CORE
Metadata, citation and similar papers at core.ac.uk
Provided by DSpace at VU

TI 2009-110/4
Tinbergen Institute Discussion Paper

Spot Variance Path Estimation and its Application to High Frequency Jump Testing

Charles S. Bos
Pawel Janus
Siem Jan Koopman

VU University Amsterdam, and Tinbergen Institute.

Tinbergen Institute

The Tinbergen Institute is the institute for economic research of the Erasmus Universiteit Rotterdam, Universiteit van Amsterdam, and Vrije Universiteit Amsterdam.

Tinbergen Institute Amsterdam
Roetersstraat 31
1018 WB Amsterdam
The Netherlands
Tel.: +31(0)20 551 3500
Fax: +31(0)20 551 3555

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

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

Spot variance path estimation and its application to high frequency jump testing*

Charles S. Bos Paweł Janus† Siem Jan Koopman

Department of Econometrics, VU University Amsterdam

&

Tinbergen Institute, Amsterdam

December 3, 2009

Abstract

This paper considers spot variance path estimation from datasets of intraday high frequency asset prices in the presence of diurnal variance patterns, jumps, leverage effects and microstructure noise. We rely on parametric and nonparametric methods. The estimated spot variance path can be used to extend an existing high frequency jump test statistic, to detect arrival times of jumps and to obtain distributional characteristics of detected jumps. The effectiveness of our approach is explored through Monte Carlo simulations. It is shown that sparse sampling for mitigating the impact of microstructure noise has an adverse affect on both spot variance estimation and jump detection. In our approach we can analyze high frequency price observations that are contaminated with microstructure noise without the need for sparse sampling, say at fifteen minute intervals. An empirical illustration is presented for the intraday EUR/USD exchange rates. Our main finding is that fewer jumps are detected when sampling intervals increase.

Keywords: high frequency, intraday periodicity, jump testing, leverage effect, microstructure noise, pre-averaged bipower variation, spot variance.

JEL classification codes: C12, C13, C22, G10, G14.

*The authors wish to thank Dick van Dijk and Mark Podolskij as well as participants at the Tinbergen Institute Lunch Seminar, Computational and Financial Econometrics conference (Cyprus, October 2009), Applied Finance and Financial Econometrics conference (Berlin, November 2009) for many helpful comments and suggestions.

†Corresponding author: Paweł Janus, VU University Amsterdam, FEWEB, De Boelelaan 1105, 1081 HV Amsterdam, The Netherlands, phone: +31 20 598 5446, fax: +31 20 598 6020, email: pjanus@feweb.vu.nl.

1 Introduction

The increasing availability of high frequency financial data has led to the development of new measures of variation. It is common to model financial returns as a continuous-time real-valued stochastic process such as a jump-diffusion. A possible measure of variation of such a process is quadratic variation. Realized variance is an estimate of quadratic variation and can be computed from an of observed sequence of prices for some time-interval, typically one trading day, see Andersen, Bollerslev, Diebold, and Labys (2001) and Barndorff-Nielsen and Shephard (2002). Realized variance can be employed for a wide range of applications including volatility forecasting, measuring financial risk and testing for jumps in financial price data. An up-to-date review of realized variance and its applications is presented by Andersen and Benzoni (2009). In financial market applications, quadratic variation is decomposed into integrated variance and jump variation. Integrated variance is defined as the integral of so-called spot variance. Jump variation is defined as a sum of squared jumps and is estimated by the difference between estimates of quadratic variation and integrated variance. Estimated jump variation indicates a presence, if any, of a jump component in a given time-interval. It does not give an insight into the arrival time, size or direction of a jump. For this we require the spot variance path in high frequency data, see Lee and Mykland (2008). In this paper we modify an existing estimator of integrated variance for the purpose of estimating the spot variance path and with the aim to improve testing procedures for jumps.

