approx.}}{\sim}\; r_j^2 + r_{j+1}^2}\,,
\end{aligned}$$

whereas for $MedRV_t$ we derive the following:

$$\begin{aligned}
&\underbrace{\tfrac{1}{3}\,\varphi_4\,\mathrm{med}(\wp_{j-2,4}\,, \wp_{j-1,4}\,, |r_j|)^2 + \tfrac{1}{3}\,\varphi_4\,\mathrm{med}(\wp_{j-1,4}\,, |r_j|\,, |r_{j+1}|)^2 + \tfrac{1}{3}\,\varphi_4\,\mathrm{med}(|r_j|\,, |r_{j+1}|\,, \wp_{j+2,4})^2}_{\overset{\text{approx.}}{\sim}\; r_j^2} + \\
&\underbrace{\tfrac{1}{3}\,\varphi_4\,\mathrm{med}(\wp_{j-1,4}\,, |r_j|\,, |r_{j+1}|)^2 + \tfrac{1}{3}\,\varphi_4\,\mathrm{med}(|r_j|\,, |r_{j+1}|\,, \wp_{j+2,4})^2 + \tfrac{1}{3}\,\varphi_4\,\mathrm{med}(|r_{j+1}|\,, \wp_{j+2,4}\,, \wp_{j+3,4})^2}_{\overset{\text{approx.}}{\sim}\; r_{j+1}^2} \\
&= \tfrac{1}{3}\,\varphi_4\,\mathrm{med}(\wp_{j-1,4}\,, |r_j|\,, |r_{j+1}|)^2 \;+\; \tfrac{2}{3}\,\varphi_4\,\mathrm{med}(|r_j|\,, |r_{j+1}|\,, \wp_{j+2,4})^2 \\
&\quad + \tfrac{2}{3}\,\varphi_4\,\mathrm{med}(|r_j|\,, |r_{j+1}|\,, \wp_{j+2,4})^2 \;+\; \tfrac{1}{3}\,\varphi_4\,\mathrm{med}(|r_{j+1}|\,, \wp_{j+2,4}\,, \wp_{j+3,4})^2\,.
\end{aligned}$$

□

General Definition of SIVQE:

So far, we have elaborated the functionality of *SIVQE* for two scenarios of observable and latent return intervals within a trading day. Incorporating the just mentioned ideas for any observable and latent return patterns in the concept of each integrated variance estimator, we define **sustained bipower variation** ($SBP_{t,i}$), **sustained threshold bipower variation** ($STBP_{t,i}$), **sustained MinRV** ($SMinRV_t$) and **sustained MedRV** ($SMedRV_t$). For brevity, we summarize the estimator for $SBP_{t,i}$, $STBP_{t,i}$ and $SMinRV_t$ in $S\widehat{IV}_{t,\nu}$ as follows:

$$S\widehat{IV}_{t,\nu} \;=\; \xi_\nu \sum_{j_\nu}^{M} \left(\mathbb{1}_{a1}\,\hat{r}^2_{j,\nu} \;+\; \tfrac{1}{2}\,\mathbb{1}_{a2}\,\hat{r}^2_{j,\nu,\wp_{a2}} \;+\; \tfrac{1}{2}\,\mathbb{1}_{a3}\,\hat{r}^2_{j,\nu,\wp_{a3}} \right)\,, \tag{10}$$

where $\nu = 1, 2, 3$. The indicator functions $\mathbb{1}_{a1}, \mathbb{1}_{a2}$ and $\mathbb{1}_{a3}$ are defined as:

$$
\begin{aligned}
\mathbb{1}_{a1} &= \begin{cases} 1 & \text{if } (|r_{j-(1+i)}||r_j| > 0) \lor (|r_{j-(1+i)}| = |r_j| = 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{a2} &= \begin{cases} 1 & \text{if } (|r_{j-(1+i)}| > 0 \ \land \ |r_j| = 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{a3} &= \begin{cases} 1 & \text{if } (|r_{j-(1+i)}| = 0 \ \land \ |r_j| > 0) \\ 0 & \text{else} \end{cases}.
\end{aligned} \tag{11}
$$

For the respective ν, the following specifications apply:

ν	ξ_ν	j_ν	$\hat{r}^2_{j,\nu}$	$\hat{r}^2_{j,\nu,\wp_{a2}}$	$\hat{r}^2_{j,\nu,\wp_{a3}}$	i
1	$\frac{M}{M-1-i}\frac{\pi}{2}$	$2+i$	$\lvert r_{j-(1+i)}\rvert\lvert r_j\rvert$	$\lvert r_{j-(1+i)}\rvert\wp_{j,1}$	$\wp_{j-(1+i),1}\lvert r_j\rvert$	≥ 0
2	$\frac{M}{M-1-i}\frac{\pi}{2}$	$2+i$	$\Psi_1(r_{j-(1+i)},\theta)\Psi_1(r_j,\theta)$	$\Psi_1(r_{j-(1+i)},\theta)\wp_{j,2}$	$\wp_{j-(1+i),2}\Psi_1(r_j,\theta)$	≥ 0
3	$\frac{M}{M-1}\frac{\pi}{\pi-2}$	2	$\min(\lvert r_{j-1}\rvert,\lvert r_j\rvert)^2$	$\min(\lvert r_{j-1}\rvert,\wp_{j,3})^2$	$\min(\wp_{j-1,3},\lvert r_j\rvert)^2$	0

To specify $SMedRV_t$, we need further conditions as we have to consider three consecutive interval returns. We define it as:

$$
\begin{aligned}
SMedRV_t = \frac{\pi}{6-4\sqrt{3}+\pi}\left(\frac{M}{M-2}\right)\sum_{j=3}^{M}\Big[&\mathbb{1}_{b1}\,\mathrm{med}\left(|r_{j-2}|\,,\,|r_{j-1}|\,,\,|r_j|\right)^2 \\
&+\mathbb{1}_{b2}\,\mathrm{med}\left(\wp_{j-2,4}\,,\,\wp_{j-1,4}\,,\,|r_j|\right)^2\tfrac{1}{3} + \mathbb{1}_{b3}\,\mathrm{med}\left(\wp_{j-2,4}\,,\,|r_{j-1}|\,,\,\wp_{j,4}\right)^2\tfrac{1}{3} \\
&+\mathbb{1}_{b4}\,\mathrm{med}\left(|r_{j-2}|\,,\,\wp_{j-1,4}\,,\,\wp_{j,4}\right)^2\tfrac{1}{3} + \mathbb{1}_{b5}\,\mathrm{med}\left(|r_{j-2}|\,,\,|r_{j-1}|\,,\,\wp_{j,4}\right)^2\tfrac{2}{3} \\
&+\mathbb{1}_{b6}\,\mathrm{med}\left(|r_{j-2}|\,,\,\wp_{j-1,4}\,,\,|r_j|\right)^2\tfrac{2}{3} + \mathbb{1}_{b7}\,\mathrm{med}\left(\wp_{j-2,4}\,,\,|r_{j-1}|\,,\,|r_j|\right)^2\tfrac{2}{3}\Big],
\end{aligned}
$$

The indicator functions $\mathbb{1}_{b1}, \mathbb{1}_{b2}, \ldots, \mathbb{1}_{b7}$ are defined as:

$$
\begin{aligned}
\mathbb{1}_{b1} &= \begin{cases} 1 & \text{if } (|r_{j-2(1+i)}||r_{j-(1+i)}||r_j| > 0) \lor (|r_{j-2(1+i)}| = |r_{j-(1+i)}| = |r_j| = 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{b2} &= \begin{cases} 1 & \text{if } (|r_{j-2(1+i)}| = |r_{j-(1+i)}| = 0 \ \land \ |r_j| > 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{b3} &= \begin{cases} 1 & \text{if } (|r_{j-2(1+i)}| = 0 \ \land \ |r_{j-(1+i)}| > 0 \ \land \ |r_j| = 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{b4} &= \begin{cases} 1 & \text{if } (|r_{j-2(1+i)}| > 0 \ \land \ |r_{j-(1+i)}| = |r_j| = 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{b5} &= \begin{cases} 1 & \text{if } ((|r_{j-2(1+i)}| \ \land \ |r_{j-(1+i)}| > 0) \ \land \ |r_j| = 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{b6} &= \begin{cases} 1 & \text{if } (|r_{j-2(1+i)}| > 0 \ \land \ |r_{j-(1+i)}| = 0 \ \land \ |r_j| > 0) \\ 0 & \text{else} \end{cases}, \\
\mathbb{1}_{b7} &= \begin{cases} 1 & \text{if } (|r_{j-2(1+i)}| = 0 \ \land \ (|r_{j-(1+i)}| \ \land \ |r_j| > 0)) \\ 0 & \text{else} \end{cases},
\end{aligned} \tag{12}
$$

where i is set to zero for $SMedRV_t$. In order to make $S\widehat{IV}_{t,\nu}$ ($\nu = 1, 2, 3, 4$) feasible, we suggest to set $\wp_{j,\nu} = E(|r_j|)$. $\wp_{j,\nu}$ can be consistently estimated on day t by $\hat{\wp}_{j,\nu}$, formulated as:

$$\hat{\wp}_{j,\nu} = \frac{1}{N} \sum_{j=1}^{M} |r_j| \,, \quad \text{with} \qquad N = \sum_{j=1}^{M} \mathbb{1}_{\{|r_j|>0\}} \,. \tag{13}$$

The estimator in equation (13) is unbiased if there are no price jumps. However, this is not true in the presence of price jumps. In such cases, $\hat{\wp}_{j,\nu}$ is estimated based on the local Kernel smoothed and jump controlled spot variance estimator by CPR (2009):

$$\hat{\wp}_{j,\nu} = \sqrt{\tfrac{2}{\pi}\, \Theta_\tau^\delta} \,, \tag{14}$$

where Θ_τ^δ is the specified in equation (8).

In *Example 1* and *2*, we illustrated the effect of zero-returns on each $\widehat{IV}_{t,\nu}$. Yet, there is a similar or even more severe effect on the estimators for integrated quarticity. To formulate robustified estimators of integrated quarticity, the same idea as outlined above applies. Each so-called sustained integrated quarticity estimator ($S\widehat{IQ}_{t,\nu}$) is explicitly defined in the appendix.

By reviewing the newly defined integrated variance and quarticity estimators, it seems that every potential source of bias due to flat prices and no trading is eliminated, however, not for *SCPR*. The computation of the threshold (θ) is executed beforehand and employs the original return series with all the zero-returns. Given equation (8), the local Kernel smoothed and jump controlled variance estimator Θ_τ^δ converges to an extremely small value for an increasing fraction of zero-returns. As a consequence, the final threshold (θ) turns out to be artificially small. Therefore, it is advisable to exclude all zero-returns for the computation of the threshold function.[8]

Asymptotic Behavior of SIVQE:

Before proceeding with implementing *SIVQE* in a Monte Carlo experiment, two additional issues have to be addressed. The first question is whether the asymptotic results remain unaffected in the ideal case of no zero-returns. This means, we need to question whether $S\widehat{IV}_{t,\nu}$ and $S\widehat{IQ}_{t,\nu}$ (for $\nu = 1, 2, 3, 4$) converge in probability to integrated variance in equation (2) and integrated quarticity in equation (5), and whether the asymptotic distribution coincide with their original counterpart. The answer to this is that in the ideal case of no zero-returns we can rewrite each $S\widehat{IV}_{t,\nu}$ and $S\widehat{IQ}_{t,\nu}$ estimator in its corresponding original format. Therefore, the existing asymptotic results hold. Clearly, the asymptotic distribution coincides as well and we can implement the same test statistic as before.

The second concern is to show that in case of observable and latent return process fragments, the observable integrated variance and quarticity are underestimated by the original estimators, whereas not by *SIVQE*. For this, the following proposition is formulated. We decompose RV_t and $(S)BP_{t,0}$ into observable and latent quadratic variation.

[8]Previous simulation results showed that the threshold (θ) gets artificially close to zero due to an increasing fraction of zero-returns.

Proposition 3: *Assume a price process as of equation (1) without price jumps. Furthermore, assume that the return process consists of observable* (o) *and latent* (ℓ) *fragments, whereas the two states are described by a Bernoulli process,*

$$r_j = \begin{cases} r_j^{(o)} & \textit{if} \quad \gamma_j = 1\,, \\ r_j^{(\ell)} & \textit{if} \quad \gamma_j = 0\,. \end{cases}$$

γ_j *is a Bernoulli sequence independent of the return process,* $E(\gamma_j = 1) = \zeta$ *and* $\zeta \in (0,1]$*.* $r_j^{(o)}$ *is an observable return, whereas* $r_j^{(\ell)}$ *is latent. Then realized variance decomposes as:*

$$RV_t \;=\; \sum_{j=1}^{M} r_j^{2\,(o)}\,\gamma_j \;+\; \sum_{j=1}^{M} r_j^{2\,(\ell)}\,(1-\gamma_j) \;=\; RV_t^{(o)} \;+\; RV_t^{(\ell)}\,,$$

where $RV_t^{(o)}$ *stands for the observable quadratic variation. Correspondingly,* $RV_t^{(\ell)}$ *is the latent one. Further, bipower variation decomposes as in the following:*

$$\begin{aligned} BP_{t,0} \;&=\; \frac{M}{M-1}\Bigg[\varphi_1\sum_{j=2}^{M}|r_{j-1}^{(o)}|\,|r_j^{(o)}|\,\gamma_{j-1}\,\gamma_j \;+\; \varphi_1\sum_{j=2}^{M}|r_{j-1}^{(\ell)}|\,|r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j \\ &\qquad +\; \varphi_1\sum_{j=2}^{M}|r_{j-1}^{(o)}||r_j^{(\ell)}|\,\gamma_{j-1}\,(1-\gamma_j) \;+\; \varphi_1\sum_{j=2}^{M}|r_{j-1}^{(\ell)}||r_j^{(\ell)}|\,(1-\gamma_{j-1})\,(1-\gamma_j)\Bigg] \\ &=\; BP_{t,0}^{(o)} \;+\; BP_{t,0}^{(\ell)^*,1} \;+\; BP_{t,0}^{(\ell)^*,2} \;+\; BP_{t,0}^{(\ell),3}\,, \end{aligned}$$

where $BP_{t,0}^{(o)}$ *is the observable quadratic variation, whereas the rest is the latent part. If* $0 < \zeta < 1$*,* $RV_t^{(o)} > BP_{t,0}^{(o)}$*. Finally, sustained bipower variation decomposes as:*

$$\begin{aligned} SBP_{t,0} \;&=\; \frac{M}{M-1}\Bigg[\varphi_1\sum_{j=2}^{M}|r_{j-1}^{(o)}|\,|r_j^{(o)}|\,\gamma_{j-1}\,\gamma_j \;+\; \varphi_1\sum_{j=2}^{M}\tfrac{1}{2}\,\wp_{j-1}\,|r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j \\ &\qquad +\varphi_1\sum_{j=2}^{M}\tfrac{1}{2}\,|r_{j-1}^{(o)}|\,\wp_j\,\gamma_{j-1}\,(1-\gamma_j) \;+\; \varphi_1\sum_{j=2}^{M}\tfrac{1}{2}\,|r_{j-1}^{(\ell)}|\,|r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j \\ &\qquad +\varphi_1\sum_{j=2}^{M}\tfrac{1}{2}\,|r_{j-1}^{(o)}|\,|r_j^{(\ell)}|\,\gamma_{j-1}\,(1-\gamma_j) \;+\; \varphi_1\sum_{j=2}^{M}|r_{j-1}^{(\ell)}|\,|r_j^{(\ell)}|\,(1-\gamma_{j-1})\,(1-\gamma_j)\Bigg] \\ &=\; SBP_{t,0}^{(o)} \;+\; sBP_{t,0}^{(o)^*,1} \;+\; sBP_{t,0}^{(o)^*,2} \;+\; sBP_{t,0}^{(\ell)^*,1} \;+\; sBP_{t,0}^{(\ell)^*,2} \;+\; SBP_{t,0}^{(\ell),3}\,, \end{aligned}$$

where $SBP_{t,0}^{(o)} + sBP_{t,0}^{(o)^*,1} + sBP_{t,0}^{(o)^*,2}$ *is the observable quadratic variation, which is approximatively equal to* $RV_t^{(o)}$ *for* $\zeta \in (0,1]$*.*

The proof can be found in the appendix. *Proposition 3* states that $BP_{t,0}$ underestimates the actual quadratic variation of the observable return process, whereas RV_t and $SBP_{t,0}$ does not. In fact we can show that the difference between $RV_t^{(o)}$ and $SBP_{t,0}^{(o)} + sBP_{t,0}^{(o)^*,1} + sBP_{t,0}^{(o)^*,2}$ converges for $M \to \infty$ to zero.

Proposition 4: *Given Proposition 3,*

$$RV_t^{(o)} + RV_t^{(\ell)} \;\;\xrightarrow{p}\;\; IV_t^{(o)} + IV_t^{(\ell)} \quad \textit{and} \quad RV_t^{(o)} \;\;\xrightarrow{p}\;\; IV_t^{(o)}\,,$$

for $M \to \infty$. $IV_t^{(o)}$ is the corresponding integrated variance for the observable return process, and $IV_t^{(\ell)}$ for the latent one. Besides, the difference between the sum of observable unbiased spot variance estimates of RV_t, and the sum of observable unbiased spot variance estimates of $SBP_{t,0}$ converges for $M \to \infty$ to zero:

$$\sum_{j=1}^{M} \sigma_j^{2\,(o)} - \sum_{j=2}^{M} \Big[\sigma_{j-1}{}^{(o)} \sigma_j{}^{(o)} + \tfrac{1}{2} \sigma_{j-1}{}^{(o)^*,1} \sigma_j{}^{(o),1} + \tfrac{1}{2} \sigma_{j-1}{}^{(o),2} \sigma_j{}^{(o)^*,2} \Big] \xrightarrow{p} 0 \, .$$

Proof. Given the specification of sustained bipower variation in equation (10), the fact that each increment is an unbiased estimate of the spot variance for a specific interval and the general convergence result of BNS (2004, p.10), the difference does converge to zero if the same quadratic variation fragments are approximated and the assumptions of *Proposition 3* hold. □

Analogously, we can extend *Proposition 3* and *4* for $i \geq 0$, $TBP_{t,i}$, $STBP_{t,i}$, $TriP_{t,i}$, $STriP_{t,i}$, $TTriP_{t,i}$, and $STTriP_{t,i}$. The derivable underestimation of integrated quarticity for the observable process by $TriP_{t,i}$ and $TTriP_{t,i}$ is even more severe. This is due to the fact that the probability of three consecutive observable return events is only ζ^3. The sustained estimators $STriP_{t,i}$ and $STTriP_{t,i}$ correct for this bias. Finally, we conclude that the same asymptotic distribution holds for the multipower variation based methods by separately looking at the observable and latent fragments. However, showing the same bias (correction) for *(S)ADS-Min* and *(S)ADS-Med* is not trivial and therefore left to future research. Additionally note, if a zero-return is caused by flat prices, we implicitly assume in *Proposition 3* and *4* that prices stay flat within the interval as well. In practice this might not always be true.

3.2 Simulation Study

In the following Monte Carlo experiments we are interested in analyzing the accuracy of the limit distribution of $Z_{t,\nu}$ and the correct detection rate of days with jumps and without jumps by employing *SIVQE*. For that, a Heston type price process (see Heston, 1993) with and without jumps is simulated:

$$\begin{aligned} \frac{dX(t)}{X(t)} &= \mu dt + \sqrt{v(t)} dW_X(t) + \kappa(t) dq(t) \, , \\ dv(t) &= (\varsigma - \varpi v(t)) dt + \eta \sqrt{v(t)} dW_v(t) \, , \end{aligned}$$

where μ is the drift, $W_{(\cdot)}(t)$ are standard Brownian motions, $corr(dW_X, dW_v) = \rho$ is the leverage correlation, $v(t)$ is a stochastic volatility factor, $\kappa(t)dq(t)$ is a compound Poisson process with a constant jump intensity λ_{jmp} and a random jump size distributed as $N(0, \sigma_{jmp}^2)$.[9] Generally, we simulate one setting without jumps, one with small and rare jumps ($\sigma_{jmp} = 0.0134$), and another with large and rare jumps ($\sigma_{jmp} = 0.1$). Moreover, we compute 5 (15) minute interval returns

[9]The parameters and the simulation horizon (30 years with 255 trading days per year and 7.5 trading hours per day) are chosen according to Schulz and Mosler (2010). Parameter settings: $\mu = 0.0304$, $\varsigma = 0.0064$, $\varpi = 0.012$, $\eta = 0.0711$, $\rho = -0.622$, $\sigma_{jmp} = \{0.0134, 0.1\}$ and $\lambda_{jmp} = 0.058$. Each second a price is simulated with the Euler scheme.

and process the series with the zero-return algorithm of Schulz and Mosler (2010) in order to proceed with the analysis of how robust *SIVQE* works in case of flat prices and no trading. The sampling frequency of 5 minutes is chosen as this is a very common sampling length in many empirical studies. Additionally, we compute the sampling length of 15 minutes due to the present empirical high-frequency dataset. For sensitivity interests, we vary the fraction of zero-returns in the return series from a very low level to a high level.