Most contributions for estimating spot variance are based on nonparametric methods. Under the absence of jumps and microstructure noise, estimation of spot variance is discussed, among others, in Foster and Nelson (1996), Andreou and Ghysels (2002), Alvarez, Panloup, Pontier, and Savy (2008), Fan and Wang (2008) and Kristensen (2009). Kinnebrock (2008), Mykland and Zhang (2008) and Ogawa and Sanfelici (2008) consider estimation when price observations are contaminated with microstructure noise. Bandi and Reno (2009) discuss estimation when price data contain either microstructure noise or jumps by making use of different consistent estimators of integrated variance. Boudt, Croux, and Laurent (2008) and Lee and Mykland (2008) estimate spot variance in the presence of jumps, but they rule out microstructure noise. In this paper, however, we adapt the pre-averaging approach and estimate the spot variance path when jumps and microstructure noise are both present in high frequency price observations. In a Monte Carlo study we find that leverage effects do not have any adverse effect on the estimation performance. However, we also find that diurnal patterns can lead to systematic bias in the estimation. We therefore modify the method further to extract a periodic diurnal component in a robust manner.

Earlier theoretical and empirical contributions have focussed on modifying estimation methods for integrated variance using high frequency data to allow for jumps, leverage effects, microstructure distortions and periodic variance patterns. This work has led to mild assumptions for the spot variance process; for example, see Barndorff-Nielsen, Graversen, Jacod, and Shephard (2006). The application of the multipower variation estimation methods for testing for jumps using high frequency data has

shown that it is necessary to allow for jumps into models for financial data. The importance of jumps is revealed by Huang and Tauchen (2005) who found that jumps account for around 7% of S&P cash index variation. It implies that jumps in financial returns have a severe impact on risk management, option pricing and derivative hedging. For example, Tauchen and Zhou (2006) have shown that for investment-grade bond spread indices, jump variation yields more forecasting power than interest rates and other volatility measures. It illustrates the importance of developing sound statistical methods for the detection and characterization of jumps in financial returns.

Related work on high frequency jump testing with the use of nonparametric methods is presented by Barndorff-Nielsen and Shephard (2006) where jumps are detected on the basis of a standardized difference between quadratic variation and integrated variance. A similar test for the robustness of microstructure noise is recently proposed by Podolskij and Vetter (2009a). Jiang and Oomen (2008) propose a test based on higher order return moments, that can also be employed in the presence of microstructure noise. Aït-Sahalia and Jacod (2009) use multipower variation at different sampling intervals to detect jumps. These tests are able to detect the presence of jumps in a fixed time interval. Alternatively, Andersen, Bollerslev, and Dobrev (2007) and Lee and Mykland (2008) have developed tests for finding a jump in each realized price increment. It reveals the timing as well as the distributional characteristics of realized jumps in terms of mean, variance or intensity of the jump process. Moreover, realized jumps can provide an insight into the nature of jumps for asset classes or common movements (co-jumps) between assets, see Lahaye, Laurent, and Neely (2009). Parametric methods to test for jumps are adapted by Duan and Fülöp (2007) and Bos (2008) where jumps are detected as part of the estimation method.

Boudt et al. (2008) have shown that pronounced intraday periodicity leads to the distorted jump inference based on the Lee and Mykland (2008) test. They modify the test to account for diurnal variance patterns. Both tests rule out microstructure noise and rely on an estimate of the spot variance path that is not robust to microstructure noise. Empirical applications of these tests are therefore based on sampling intervals that mitigate noise effects. However, sampling at high frequencies is of key importance for spot variance estimation and jump detection. Since financial time series are now widely available at high frequencies (see Table 1 of Shephard and Sheppard, 2009) sampling at low frequencies discards available information. For instance, if asset prices are observed each second, then sampling every 15 minutes discards 899 out of every 900 observations.

In this paper we adopt a consistent estimator of integrated variance to estimate the spot variance path based on price observations which are sampled at ultra-high frequencies. The estimated spot variance path in a jump-diffusion framework with microstructure noise allows us to extend the jump test statistic of Lee and Mykland (2008). We study the impact of leverage effects, microstructure noise and diurnal patterns on our proposed spot variance estimate as well as on our extended jump statistic by means of a Monte Carlo study.

We conduct an empirical study for the intraday EUR/USD exchange rates at different sampling

intervals over the period of July 2 to December 31, 2007. The difference between the daily estimates of quadratic variation and integrated variance indicates the presence of jumps at the considered sampling intervals. However, when sampling intervals increase, we are not able to detect jumps for some days even though the estimated relative jump variation has a substantial level. As shown in the simulations, the accuracy of the spot variance estimation and the power in jump detection both increase when based on finer sampling intervals.

The remainder of the paper is organized as follows. Section 2 discusses the estimation of spot variance path. Section 3 deals with the robust estimation of the intraday return periodicity. Section 4 reviews the Lee-Mykland jump test and proposes an adjusted statistic. In Section 5 we present and discuss the results of a Monte Carlo study. Section 6 presents an empirical illustration, while Section 7 concludes.

2 Integrated and Spot Variance

2.1 General framework

For the modeling of asset prices in continuous-time, we adopt a jump-diffusion process in a similar way as Barndorff-Nielsen and Shephard (2004), Jiang and Oomen (2008), Lee and Mykland (2008) and many others. The efficient log price process X_t is assumed to follow a Brownian semi-martingale plus jumps defined as

$$\mathrm{d}X_t = \sigma_t p_t \mathrm{d}W_t + \kappa_t \mathrm{d}N_t, \qquad t \geq 0, \tag{2.1}$$

where $\sigma_t > 0$ is the stochastic part of spot volatility, $p_t > 0$ is the deterministic diurnal part of spot volatility, W_t is a standard Brownian motion, κ_t is a random variable with mean $\mu_\kappa(t)$ and variance $\sigma^2_\kappa(t)$ and N_t is a counting process that represents the number of jumps in the price path up to time t. We allow for leverage effects, that is (negative) correlation between W_t and σ_t. In this section we have $p_t = 1,\ \forall t$ in (2.1) and in Section 3 we take p_t as a deterministic function of time. Our aim is to estimate spot variance σ^2_t by means of a consistent estimator of integrated variance for a local window $[t-h,\, t]$ with $h \to 0$. We assume some degree of smoothness for the spot variance path as in Lee and Mykland (2008), Ogawa and Sanfelici (2008) and Kristensen (2009). For example, in the Monte Carlo study of Section 5, we use the Heston variance process driven by Brownian motion.

The price observations for (2.1) are assumed to be available at normalized equidistant times $0 = t_0 < t_1 < \ldots < t_n = 1$, where the interval $[0, 1]$ represents a trading day. Let $\Delta = t_i - t_{i-1} = n^{-1}$ be the distance between two adjacent price observations such that $\Delta \to 0$ when $n \to \infty$. The observed log price Y_{t_i} is the underlying efficient log price X_{t_i} plus error,

$$Y_{t_i} = X_{t_i} + \varepsilon_{t_i}, \tag{2.2}$$

where ε_{t_i} represents market microstructure noise with $\mathbb{E}[\varepsilon_{t_i}] = 0$, $\mathbb{E}[\varepsilon_{t_i}\varepsilon_{t_k}] = 0$ for $i \neq k$ and $\mathbb{E}[\varepsilon^2_{t_i}] = \varpi^2$. We further assume that the efficient price X_{t_i} and the microstructure noise ε_{t_i} are independent

of each other at all lags and leads. The variance and autocovariance structure of price increments is then given by

$$\mathbb{E}\left[(Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-k}} - Y_{t_{i-1-k}})\right] = \begin{cases} \int_{t_{i-1}}^{t_i} \sigma_s^2 \mathrm{d}s + 2\varpi^2, & \text{if} \quad k = 0, \\ -\varpi^2, & \text{if} \quad k = 1, \\ 0, & \text{if} \quad k > 1. \end{cases} \tag{2.3}$$

The first order autocorrelation coefficient of price increments is negative with a lower limit of $-1/2$ and all higher order autocorrelations are zero.

2.2 Integrated variance

The quadratic variation (QV) of the efficient price process (2.1) over the interval $[0, 1]$ (trading day) is defined as

$$\mathrm{QV}_{[0,1]} = \operatorname*{plim}_{\Delta \to 0} \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2, \tag{2.4}$$

which in case of price process (2.1) can be decomposed into integrated variance (IV) and jump variation (JV)

$$\mathrm{QV}_{[0,1]} = \mathrm{IV}_{[0,1]} + \mathrm{JV}_{[0,1]}, \tag{2.5}$$

where

$$\mathrm{IV}_{[0,1]} = \int_0^1 \sigma_t^2 \mathrm{d}t \qquad \text{and} \qquad \mathrm{JV}_{[0,1]} = \sum_{j=1}^{N_1} \kappa_j^2. \tag{2.6}$$