The analysis of the limit result of $Z_{t,\nu}$ in finite samples is graphed in figure 1 for 5 minute interval returns. In advance, it is worth repeating that in the no-jump case *(S)BNS* and *(S)CPR* are the same, as no trimming of the return series can be justified. Furthermore, in each panel of figure 1, the simulation results of the original methods serve as direct contrast. In the upper panels of figure 1, the method of *BNS/CPR* in combination with and without *SIVQE* is plotted for (from the left panel on) 10%, 20%, 40% and 60% zero-returns. Astonishing is to what extent the bias is reduced by implementing *SIVQE*. That is, even for a very high fraction of zero-returns, the limit results of $Z_{t,1/2}$ seem to be valid. Applying *SIVQE* to *ADS-Min* and *ADS-Med* (see middle and lower panels of figure 1) likewise yields a considerable reduction in bias, though for the highest fraction of zero-returns not as pertinent as for *SBNS/SCPR*.

In the next step, we analyze the rate of correctly detecting days with jumps ($\hat{=}$*(j)*) and days without jumps ($\hat{=}$*(nj)*) for the scenario large and rare jumps (# 442), and 5 minute interval returns (see table 1). To be more precise, we are interested in the question whether the jump and no-jump detection rate diverge with an increasing fraction of zero-returns from the ideal case, i.e. 0% zero-returns. For *SBNS*, *SCPR* and *SADS-Min* an increasing fraction of zero-returns seem to have no impact on both detection rates. That means, *SBNS*, *SCPR* and *SADS-Min* are extraordinary robust against zero-returns, even if they account more than 50%. Not as strong but still highly robust are the detection rates for *SADS-Med*, a result in line with figure 1. Note that for further comparison purposes, the detection rates without *SIVQE* are reported in brackets below respectively in table 1. We notice that with an increasing fraction of zero-returns the detection rates with and without *SIVQE* greatly diverge. And second, by computing the number of overall detected jump days,[10] we can derive that the number of detected jump days rises with an increasing fraction of zero-returns without applying *SIVQE*, i.e. the number of spurious jump days increases with a rising occurrence of flat prices and no trading. For the scenario small and rare jumps (# 435), the conclusions remain qualitatively similar to the scenario with large and rare jumps (see table 2).

To analyze the overall performance of each method with respect to detecting days with and without price jumps across different levels of zero-returns, we graph the nonparametric sensitivity index $\mathcal{A}$ for 5 minute interval returns. The outcome for scenario large (small) and rare jumps can be found in the upper (lower) panel of figure 2. For large and rare jumps (upper panel in figure 2), we observe that for 1% fraction of zero-returns very similar results of $\mathcal{A}$ across each method with and without *SIVQE* are obtained. Not observable in the panel is that even for this small fraction of zero-returns the new approach performs already slightly better in all

[10] Here: ('#simulated-jumps' × '(j)' + ('#trad.-days' − '#simulated-jumps')× 'β-error').

cases than without *SIVQE*. By focusing on each method without *SIVQE*, we can state that for an increasing fraction of zero-returns, $\mathcal{A}$ rapidly decreases across the respective methods. The negative slope of the $\mathcal{A}$ curve is steep for *BNS* and *CPR* and less steep for both *ADS-Min* and *ADS-Med*. Turning now to the results for *SIVQE*, it can be seen that up to 15% fraction of zero-returns $\mathcal{A}$ stays on almost the same level across each method. For an even larger fraction of zero-returns, $\mathcal{A}$ for *SBNS*, *SCPR* and *SADS-Min* remains on almost the same level, whereas for *SADS-Med* it gets sooner or later only slightly worse. Additionally, we observe that each method with *SIVQE* performs better than the approach proposed by Schulz and Mosler (2010) ($\hat{=}SM$). [11]

The direction of the overall performance results do not change for the scenario small and rare jumps (see lower panel in figure 2). Not explicitly reported are the results for 15 minute sampling intervals as they give no additional input to the qualitative conclusions already drawn by 5 minute sampling intervals.[12]

In summary we can say that *SIVQE* shows valuable properties in this simulation study. *SIVQE* in combination with the methods *BNS*, *CPR*, *ADS-Min* and *ADS-Med* is quite robust against zero-returns. It keeps good size and detection rates with respect to the case where no zero-returns are present.

4 Empirical Analysis: Electricity Forward Contracts

4.1 Data

The high-frequency dataset we are working with is the same as employed by Schulz and Mosler (2010), covering the period from May 1st 2002 to June 30th 2008. It consists of initially season forward contracts and later on quarter forward contracts, differing in length of the delivery period. In 2004, quarter forwards were introduced as a replacement of the season forwards. To receive a long time series both contracts are treated the same. A further institutional detail is that contracts are traded on weekdays from 8:00am to 3:30pm. Each traded contract has a finite life cycle. To create one time series, we merged periods of contracts shortest to maturity up to seven days before settlement, as the heaviest trading activity regarding number of trades per day are observed within this life cycle of a forward contract. Not included in the time series are inactive trading days, like overnights, weekends, holidays and several trading days with extremely low trading activity. This proceeding results in a time series of high-frequency transaction prices over 1390 active trading days. The computation of realized variance as in equation (3) requires to compute sufficiently small interval returns over equidistant time grids. For each time grid, we assign a price with the previous tick method by Hansen and Lunde

[11] The approach by Schulz and Mosler (2010) is based on *BNS*. It aims to reduce the distorting impact of zero-returns by implicitly maximizing the number of increments in $BP_{t,i}$ and $TriP_{t,i}$ unequal to zero with an optimal choice of i. Thereafter, the general proceeding of computing the test statistic, jump factor and integrated variance applies. Details on the optimal choice of i are: (a) fix the number of intraday sampling intervals M (effective for the full-sample); (b) $\max_{\{i \in I\}} \frac{TriP_{t,i}}{(BP_{t,i})^2}$, where $I = \left\{1, 2, ..., \left\lfloor \frac{M}{2} \right\rfloor\right\}$.

[12] The simulation results can be obtained upon request from the author.

(2003, 2006) and then compute continuously compounded interval returns. To be in line with the proposed robust approach and not to receive biased interval returns, we only apply the previous tick method if there is actually observed price data within an interval before a grid point. Important for the computation of the return series: if at least one grid is without an assigned price, the corresponding interval return is set to zero. According to Schulz and Mosler (2010), we set the sampling interval length to 15 minutes, producing 30 interval returns per day. For 15 minute sampling intervals, the time series consist of 48% zero-returns, whereof 13% (35%) are due to flat prices (no trading).

4.2 Detecting Price Jumps with Original Methods

Initial empirical results for each original jump detection method is reported in table 3a, where we can find the proportion of detected jump days across different levels of significance (α). As Schulz and Mosler (2010) already show, *BNS* yields for small α an overproportionally large amount of jump days. Not surprising is the fact that *CPR* surpasses this amount as *CPR* corrects for one drawback of bipower variation for finite M, closely discussed in section 2.2, but still suffers from zero-returns. The smallest proportion of detected jump days is produced by the methods *ADS-Min* and *ADS-Med*. In order to finer compare the empirical results across methods, we question whether the 5% most potential jump factors ($\tilde{J}_{t,\nu}$) of each method correspond in occurrence time.[13] For this, we formulate the parity measure EJ_i:

$$EJ_i = 2\frac{\sum_{t=1}^{T} JG_t}{\sum_{t=1}^{T}\left(\mathbb{1}_{\{\tilde{J}_{t,\nu}>0\}} + \mathbb{1}_{\{\tilde{J}_{t,k}>0\}}\right)}, \quad JG_t = \begin{cases} 1 & \text{if } (\tilde{J}_{t,\nu} > 0) \wedge (\tilde{J}_{t,k} > 0) \\ 0 & \text{else} \end{cases}, \tag{15}$$

where $\nu \neq k,\ \nu, k = \{1, 2, ..., 4\}$ and $i = \{1, 2, 3\}$. The upper left part of table 4 reports the output of EJ_i. Obviously, the occurrence time of the potential jump factors greatly diverges across methods. Highlighting is also the large parity of *BNS* and *CPR*. This is due to the fact that they are from a methodic point of view most congruent. Further details about the respective size of the potential jump factors as well as trading activity[14] on these potential jump days can be found in table 5a. The largest average proportionate contribution of the potential jump factors to realized variance yields *CPR*, followed by *BNS*, *ADS-Min* and *ADS-Med*. An almost reversed order applies to trading activity. *ADS-Min* and *ADS-Med* show that the potential jump days are on average characterized by slightly greater trading activities in comparison to the full-sample (compare to the last row in table 5).

The considerable discrepancy across methods is mainly due to their individual sensitivity to zero-returns, i.e. flat prices and no trading. From the previous simulation results, and Schulz and Mosler (2010), we know that zero-returns have an impact on detecting price jumps in realized variance, i.e. intraday zero-returns positively distort the test statistic $Z_{t,\nu}$. It arises

[13] Procedure to compute for each method the series with the 5% most potential jump factors ($\tilde{J}_{t,\nu}$): (a) compute the test statistic across all active trading days; (b) compute the corresponding jump factor for the largest 5% of all test statistic values; (c) set on the remaining days a jump factor of zero.