The convergence results of the IV estimators rely on the continuous-time framework. For instance, the realized bipower variation (BPV) defined as

$$\mathrm{BPV}_{[0,1]} = \frac{\pi}{2} \sum_{i=2}^{n} |X_{t_i} - X_{t_{i-1}}||X_{t_{i-1}} - X_{t_{i-2}}|, \tag{2.7}$$

is a consistent estimator of $\mathrm{IV}_{[0,1]}$ in the absence of microstructure noise, see Barndorff-Nielsen and Shephard (2004). However, increasing the sampling frequency in the presence of microstructure noise leads to a severe bias in BPV, see Huang and Tauchen (2005). For consistent estimation of IV in the presence of jumps and noise, one can use the pre-averaging approach. We use the pre-averaged bipower variation (PBPV) estimator, see Jacod, Li, Mykland, Podolskij, and Vetter (2009) and Podolskij and Vetter (2009a,b).

For applying PBPV, we select $\Theta \in (0, \infty)$ and integer k_n such that

$$k_n \sqrt{\Delta} = \Theta + o(n^{-\frac{1}{4}}), \tag{2.8}$$

and we select a weight function $g(u)$ on the $[0, 1]$ interval for some variable u with $g(0) = g(1) = 0$. Jacod et al. (2009) and Podolskij and Vetter (2009a,b) take $g(u) = \min(u, 1 - u)$ for $0 \leq u \leq 1$. The PBPV estimator is then given by

$$\mathrm{PBPV}_{[0,1]}(l, r) = n^{\frac{l+r}{4}-1} \sum_{i=0}^{n-2k_n+1} |\bar{Y}_i|^l |\bar{Y}_{i+k_n}|^r, \tag{2.9}$$

where

$$\bar{Y}_i = \sum_{j=1}^{k_n-1} g\left(j/k_n\right)\left(Y_{t_{i+j}} - Y_{t_{i+j-1}}\right), \tag{2.10}$$

and for $l, r > 0$. The estimator (2.9) is the (l, r) order of pre-averaged bipower variation. We only consider the cases $(l, r) = (2, 0)$ and $(l, r) = (1, 1)$. The PBPV estimator (2.9) is biased due to microstructure noise (see Theorem 1 and 2 of Podolskij and Vetter, 2009a). The bias term depends on the microstructure noise variance ϖ^2 which we estimate on the basis of (2.3), that is

$$\widehat{\varpi}^2 = -\frac{1}{n-1}\sum_{i=2}^{n}(Y_{t_i} - Y_{t_{i-1}})(Y_{t_{i-1}} - Y_{t_{i-2}}), \tag{2.11}$$

see Oomen (2006). The bias-corrected PBPV is then given by

$$\widehat{\text{PBPV}}_{[0,1]}(l, r) = \frac{\mu_l^{-1}\mu_r^{-1}}{\Theta\varphi_2}\text{PBPV}_{[0,1]}(l, r) - \frac{\varphi_1}{\Theta^2\varphi_2}\widehat{\varpi}^2, \tag{2.12}$$

where $\mu_p = \mathbb{E}[|z|^p]$ for $z \sim \mathcal{N}(0, 1)$ such that $\mu_1 = \sqrt{2}/\sqrt{\pi}$ and $\mu_0 = \mu_2 = 1$ and where

$$\varphi_1 = \int_0^1 (g'(u))^2 \mathrm{d}u, \qquad \varphi_2 = \int_0^1 (g(u))^2 \mathrm{d}u,$$

with $g'(u)$ as the first derivative of $g(u)$ with respect to u. In case $g(u) = \min(u, 1-u)$, we have $\varphi_1 = 1$ and $\varphi_2 = \frac{1}{12}$. The estimate $\widehat{\text{PBPV}}_{[0,1]}(l, r)$ with $(l, r) = (2, 0)$ measures daily quadratic variation $\text{QV}_{[0,1]}$ and with $(l, r) = (1, 1)$ measures daily integrated variance $\text{IV}_{[0,1]}$. The relative jump variation (RJV) is estimated by

$$\widehat{\text{RJV}}_{[0,1]} = 100 \times \frac{\widehat{\text{PBPV}}_{[0,1]}(2, 0) - \widehat{\text{PBPV}}_{[0,1]}(1, 1)}{\widehat{\text{PBPV}}_{[0,1]}(2, 0)}, \tag{2.13}$$

and measures the percentage contribution of jumps in total price variation for a trading day. Finally, Jacod et al. (2009) show that for the special case with $g(u) = \min(u, 1-u)$, $\sigma_t \equiv \sigma$ and $\mathrm{d}N_t = 0, \forall t$ in (2.1), an optimal Θ is obtained by minimizing the asymptotic variance of the estimator, and they obtain $\Theta = 4.777\, \varpi/\sigma$. We follow the authors in their suggestion to fix $\Theta = 1/3$, such that $k_n = [1/3\sqrt{n}\,]$; for instance, $k_n = 98$ when $n = 86400$ and $\Delta = 1sec$ in a $24h$ market.