[14] In this study, trading activity is determined via number of trades per day, number of price changes per day, number of traded contracts and intertrade duration.

from the Monte Carlo experiments that an ascending fraction of zero-returns amplifies the degree of the bias. Therefore, it is intuitive that trading activity is a supportive indicator for specifying, whether a jump is spurious or true. In light of the just mentioned, let us anew analyze the trading activity on the potential jump days in table 5a. It appears that there are indications for more spurious jumps detected by *BNS* and *CPR* than by *ADS-Min* and *ADS-Med*, as *BNS* and *CPR* mostly show that the potential jump days are characterized by remarkably low trading activity.

4.3 Detecting Price Jumps with Robust Approach

This section discusses the empirical results of the each jump detection method in combination with *SIVQE*. We begin with the proportion of jump days given a certain level of significance, reported in table 3b. Coherent with the output of the Monte Carlo experiment, *SBNS*, *SCPR*, *SADS-Min* and *SADS-Med* yield much less jump days as in the original setting, and, in line with the theory, *SCPR* slightly more than *SBNS*. Furthermore consistent with the derivable simulation results as of table 1 and 2, *SADS-Med* produces the largest fraction of jump days and *SADS-Min* the lowest.

Not only do we receive overall less proportions of jump days given a certain α, but also a larger parity in occurrence time of the 5% most potential jump factors, reported in the lower right part of table 4. The latter statement is graphically illustrated in figure 3. In the bottom panel, the cumulative sum of $\tilde{J}_{t,\nu}^{1/2}$ is graphed over time for *SIVQE*, which can be directly compared to the result of the original methods (top panel). We may say that by applying *SIVQE*, we overall receive more consistent conclusions across methods.

A further interesting question is to what extent the characteristic of these potential jump days change from the original setting. In table 5b, the size of the potential jump factors and trading activity on these potential jump days are specified. Starting with the size: despite the decline of the average proportionate contribution of the potential jump factors to realized variance, the average total size of these potential jump factors increases for *SBNS* and *SCPR*. For *SADS-Min* and *SADS-Med* it roughly remains on the same level. This is due to the fact that the potential jump factors are now on days with on average higher trading activity and realized variance. As mentioned above, this seems more plausible as in the original setting. Applying trading activity as a qualitative variable, all methods with *SIVQE* have potential jump factors on days with above average trading activity. Likewise interesting in table 5 is the parity in occurrence time of the potential jump factors with the upper 5% right tail of the empirical distribution of $\max_t\{|r_j|\}$, denoted by *rtp*. As expected, a considerable intersection can be observed for *SIVQE* and a low one working with the original methods. The visualization of this result can be found in figure 3 by comparing the top and bottom panel with the middle panel, which graphs $\max_t\{|r_j|\}$ over time.

Yet to be clarified is the question which method should be preferred in this specific case. Two arguments speak for working with *SCPR*. First, the simulation results of *SCPR* show slightly better overall performance ($\mathcal{A}$) for larger fractions of zero-returns. Second, the qualitative indicators are strongest. Therefore, it seems preferable to implement *SCPR*.

5 Conclusion

This paper investigates a selection of comparable methods disentangling contributions from price jumps to realized variance. Employed methods are by BNS (2004a, 2006a), CPR (2009) and ADS (2009). Zero-returns, caused by flat prices and no trading, are a pivotal source of distortion in each method. Therefore, we introduce a new approach to robustify each method to zero-returns. It is called sustained integrated variance and quarticity estimation *(SIVQE)*.

Under ideal conditions of no zero-returns, we show that the asymptotic distribution of each method with *SIVQE* remains the same with respect to its original counterpart. Besides, in describing the return process with a Bernoulli process we show that the multipower variation based integrated variance and quarticity estimators by BNS (2004a, 2006a) and CPR (2009) underestimate the actual quadratic variation of the observable return variation. However, implementing the methods by BNS (2004a, 2006a) or CPR (2009) with *SIVQE* does not yield an underestimation. Furthermore, we show that in case of no price jumps the difference between realized variance and the robustified multipower variation based integrated variance estimator of the observable variation converges in probability to zero for a decreasing sampling length.

SIVQE is tested in a Monte Carlo experiment under imperfect market conditions, reflecting different levels of flat price and no trading bias. The convergence criteria under the null hypothesis of each test statistic are quite robust against an increasing fraction of zero-returns. The investigation of the accuracy of (no-) jump day detection rates shows that the detection rates are considerably robust against an increasing fraction of zero-returns. To analyze the overall performance of detecting days with or without jumps, we employ a nonparametric sensitivity index ($\mathcal{A}$) typically used in signal detection theory. The simulation results for $\mathcal{A}$ yield that *SIVQE* in combination with the corresponding method definitely performs better across all zero-return levels than implementing the original methods. Besides, it even constitutes a better performer than the approach proposed by Schulz and Mosler (2010).

In an empirical analysis, using high-frequency data of electricity forward contracts traded on the Nord Pool Energy Exchange, measured contributions from price jumps to realized variance are compared with regard to size and occurrence time. The empirical study shows considerable differences across the original methods, foremost amongst the 5% most potential jump factors. In addition, these potential jump days are typically characterized with below average trading activity and a small amount of extreme price movements, indicating a large fraction of spurious price jumps. With the same proceeding, we analyze the empirical results for *SIVQE* and find more plausible conclusions. For the present high-frequency dataset there are indications for preferring *SIVQE* with the method of CPR (2009) as the qualitative indicators are strongest.

The introduction of *SIVQE* raises further interesting research questions. Of interest would be to elaborate on alternative estimators for the sustainer, utilized either in the integrated variance or quarticity estimators. Moreover, it would be interesting to extend the idea of *SIVQE* for a multivariate setting, i.e. for existing realized covariance or covariation estimators.

References

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

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

[3] Andersen, T. G., Bollerslev, T., and Diebold, F. X. (2002) Parametric and nonparametric volatility measurement. *Handbook of Financial Econometrics*, Y. Aït-Sahalia and L.P. Hansen.

[4] Andersen, T. G., Bollerslev, T., and Huang, X. (2007) A semiparametric framework for modeling and forecasting jumps and volatility in speculative prices. Working paper, Kellogg School of Management, Northwestern University.

[5] Andersen, T. G., Dobrev, D., and Schaumburg, E. (2009) Jump-robust volatility estimation using nearest neighbor truncation. Working paper 15533, National Bureau of Economic Research, Cambridge.

[6] Andersen, T. G., Dobrev, D., and Schaumburg, E. (2008) Duration-based volatility estimation. Working paper, Kellogg School of Management, Northwestern University.

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

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

[9] Barndorff-Nielsen, O. E. and Shephard, N. (2006) Multipower variation and stochastic volatility. *Stochastic Finance*, **Part I**, 73–82.

[10] Barndorff-Nielsen, O. E. and Shephard, N. (2005) How accurate is the asymptotic approximation to the distribution of realized volatility. *Identification and Inference for Econometric Models. Essays in Honor of Thomas Rothenberg.*, Cambridge: Cambridge University Press.

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

[12] Barndorff-Nielsen, O. E. and Shephard, N. (2004) Measuring the impact of jumps on multivariate price processes using multipower variation. Discussion paper, Nuffield College, Oxford University.

[13] Bollerslev, T., Kretschmer, U., Pigorsch, C., and Tauchen, G. (2009) A discrete-time model for daily s&p 500 returns and realized variations: Jumps and leverage effects. *Journal of Econometrics*, **150**, 151–166.

[14] Christensen, K., Oomen, R. C., and Podolskij, M. (2009) Realised quantile-based estimation of the integrated variance. Creates research papers, School of Economics and Management, University of Aarhus.

[15] Corsi, F., Pirino, D., and Renò, R. (2009) Threshold bipower variation and the impact of jumps on volatility forecasting. Working paper, Università di Siena.

[16] Eraker, B. (2004) Do stock prices and volatility jump? reconciling evidence from spot and option prices. *Journal of Finance*, **59**, 1367–1404.

[17] Eraker, B., Johannes, M. S., and Polson, N. (2003) The impact of jumps in volatility and returns. *Journal of Finance*, **58**, 1269–1300.

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

[19] Hansen, P. R. and Lunde, A. (2003) An optimal and unbiased measure of realized variance based on intermittent high-frequency data. Unpublished paper, Department of Economics, Stanford University.

[20] Heston, S. L. (1993) A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, **6**, 327–343.

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

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

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

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

[25] Mancini, C. and Renò, R. (2008) Threshold estimation of jump-diffusion models and interest rate modeling. Working paper, University of Florence, Dipartimento di Matematica, and University of Siena, Department of Economics.

[26] Phillips, P. C. B. and Yu, J. (2008) Information loss in volatility measurement with flat price trading. Working paper, Cowles Foundation for Research in Economics Yale University.

[27] Schulz, F. C. and Mosler, K. (2010) The effect of infrequent trading on detecting price jumps. Working paper, University of Cologne, Germany.

[28] Zhang, J. and Mueller, S. T. (2005) A note on roc analysis and non-parametric estimate of sensitivity. *Psychometrika*, **70**, 203–212.

Appendix

A1. Sustained Integrated Quarticity Estimators

Formalization of **sustained tripower quarticity** ($STriP_{t,i}$, $\nu = 1$) and **sustained threshold tripower quarticity** ($STTriP_{t,i}$, $\nu = 2$), summarized in $S\widehat{IQ}_{t,\nu}$ ($\nu = 1, 2$):

$$
\begin{aligned}
S\widehat{IQ}_{t,\nu} \;&=\; \Delta_i \sum_{j=1+2(1+i)}^{M} \Big[\mathbb{1}_{b1} \left(\breve{r}_{j-2(1+i),\nu}\, \breve{r}_{j-(1+i),\nu}\, \breve{r}_{j,\nu} \right) \\
&+\; \mathbb{1}_{b2} \left(\omega_{j-2(1+i),\nu}\, \omega_{j-(1+i),\nu}\, \breve{r}_{j,\nu} \right) \tfrac{1}{3} \;+\; \mathbb{1}_{b3} \left(\omega_{j-2(1+i),\nu}\, \breve{r}_{j-(1+i),\nu}\, \omega_{j,\nu} \right) \tfrac{1}{3} \\
&+\; \mathbb{1}_{b4} \left(\breve{r}_{j-2(1+i),\nu}\, \omega_{j-(1+i),\nu}\, \omega_{j,\nu} \right) \tfrac{1}{3} \;+\; \mathbb{1}_{b5} \left(\breve{r}_{j-2(1+i),\nu}\, \breve{r}_{j-(1+i),\nu}\, \omega_{j,\nu} \right) \tfrac{2}{3} \\
&+\; \mathbb{1}_{b6} \left(\breve{r}_{j-2(1+i),\nu}\, \omega_{j-(1+i),\nu}\, \breve{r}_{j,\nu} \right) \tfrac{2}{3} \;+\; \mathbb{1}_{b7} \left(\omega_{j-2(1+i),\nu}\, \breve{r}_{j-(1+i),\nu}\, \breve{r}_{j,\nu} \right) \tfrac{2}{3} \Big] \,.
\end{aligned}
$$

For $\nu = 1$, $\breve{r}_{j,1} = |r_j|^{4/3}$, and for $\nu = 2$, $\breve{r}_{j,2} = \Psi_{4/3}(r_j, \theta)$. The indicator functions $\mathbb{1}_{b1}, \ldots, \mathbb{1}_{b7}$ have the same definition as in equation (12). In light of the concept of (threshold) tripower quarticity, we suggest likewise a reasonable proxy for $\omega_{j,\nu}$, i.e. $\omega_{j,\nu} = E\left(\breve{r}_{j,\nu}\right)$. In the absence of price jumps, $\omega_{j,\nu}$ can be estimated on day t by:

$$
\hat{\omega}_{j,\nu} = \hat{\sigma}_{\nu}^{4/3}\, 2^{2/3}\, \frac{\Gamma(7/6)}{\Gamma(1/2)}\,, \quad \text{with} \quad \hat{\sigma}_{\nu} = \sqrt{\frac{\pi}{2}}\, \frac{1}{N} \sum_{j=1}^{M} |r_j|\,, \quad \text{where} \quad N = \sum_{j=1}^{M} \mathbb{1}_{\{|r_j|>0\}}\,.
$$

In the presence of price jumps,

$$
\hat{\omega}_{j,\nu} = \hat{\sigma}_{\nu}^{4/3}\, 2^{2/3}\, \frac{\Gamma(7/6)}{\Gamma(1/2)}\,, \quad \text{with} \quad \hat{\sigma}_{\nu} = \sqrt{\Theta_{\tau}^{\delta}}\,,
$$

where Θ_{τ}^{δ} is specified in equation (8).

Formalization of **sustained MinRQ** ($SMinRQ_t$):

$$
\begin{aligned}
SMinRQ_t \;&=\; M \frac{\pi}{3\pi - 8} \left(\frac{M}{M-1} \right) \sum_{j=2}^{M} \Big[\mathbb{1}_{a1} \min\left(|r_{j-1}|\,, |r_j|\right)^4 \\
&+\; \mathbb{1}_{a2} \min\left(|r_{j-1}|\,, \wp_{j,3}\right)^4 \tfrac{1}{2} + \mathbb{1}_{a2} \min\left(\wp_{j-1,3}\,, |r_j|\right)^4 \tfrac{1}{2} \Big]\,,
\end{aligned}
$$

where the indicator functions $\mathbb{1}_{a1}, \mathbb{1}_{a2}$ and $\mathbb{1}_{a3}$ are defined as in equation (11) with $i = 0$. $\wp_{j,3}$ is estimated as of equation (13) in the absence of price jumps, and as of equation (14) in the presence of price jumps.

Formalization of **sustained MedRQ** ($SMedRQ_t$):

$$
\begin{aligned}
SMedRQ_t \;&=\; M \frac{3\pi}{9\pi + 72 - 52\sqrt{3}} \left(\frac{M}{M-2} \right) \sum_{j=3}^{M} \Big[\mathbb{1}_{b1} \operatorname{med}\left(|r_{j-2}|\,, |r_{j-1}|\,, |r_j|\right)^4 \\
&+\; \mathbb{1}_{b2} \operatorname{med}\left(\wp_{j-2,4}\,, \wp_{j-1,4}\,, |r_j|\right)^4 \tfrac{1}{3} \;+\; \mathbb{1}_{b3} \operatorname{med}\left(\wp_{j-2,4}\,, |r_{j-1}|\,, \wp_{j,4}\right)^4 \tfrac{1}{3} \\
&+\; \mathbb{1}_{b4} \operatorname{med}\left(|r_{j-2}|\,, \wp_{j-1,4}\,, \wp_{j,4}\right)^4 \tfrac{1}{3} \;+\; \mathbb{1}_{b5} \operatorname{med}\left(|r_{j-2}|\,, |r_{j-1}|\,, \wp_{j,4}\right)^4 \tfrac{2}{3} \\
&+\; \mathbb{1}_{b6} \operatorname{med}\left(|r_{j-2}|\,, \wp_{j-1,4}\,, |r_j|\right)^4 \tfrac{2}{3} \;+\; \mathbb{1}_{b7} \operatorname{med}\left(\wp_{j-2,4}\,, |r_{j-1}|\,, |r_j|\right)^4 \tfrac{2}{3} \Big]\,,
\end{aligned}
$$

where the indicator functions $\mathbb{1}_{b1}, \ldots, \mathbb{1}_{b7}$ are defined as in equation (12) with $i = 0$. $\wp_{j,4}$ can be estimated as of equation (13) in the absence of price jumps, and as of equation (14) in the presence of price jumps.

A2. Proof of Proposition 3

Assume a price process as of equation (1) without price jumps. Furthermore, assume that the return process consists of observable (o) and latent (ℓ) fragments. The two states are described by a Bernoulli process,

$$r_j = \begin{cases} r_j^{(o)} & \text{if} \quad \gamma_j = 1 \,, \\ r_j^{(\ell)} & \text{if} \quad \gamma_j = 0 \,, \end{cases} \tag{16}$$

where γ_j is a Bernoulli sequence independent of the return process, with $\mathrm{E}(\gamma_j = 0) = 1 - \zeta$, $\mathrm{E}(\gamma_j = 1) = \zeta$, and $\zeta \in (0, 1]$. $r_j^{(o)}$ is an observable return, whereas $r_j^{(\ell)}$ is latent. Time change is random and discontinuous. The specification of r_j implies that

$$r_j = r_j^{(o)} \gamma_j + r_j^{(\ell)} (1 - \gamma_j) \qquad \text{or} \qquad r_j^2 = r_j^{2(o)} \gamma_j + r_j^{2(\ell)} (1 - \gamma_j) \,. \tag{17}$$

In light of Theorem 2.1 by Phillips and Yu (2008) the model for r_j in equation (16) preserves the martingale property. According to the specification of the return process in equation (17), realized variance is for $[t\text{-}1, t]$:

$$RV_t = \sum_{j=1}^{M} r_j^2 = \sum_{j=1}^{M} \Big[r_j^{2(o)} \gamma_j + r_j^{2(\ell)} (1 - \gamma_j) \Big] = \sum_{j=1}^{M} r_j^{2(o)} \gamma_j + \sum_{j=1}^{M} r_j^{2(\ell)} (1 - \gamma_j) = RV_t^{(o)} + RV_t^{(\ell)} \,.$$

$r_j^{(o)}$ occurs with probability ζ, and so does the corresponding spot variance estimate $r_j^{2(o)}$. Therefore, the observable quadratic variation $RV_t^{(o)}$ is figuratively linked to the probability ζ. Correspondingly, the latent quadratic variation $RV_t^{(\ell)}$ is linked to the probability $(1 - \zeta)$. In other words, the variation produced by $r_j^{(o)}$ is completely captured by $RV_t^{(o)}$. The estimation of the spot variances with $r_j^{2(o)}$ is not influenced by the previous or following event. This makes the spot variance estimates produced by realized variance robust to the influence of previous or following events.

Proceeding likewise with incorporating the return process as of equation (17) in bipower variation yields,

$$\begin{aligned} BP_{t,0} &= \frac{M}{M-1} \varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(o)} \gamma_{j-1} + r_{j-1}^{(\ell)} (1 - \gamma_{j-1})| \, |r_j^{(o)} \gamma_j + r_j^{(\ell)} (1 - \gamma_j)| \\ &= \frac{M}{M-1} \varphi_1 \sum_{j=2}^{M} \Big[|r_{j-1}^{(o)}| \gamma_{j-1} + |r_{j-1}^{(\ell)}| (1 - \gamma_{j-1}) \Big] \Big[|r_j^{(o)}| \gamma_j + |r_j^{(\ell)}| (1 - \gamma_j) \Big] \\ &= \frac{M}{M-1} \Bigg[\varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(o)}| |r_j^{(o)}| \gamma_{j-1} \gamma_j + \varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(\ell)}| |r_j^{(o)}| (1 - \gamma_{j-1}) \gamma_j \\ &\qquad + \varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(o)}| |r_j^{(\ell)}| \gamma_{j-1} (1 - \gamma_j) + \varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(\ell)}| |r_j^{(\ell)}| (1 - \gamma_{j-1}) (1 - \gamma_j) \Bigg] \\ &= BP_{t,0}^{(o)} + BP_{t,0}^{(\ell)^*,1} + BP_{t,0}^{(\ell)^*,2} + BP_{t,0}^{(\ell),3} \,, \end{aligned}$$

where $\mathrm{E}(\gamma_{j-1} = 1, \gamma_j = 1) = \zeta^2$, and $\mathrm{E}(\gamma_{j-1} = 0, \gamma_j = 1) = \mathrm{E}(\gamma_{j-1} = 1, \gamma_j = 0) = \zeta(1 - \zeta)$. $BP_{t,0}^{(o)}$ is the observable quadratic variation. Each increment in $BP_{t,0}^{(o)}$ represents an unbiased estimate of the spot variance. This observable spot variance estimator requires that two consecutive events yield the state ‘(o)’, which happens only with probability ζ^2. That means, the occurrence of one event ‘(o)’ does not guarantee a spot variance estimate contributing to $BP_{t,0}^{(o)}$. All other consecutive events like ‘(ℓ)-(o)’, ‘(o)-(ℓ)’ and ‘(ℓ)-(ℓ)’ represent the latent spot variance estimates of either partly observable and latent, or solely latent return variations, and have a joint probability of $1-\zeta^2$. The latent part of $BP_{t,0}$ amounts to $BP_{t,0}^{(\ell)^*,1} + BP_{t,0}^{(\ell)^*,2} + BP_{t,0}^{(\ell),3}$ and can be linked to the probability $1-\zeta^2$. Therefore, if $0 < \zeta < 1$, $RV_t^{(o)} > BP_{t,0}^{(o)}$. If and only if $\zeta = 1$, $BP_{t,0}^{(o)}$ is in limit equal to $RV_t^{(o)}$.