2.3 Spot variance

The spot variance is defined as the derivative of the integrated variance

$$\sigma_t^2 = \lim_{h \to 0} \frac{\text{IV}_{[t-h,t]}}{h}, \tag{2.14}$$

where $\text{IV}_{[t-h,t]} = \int_{t-h}^{t} \sigma_s^2 \mathrm{d}s$, $h > 0$, and can be estimated by

$$\widehat{\sigma}_t^2 = \frac{n\widehat{\text{IV}}_{[t-h,t]}}{h_n}, \tag{2.15}$$

where $\widehat{\text{IV}}_{[t-h,t]}$ is an unbiased and consistent estimator of $\text{IV}_{[t-h,t]}$, and $h_n = hn$, with $h_n \to \infty$ satisfying $h_n/n \to 0$. The spot variance can be estimated for each observation at time $t_i = \lfloor t/\Delta \rfloor$. In the absence of microstructure noise, spot variance is consistently estimated by

$$\widehat{\sigma}_{t_i}^2 = \frac{\mu_1^{-2} n}{h_n} \sum_{j=i-h_n+1}^{i} |X_{t_j} - X_{t_{j-1}}||X_{t_{j-1}} - X_{t_{j-2}}|, \tag{2.16}$$

for the ith observation. However, to estimate σ_t^2 in the presence of jumps and additive noise, we adopt (2.12) to obtain

$$\widehat{\sigma}_{t_i}^2 = \frac{\mu_1^{-2} n^{1/2}}{\Theta \varphi_2 h_n} \sum_{j=i-h_n}^{i-2k_n} |\bar{Y}_j||\bar{Y}_{j+k_n}| - \frac{\varphi_1}{\Theta^2 \varphi_2} \widehat{\varpi}^2, \tag{2.17}$$

for the ith observation. Since k_n is of order $n^{1/2}$, we consider $h_n = n^\beta$ with $\beta \in (1/2, 1]$. In case $i - h_n < 0$, we take a practical stance and use price realizations from a previous day, disregarding issues related to overnight returns. We refer to Bandi and Reno (2009) for a general theory of spot variance estimation.

The choice of the local window h for $\mathrm{IV}_{[t-h,t]}$ estimation is subject to a trade-off between bias and variance of the spot variance estimator. A small h leads to an increased variation of the estimator but it does not restrict σ_t^2 to be constant over a larger time interval. A large h assumes σ_t^2 to be constant over a large window and increases bias when spot variance is more volatile. The simulation study in Section 5.3 will show that different spot variance specifications and different choices of window length in (2.17) lead to different levels of estimation error.

3 Intraday Periodicity

3.1 Preliminaries

The intraday periodic patterns in return variance are mostly due to the opening, lunch-break and closing of the own market and other markets. The variance and autocorrelation functions of absolute intraday returns are typically U-shaped and persistent, see Andersen and Bollerslev (1997). We allow for diurnal patterns by decomposing spot volatility in the price process (2.1) into a stochastic variable σ_t and a smooth deterministic multiplier p_t.

3.2 Model with periodicity and noise

The efficient log price process (2.1) has spot volatility $\sigma_t p_t$ where p_t is higher (lower) than unity when intraday trading activity due to periodic effects increases (decreases). We further assume that

(i) $\int_0^1 p_t^2 \mathrm{d}t = 1$;

(ii) the diurnal pattern p_t is the same for each trading day;

(iii) the efficient price increment $X_{t_i} - X_{t_{i-1}}$ is observed with noise term $\varepsilon_{t_i} - \varepsilon_{t_{i-1}}$ that has mean zero and variance $2\varpi^2 p_{t_i}^2$.