Now, we will define $SBP_{t,0}$ with the return process of equation (17). Before, let us rewrite $BP_{t,0}$ with respect to the decomposition of equation (9):

$$
\begin{aligned}
BP_{t,0} \;=\; \frac{M}{M-1}\,\varphi_1 \sum_{j=2}^{M}\Big[& |r_{j-1}^{(o)}|\,|r_j^{(o)}|\,\gamma_{j-1}\,\gamma_j + \tfrac{1}{2}\,|r_{j-1}^{(\ell)}||r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j \\
& + \tfrac{1}{2}\,|r_{j-1}^{(\ell)}|\,|r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j + \tfrac{1}{2}\,|r_{j-1}^{(o)}||r_j^{(\ell)}|\,\gamma_{j-1}\,(1-\gamma_j) \\
& + \tfrac{1}{2}\,|r_{j-1}^{(o)}||r_j^{(\ell)}|\,\gamma_{j-1}\,(1-\gamma_j) + |r_{j-1}^{(\ell)}||r_j^{(\ell)}|\,(1-\gamma_{j-1})\,(1-\gamma_j)\Big] .
\end{aligned}
$$

Now we can define $SBP_{t,0}$, incorporating the specifications of equation (10):

$$
\begin{aligned}
SBP_{t,0} \;=\; \frac{M}{M-1}\Bigg[& \varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(o)}|\,|r_j^{(o)}|\,\gamma_{j-1}\,\gamma_j + \varphi_1 \sum_{j=2}^{M} \tfrac{1}{2}\,\wp_{j-1}\,|r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j \\
& + \varphi_1 \sum_{j=2}^{M} \tfrac{1}{2}\,|r_{j-1}^{(o)}|\,\wp_j\,\gamma_{j-1}\,(1-\gamma_j) + \varphi_1 \sum_{j=2}^{M} \tfrac{1}{2}\,|r_{j-1}^{(\ell)}|\,|r_j^{(o)}|\,(1-\gamma_{j-1})\,\gamma_j \\
& + \varphi_1 \sum_{j=2}^{M} \tfrac{1}{2}\,|r_{j-1}^{(o)}|\,|r_j^{(\ell)}|\,\gamma_{j-1}\,(1-\gamma_j) + \varphi_1 \sum_{j=2}^{M} |r_{j-1}^{(\ell)}|\,|r_j^{(\ell)}|(1-\gamma_{j-1})\,(1-\gamma_j)\Bigg] \\
=\; & SBP_{t,0}^{(o)} + sBP_{t,0}^{(o)*,1} + sBP_{t,0}^{(o)*,2} + sBP_{t,0}^{(\ell)^*,1} + sBP_{t,0}^{(\ell)^*,2} + SBP_{t,0}^{(\ell),3} ,
\end{aligned}
$$

where $SBP_{t,0}^{(o)}$ is the observable and $sBP_{t,0}^{(o)*,1} + sBP_{t,0}^{(o)*,2}$ the sustained observable quadratic variation times 0.5. The probability of the event ‘(o)-(o)’ is again ζ^2, which accounts to $SBP_{t,0}^{(o)}$. The events ‘(o)-(ℓ)’ and ‘(ℓ)-(o)’ occur with probability $\zeta - \zeta^2$, respectively. The resulting probability of a sustained observable event is $\zeta - \zeta^2$ contributing either to $sBP_{t,0}^{(o)*,1}$ or $sBP_{t,0}^{(o)*,2}$. Therefore, the joint probability of ‘(o)-(o)’, and ‘(o)-(ℓ)’ or ‘(ℓ)-(o)’ amounts to $\zeta^2 + (\zeta - \zeta^2) = \zeta$. With *SIVQE* we ensure that $SBP_{t,0}^{(o)} + sBP_{t,0}^{(o)*,1} + sBP_{t,0}^{(o)*,2}$ approximates the same amount as $RV_t^{(o)}$. The latent part of $SBP_{t,0}$ amounts to $sBP_{t,0}^{(\ell)^*,1} + sBP_{t,0}^{(\ell)^*,2} + SBP_{t,0}^{(\ell),3}$, which now can be linked to the probability $1-\zeta$.

□

Figures

Figure 1: QQ plots of $Z_{t,\nu}$ statistic for 5 minute sampling intervals, with(out) *SIVQE*

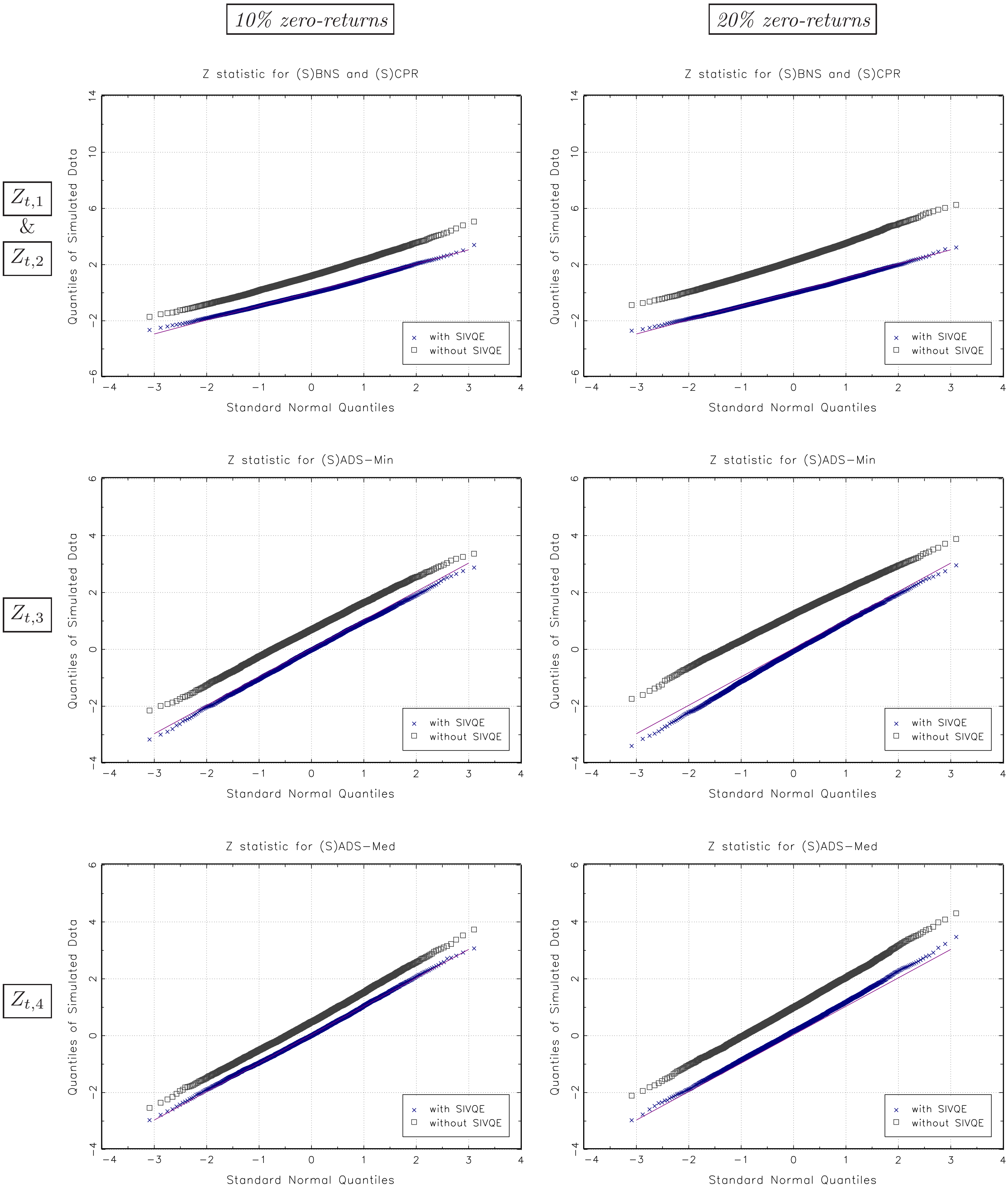
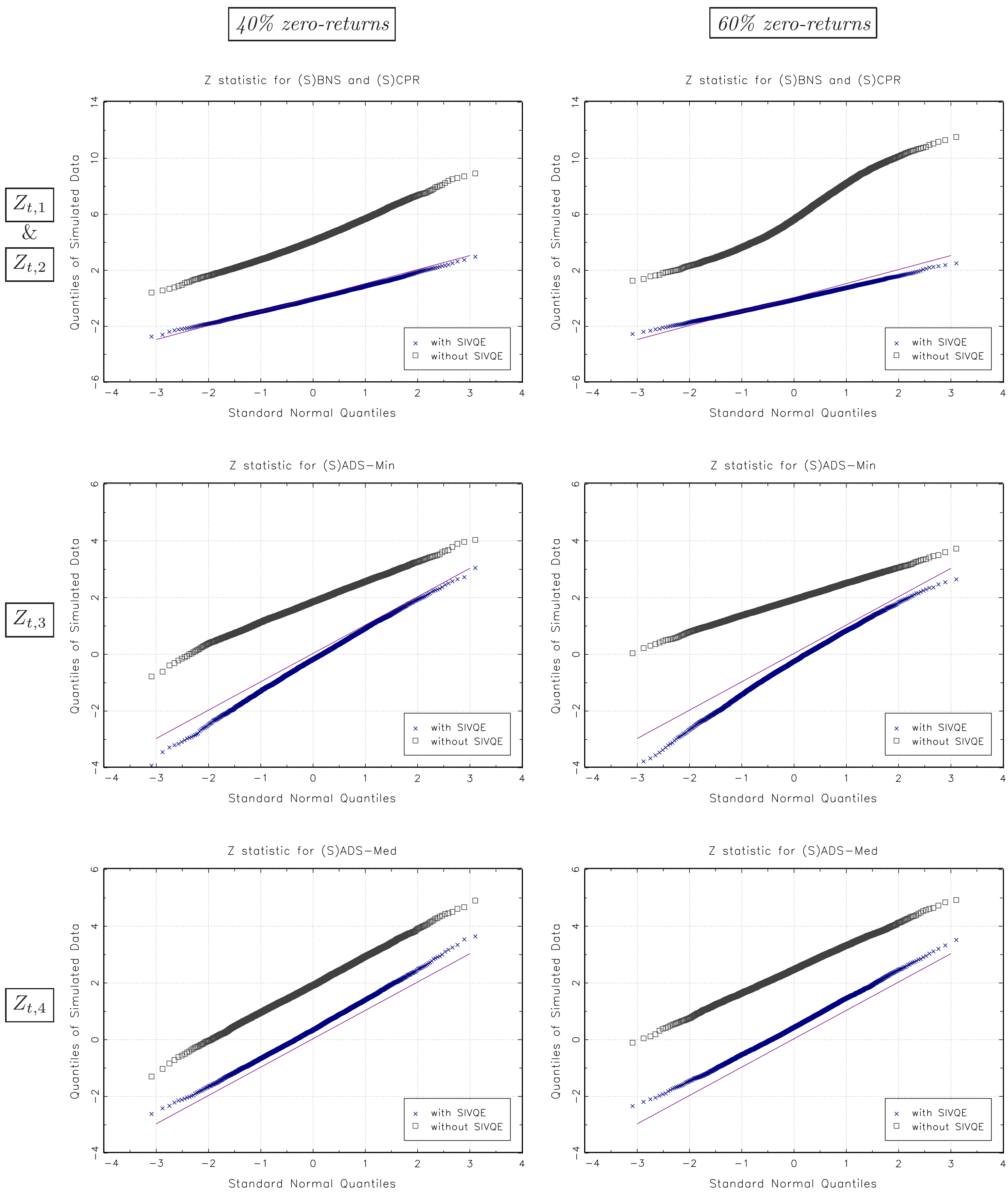
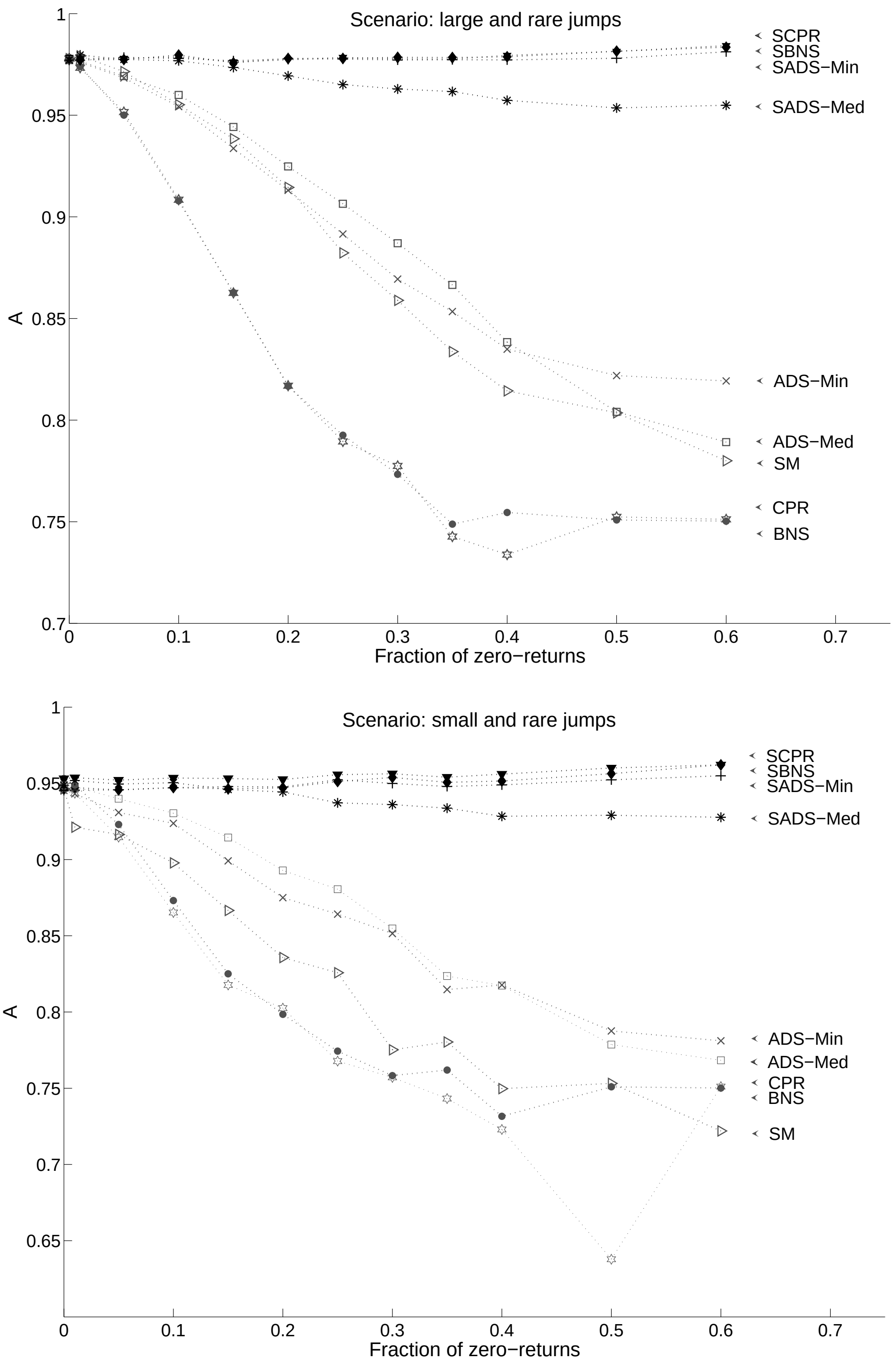


Figure 1: continued



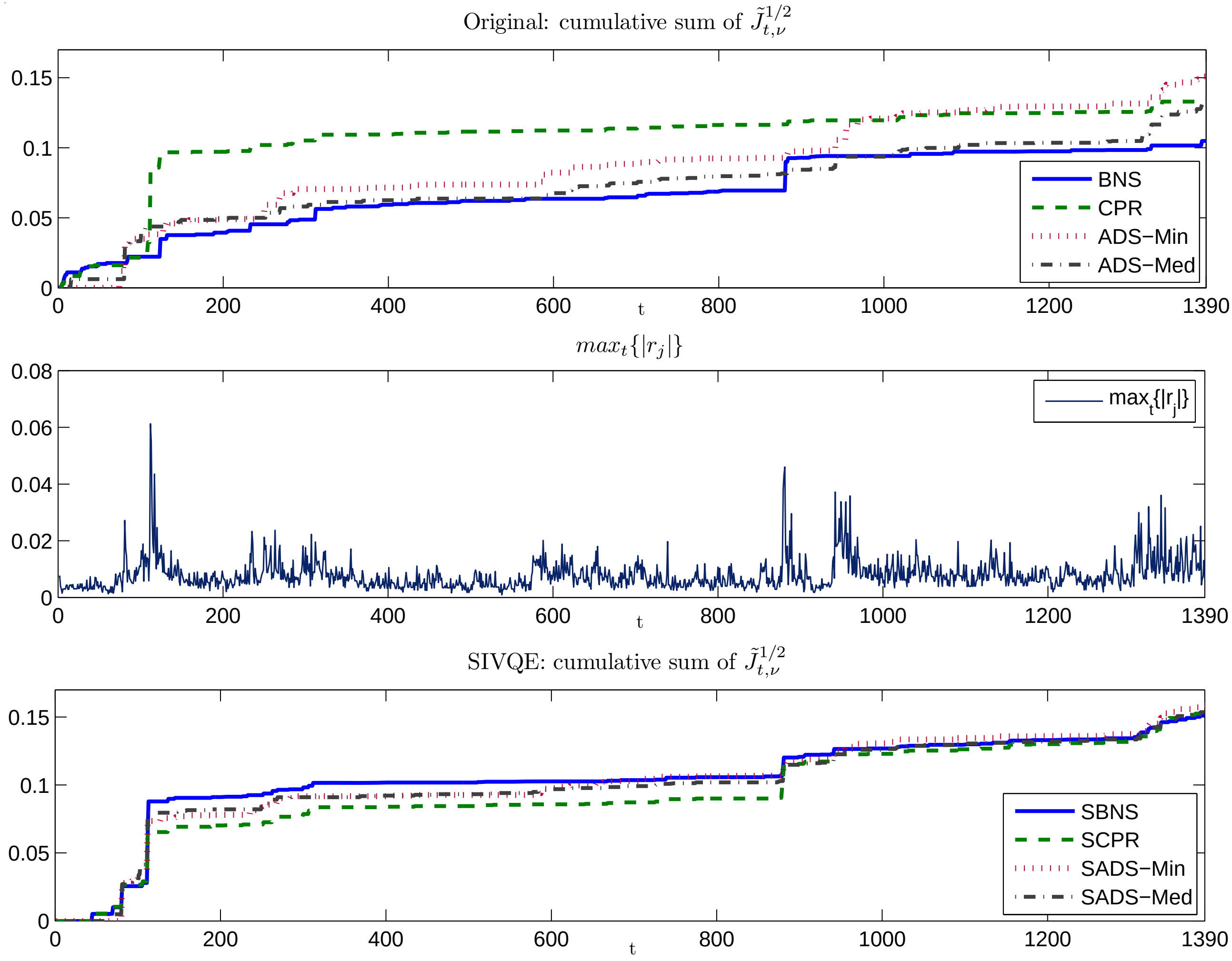
Remarks: Simulated realization of the Heston type price process without jumps for 7650 days with parameter specifications and algorithm for zero-returns as of section 3.2. Daily $Z_{t,\nu}$ statistic with(out) *SIVQE* and 5 minute sampling intervals. In the correspondingly labeled panels we simulate 10%, 20%, 40% and 60% zero-returns. In the upper, middle and lower panels, daily $Z_{t,1/2}$, $Z_{t,3}$ and $Z_{t,4}$ statistic with(out) *SIVQE* is depicted. The ordinate labels the quantiles of the simulated input sample, the abscissa the standard normal quantiles. The solid bisecting line graphs the theoretical result. The crosses (squares) represent the results with(out) *SIVQE*.