In practice, assumption (i) is sufficient to ensure that the integrated variance $\mathrm{IV}_{[0,1]}$ estimate for an entire trading day is not affected by intraday periodicity p_t, see also the discussion in Andersen (2004). However, in case we estimate $\sigma_t p_t$ using observations from an interval smaller than one trading day, diurnal patterns do matter. Assumption (iii) is motivated by Kalnina and Linton (2008) who have

considered equity data and estimated U-shaped intraday variance patterns in microstructure noise. However, it does imply that the noise-to-signal ratio

$$\mathrm{NSR}_{t_i} = \frac{2\varpi^2 p_{t_i}^2}{\sigma_{t_i}^2 p_{t_i}^2 \Delta + 2\varpi^2 p_{t_i}^2} = \frac{2\varpi^2}{\sigma_{t_i}^2 \Delta + 2\varpi^2}, \tag{3.1}$$

is not subject to a diurnal shape. Otherwise, the constant Θ in the pre-averaged based estimator (2.8) would be a function of the diurnal shape p_t. The issue of periodic behavior of Θ is not considered here. The diurnal effect is therefore a strict multiplicative effect and it does not affect the noise-to-signal ratio.

3.3 Periodicity extraction

Assumption (ii) implies that the estimation of the diurnal effect can be based on financial returns from multiple trading days. For this purpose, we denote the observed price, spot variance, microstructure noise variance and pre-averaged bipower estimator (2.12) at time t_i on day d by

$$Y_{d+t_i}, \qquad \sigma_{d+t_i}^2, \qquad \varpi_d^2, \qquad \text{and} \qquad \widehat{\mathrm{PBPV}}_d(l,r),$$

respectively, for days $d = 1, \ldots, D$. We follow Boudt et al. (2008) in benchmark estimation of the periodic component p_{t_i} on the basis of cross-sectional variances. In case of no jumps in the price process (2.1), the mean and variance of the price increment

$$Y_{d+t_i} - Y_{d+t_{i-1}} = X_{d+t_i} - X_{d+t_{i-1}} + \varepsilon_{d+t_i} - \varepsilon_{d+t_{i-1}}, \tag{3.2}$$

are given by

$$\mathbb{E}\left[Y_{d+t_i} - Y_{d+t_{i-1}}\right] = 0 \quad \text{and} \quad \mathbb{V}ar\left[Y_{d+t_i} - Y_{d+t_{i-1}}\right] = \sigma_{d+t_i}^2 p_{t_i}^2 \Delta + 2\varpi_d^2 p_{t_i}^2. \tag{3.3}$$

Given assumption (ii), price increments at time t_i for all D days have a common deterministic multiplier p_{t_i}. When all price increments are standardized based on $\sigma_{d+t_i}^2 \Delta + 2\varpi_d^2$, then the variance of the standardized increments is equal to $p_{t_i}^2$. The cross-section over days of the standardized increments at time t_i is used to estimate $p_{t_i}^2$ using all D days.

The estimation of $p_{t_i}^2$ requires estimates for $\sigma_{d+t_i}^2$ and ϖ_d^2. We treat $\sigma_{d+t_i}^2$ as constant over the dth day and $\widehat{\mathrm{PBPV}}_{[0,1]}(1,1)$ of (2.12) provides its estimate. We take (2.11) as an estimate for ϖ_d^2. To satisfy assumption (i), we normalize the diurnal estimator for $p_{t_i}^2$ by

$$\widehat{p}_{t_i}^2 = \frac{n \sum_{d=1}^{D} R_{d+t_i}^2}{\sum_{j=1}^{n} \sum_{d=1}^{D} R_{d+t_j}^2}, \tag{3.4}$$

where

$$R_{d+t_i}^2 = \frac{\left(Y_{d+t_i} - Y_{d+t_{i-1}}\right)^2}{n^{-1}\widehat{\mathrm{PBPV}}_{[0,1]}(1,1) + 2\widehat{\varpi}_d^2}, \tag{3.5}$$

for $i = 1, \ldots, n$. Our approach differs from Andersen and Bollerslev (1998) in which $\sigma_{d+t_i}^2$ is estimated on the basis of a generalized autoregressive conditional heteroskedasticity model for daily returns and hence does not use intra-daily data for the estimation of $\sigma_{d+t_i}^2$.

3.4 Robust estimation of intraday periodicity

In case jumps are present in the price process (2.1), we need to modify estimator (3.4) to keep it smooth. We propose to locally smooth the cross-sectional variances (3.4) via the weighted average

$$\widehat{\bar{p}}^2_{t_i} = \sum_{j=1}^{n} \omega_{ij} \widehat{p}^2_{t_j}, \tag{3.6}$$

where ω_{ij} is a pre-determined weight and $\sum_{j=1}^{n} \omega_{ij} = 1$, for $i = 1, \ldots, n$. We let the weights be implied by Kalman filtering and smoothing methods applied to a local level model given by

$$\widehat{p}^2_{t_i} = \bar{p}^2_{t_i} + \zeta_{t_i}, \tag{3.7}$$

$$\bar{p}^2_{t_{i+1}} = \bar{p}^2_{t_i} + \eta_{t_i}, \tag{3.8}$$

where we treat $\widehat{p}^2_{t_i}$ as the observation, $\bar{p}^2_{t_i}$ as the latent variable (local level) and with the disturbances ζ_{t_i} and η_{t_i} are serially independent, with mean zero and variances σ^2_ζ and σ^2_η respectively. The local level is modeled as a random walk process with Gaussian disturbances (or increments) η_{t_i}. Since we expect that jumps cause outliers in the "observations" $\widehat{p}^2_{t_i}$, we let the disturbances ζ_{t_i} come from a Student's t distribution. Only for the purpose of comparison, we also consider ζ_{t_i} to come from a Gaussian density. This working model for $\widehat{p}^2_{t_i}$ does not have any implications for our theoretical framework of Section 2, it only facilitates the construction of weights in (3.6).

In case all disturbances are Gaussian, the estimation of $[\sigma_\zeta, \sigma_\eta]$ and the signal extraction (3.6) can be carried out by the Kalman filter and corresponding smoother. In case we let ζ_{t_i} come from a Student's t distribution, estimation of $[\sigma_\zeta, \sigma_\eta, \nu]$ and signal extraction relies on computationally efficient Monte Carlo simulation methods that employ importance sampling techniques, see Durbin and Koopman (2001) for a detailed discussion. In both cases, the computations are carried out in `Ox` Doornik (2006) using `SsfPack` of Koopman, Shephard, and Doornik (1999).

4 Intraday Jump Testing

4.1 Lee and Mykland test

Consider the price process (2.1) with $p_t = 1$ and assume that a realization X_{t_i} from (2.1) can be observed without noise. The jump test statistic of Lee and Mykland (2008) standardizes the increment $X_{t_i} - X_{t_{i-1}}$ by a local measure of variation for the continuous part of the price process. In case a jump occurs in the $(t_{i-1}, t_i]$ interval, the price increment $X_{t_i} - X_{t_{i-1}}$ is obviously larger than we can expect from the continuous part of the price process. The Lee and Mykland jump test statistic (LMJ) exploits this notion and proposes a nonparametric statistic to test the hypothesis whether a jump has occurred in the interval $(t_{i-1}, t_i]$. The LMJ test is based on the asymptotic distribution of the maximums of absolute values of standard normal variables,

$$\frac{\max_{i \in [1,n]} |\mathcal{L}(t_i)| - C_n}{S_n}, \tag{4.1}$$

where the constants C_n and S_n are given by

$$S_n = (2\log n)^{-\frac{1}{2}} \qquad \text{and} \qquad C_n = S_n^{-1} - \frac{1}{2}(\log \pi + \log\log n)S_n, \tag{4.2}$$

and where

$$\mathcal{L}(t_i) = \frac{X_{t_i} - X_{t_{i-1}}}{\widehat{\sigma}_{t_i}\sqrt{\Delta}}, \tag{4.3}$$

is standard normally distributed under the null hypothesis of no jumps. Lee and Mykland (2008) compute the spot variance estimate $\widehat{\sigma}^2_{t_i}$ by the locally averaged bipower variation (2.16) over the window $[t_{i-h_n}, t_{i-1}]$ with sufficiently large h_n, say $h_n = 270$ for $\Delta = 5min$. The test statistic (4.1) has a standard Gumbel distribution. The null hypothesis of no jump in $(t_{i-1}, t_i]$ is rejected when $|\mathcal{L}(t_i)| > S_n\alpha^* + C_n$ at the significance level $\alpha\%$ with $\alpha^* = -\log(-\log(1-\alpha))$.