Figure 2: Nonparametric sensitivity index $\mathcal{A}$ for 5 minute sampling intervals, large/small and rare jumps, and various fraction of zero-returns



<u>Remarks:</u> Simulated realization of the Heston type price process with jumps (#: 442, 435) for 7650 days with parameter specifications and algorithm for zero-returns as of section 3.2. The curves represent $\mathcal{A}$ for either *(S)BNS*, *(S)CPR*, *(S)ADS-Min*, *(S)ADS-Med* or *SM* with $Z_{t,\nu}$ and $\alpha = 5\%$. The upper (lower) panel graphs the curves for 5 minute sampling intervals with large (small) and rare jumps. A further note to *CPR*: Θ_τ^δ did not converge after $\delta = 50$ iterations using 5 minute interval returns and 60% (60% and 50%) zero-returns with large (small) and rare jumps. For this, the test statistic was computed based on $\delta = 50$.

Figure 3: Cumulative sum of $\tilde{J}_{t,\nu}^{1/2}$ and $\max_t\{|r_j|\}$



Remarks: Sample from May 2002 to June 2008. The top (bottom) panel graphs the cumulative sum of the 5% most potential jump factors to the power of 0.5 ($\tilde{J}_{t,\nu}^{1/2}$) for *(S)BNS*, *(S)CPR*, *(S)ADS-Min* and *(S)ADS-Med*. The middle panel represents the daily maximum of absolute interval returns ($\max_t\{|r_j|\}$).

Tables

Table 1: Jump and no-jump day detection rate for simulated time series: **large and rare jumps**, 5 minute sampling intervals and varying fraction of zero-returns

Zero-Ret.	*SBNS* (nj)	*SBNS* (j)	*SCPR* (nj)	*SCPR* (j)	*SADS-Min* (nj)	*SADS-Min* (j)	*SADS-Med* (nj)	*SADS-Med* (j)
0%	0.952	0.966	0.949	0.971	0.956	0.959	0.949	0.966
1%	0.950	0.966	0.947	0.971	0.956	0.959	0.956	0.968
	(0.936)	*(0.966)*	*(0.933)*	*(0.971)*	*(0.950)*	*(0.962)*	*(0.945)*	*(0.968)*
5%	0.950	0.966	0.947	0.971	0.959	0.962	0.948	0.968
	(0.850)	*(0.973)*	*(0.841)*	*(0.975)*	*(0.918)*	*(0.966)*	*(0.919)*	*(0.968)*
10%	0.949	0.975	0.945	0.975	0.957	0.962	0.941	0.973
	(0.680)	*(0.980)*	*(0.673)*	*(0.982)*	*(0.856)*	*(0.975)*	*(0.877)*	*(0.975)*
25%	0.947	0.971	0.944	0.975	0.949	0.968	0.898	0.973
	(0.199)	*(0.993)*	*(0.185)*	*(0.998)*	*(0.618)*	*(0.980)*	*(0.663)*	*(0.984)*
50%	0.960	0.971	0.957	0.973	0.953	0.966	0.851	0.977
	(0.010)	*(1.000)*	*(0.003)*	*(1.000)*	*(0.333)*	*(0.989)*	*(0.260)*	*(0.991)*

Remarks: Simulated realization of the Heston type price process with jumps ($\sigma_{jmp} = 0.1$) for 7650 days with parameter specifications and algorithm for zero-returns as of section 3.2 (here: # 442 price jumps). Each *(nj)* (*(j)*) cell reports the percentage amount of correctly identified no-jump days (jump days) with $Z_{t,\nu}$ and $\alpha = 5\%$, using either *SBNS, SCPR*, *SADS-Min* or *SADS-Med*. In brackets, the corresponding rate without employing *SIVQE* can be found.

Table 2: Jump and no-jump day detection rate for simulated time series: **small and rare jumps**, 5 minute sampling intervals and varying fraction of zero-returns

Zero-Ret.	*SBNS* (nj)	*SBNS* (j)	*SCPR* (nj)	*SCPR* (j)	*SADS-Min* (nj)	*SADS-Min* (j)	*SADS-Med* (nj)	*SADS-Med* (j)
0%	0.949	0.864	0.944	0.892	0.960	0.846	0.948	0.874
1%	0.949	0.867	0.944	0.894	0.960	0.846	0.956	0.874
	(0.932)	*(0.876)*	*(0.927)*	*(0.899)*	*(0.953)*	*(0.848)*	*(0.944)*	*(0.876)*
5%	0.945	0.867	0.942	0.892	0.958	0.851	0.946	0.878
	(0.840)	*(0.890)*	*(0.834)*	*(0.915)*	*(0.913)*	*(0.860)*	*(0.916)*	*(0.883)*
10%	0.944	0.874	0.940	0.899	0.958	0.855	0.937	0.892
	(0.669)	*(0.910)*	*(0.660)*	*(0.929)*	*(0.860)*	*(0.894)*	*(0.871)*	*(0.901)*
25%	0.945	0.885	0.942	0.903	0.956	0.876	0.895	0.897
	(0.188)	*(0.982)*	*(0.174)*	*(0.989)*	*(0.619)*	*(0.936)*	*(0.658)*	*(0.943)*
50%	0.955	0.890	0.951	0.908	0.948	0.885	0.849	0.917
	(0.010)	*(0.995)*	*(0.004)*	*(1.000)*	*(0.340)*	*(0.952)*	*(0.260)*	*(0.970)*

Remarks: Simulated realization of the Heston type price process with jumps ($\sigma_{jmp} = 0.0134$) for 7650 days with parameter specifications and algorithm for zero-returns as of section 3.2 (here: # 435 price jumps). Each *(nj)* (*(j)*) cell reports the percentage amount of correctly identified no-jump days (jump days) with $Z_{t,\nu}$ and $\alpha = 5\%$, using either *SBNS, SCPR*, *SADS-Min* or *SADS-Med*. In brackets, the corresponding rate without employing *SIVQE* can be found. A further note to *CPR* reported in brackets below *SCPR*: Θ_τ^δ did not converge after $\delta = 50$ iterations using 5 minute interval returns and 50% zero-returns. For this, the test statistic was computed based on $\delta = 50$.

Table 3: Empirical proportion of detected jump days across different levels of significance (α) for 15 minute sampling intervals

	α = 5%	α = 1%	α = 0.1%	α = 0.01%
a) Original				
BNS	0.604	0.434	0.265	0.165
CPR	0.740	0.604	0.415	0.297
ADS-Min	0.171	0.037	0.001	0.001
ADS-Med	0.288	0.117	0.028	0.004
b) SIVQE				
S*BNS*	0.048	0.013	0.002	0.001
S*CPR*	0.063	0.023	0.006	0.003
S*ADS-Min*	0.040	0.008	0.001	0.001
S*ADS-Med*	0.124	0.047	0.008	0.001

Remarks: Sample from May 2002 to June 2008. Part *a)* of the table reports the empirical proportion of detected jump days across different levels of significance using the methods *BNS*, *CPR*, *ADS-Min* and *ADS-Med*. Part *b)* of the table correspondingly reports the outputs for *SBNS*, *SCPR*, *SADS-Min* and *SADS-Med*.

Table 4: Empirical parity in occurrence time of the 5% most potential jump factors across methods with 15 minute sampling intervals

	BNS	*CPR*	*ADS-Min*	*ADS-Med*	***S**BNS*	***S**CPR*	***S**ADS-Min*	***S**ADS-Med*
BNS	1							
CPR	0.64	1						
ADS-Min	0.06	0.10	1					
ADS-Med	0.12	0.12	0.49	1				
***S**BNS*	0.13	0.13	0.22	0.30	1			
***S**CPR*	0.12	0.12	0.26	0.35	0.88	1		
***S**ADS-Min*	0.07	0.12	0.62	0.49	0.33	0.38	1	
***S**ADS-Med*	0.09	0.10	0.46	0.61	0.42	0.46	0.59	1

Remarks: The table presents the output for the parity in occurrence time of the 5% most potential jump factors using the empirical sample from May 2002 to June 2008. It further distinguishes between jump detection methods using the original setting and *SIVQE*.

Table 5: Size of the 5% most potential jump factors, and trading activity and variation on these potential jump days

				NT_t							
	$\tilde{J}Cont_t$	mean($\tilde{J}_t$)	std($\tilde{J}_t$)	min	mean	max	NP_t	$TVol_t$	$Intd_t$	mean(RV_t)	*rtp*
a) Original											
BNS	0.825	0.00016	0.00041	39	109	247	42	545	4.79	0.00020	0.072
CPR	0.898	0.00026	0.00078	39	110	558	44	503	4.74	0.00029	0.072
ADS-Min	0.719	0.00033	0.00034	55	194	675	73	822	2.88	0.00046	0.246
ADS-Med	0.687	0.00025	0.00028	71	183	598	69	784	2.99	0.00036	0.188
b) SIVQE											
S*BNS*	0.370	0.00033	0.00063	50	234	868	83	999	2.58	0.00088	0.333
S*CPR*	0.418	0.00034	0.00053	50	240	868	84	1025	2.49	0.00080	0.377
S*ADS-Min*	0.517	0.00037	0.00058	49	198	675	75	838	2.95	0.00067	0.333
S*ADS-Med*	0.534	0.00034	0.00064	55	198	598	75	834	2.87	0.00061	0.275
Full-sample	-	-	-	30	182	868	65	805	3.23	0.00028	-

Remarks: Sample from May 2002 to June 2008. $\tilde{J}Cont_t$ represents the average proportionate contribution of the 5% most potential jump factors ($\tilde{J}_t$) to daily realized variance (RV_t), NT_t is the number of trades per day, NP_t is the average number of price changes per day, $TVol_t$ is the average number of traded contracts, $Intd_t$ is the average of daily averaged intertrade duration, and *rtp* is the parity in occurrence time of the potential jump days with the upper 5% right tail of the distribution of $\max_t\{|r_j|\}$ (calculating the parity is similar in procedure to the parity measure EJ_i in equation (15)). Part *a)* of the table reports results using the methods *BNS, CPR, ADS-Min* and *ADS-Med* in their original setting. Part *b)* of the table correspondingly presents the outputs for *SBNS, SCPR, SADS-Min* and *SADS-Med*. The last row reports some corresponding full sample averages due to comparative purposes.