The arrival of a jump of size κ^* in the interval $(t_{i-1}, t_i]$ leads to an increase of $X_{t_i} - X_{t_{i-1}}$ by κ^* and thus $\mathcal{L}(t_i)$ in (4.3) is not standard normally distributed. The jump size κ^* in $\mathcal{L}(t_i)$ is divided by $\widehat{\sigma}_{t_i}\sqrt{\Delta}$ and $\mathcal{L}(t_i) \to \infty$ when $\Delta \to 0$, so statistic becomes sufficiently large to detect the jump as $\Delta \to 0$, implying that more frequent sampling increases the power of the test (see Theorem 2 of Lee and Mykland, 2008).

4.2 A jump test correction for periodicity and microstructure noise

When the locally averaged bipower variation estimator is used for spot variance in the presence of microstructure noise, the statistic $\mathcal{L}(t_i)$ in (4.3) is not normally distributed under the null hypothesis of no jumps. Therefore, the LMJ test statistic (4.1) is not valid when microstructure noise is present. Also, when the return variation is subject to periodicity, Andersen et al. (2007) and Boudt et al. (2008) have argued that spurious jumps can be detected. Here we correct the LMJ test for both periodicity and microstructure noise.

Given the assumptions (i)-(iii) of Section 3.2 and under the null hypothesis of no jumps in the price process (2.1), the realized price increment is given by (3.2) with its mean and variance given by (3.3). We adopt these results to modify the LMJ test statistic (4.1) by replacing $\mathcal{L}(t_i)$ with $\widetilde{\mathcal{L}}(t_i)$ as given by

$$\widetilde{\mathcal{L}}(t_i) = \frac{Y_{t_i} - Y_{t_{i-1}}}{\widehat{s}_{t_i}} \quad \text{where} \quad \widehat{s}^2_{t_i} = \widehat{\sigma}^2_{t_i}\widehat{p}^2_{t_i}\Delta + 2\widehat{\varpi}^2\widehat{p}^2_{t_i}, \tag{4.4}$$

where the periodic multiplier $p^2_{t_i}$ is estimated as described in Section 3.4 and where $\widehat{\sigma}^2_{t_i}$ and $\widehat{\varpi}^2$ are given by (2.17) and (2.11), respectively. The variable $\widetilde{\mathcal{L}}(t_i)$ is standard normal and our corrected test statistic (4.1) remains to have a Gumbel distribution under the null hypothesis of no jumps.

The arrival of a jump of size κ^* in the interval $(t_{i-1}, t_i]$ leads to an increase of $Y_{t_i} - Y_{t_{i-1}}$ by κ^* and thus $\widetilde{\mathcal{L}}(t_i)$ in (4.4) is not standard normally distributed. The jump size κ^* in $\widetilde{\mathcal{L}}(t_i)$ is divided by $\widehat{s}_{t_i}$ and the discontinuous part is proportional to $\kappa^*/\sqrt{2}\varpi p_{t_i}$ as $\Delta \to 0$. Hence, jump detection in the presence of microstructure noise depends on the magnitude of price contamination. When the standard deviation of microstructure noise increases, the statistic $\widetilde{\mathcal{L}}(t_i)$ decreases, with the implication that small jumps are more difficult to detect.

5 Monte Carlo Studies

In this section we conduct a set of Monte Carlo experiments to study the finite sample properties of the intraday periodicity and the spot variance path estimation procedures as discussed in the previous sections. Furthermore we study the small-sample performance of our intraday jump testing procedure. We first present the design of the Monte Carlo study.

5.1 Design of Monte Carlo study

The price series are generated by (2.1) which depends on assumptions for spot volatility σ_t, periodic multiplier p_t and random jump process κ_t. We adopt the mean-reverting Heston (1993) process for the spot variance σ_t^2 that is given by

$$\mathrm{d}\sigma_t^2 = \psi(\theta - \sigma_t^2)\mathrm{d}t + \gamma\sigma_t \mathrm{d}W_t^\sigma, \tag{5.1}$$

where coefficient $\psi > 0$ determines the rate at which the variance σ_t^2 reverts to the long-run variance $\theta > 0$, coefficient γ determines the variation of the spot variance path and W_t^σ is a standard Brownian motion. The Brownian motions W_t in (2.1) and W_t^σ are possibly contemporaneously correlated, that is $\mathbb{E}[\mathrm{d}W_t \mathrm{d}W_t^\sigma] = \rho \mathrm{d}t$ with $\rho \in (-1, 0]$. In case $\rho \neq 0$, the variance process is subject to a leverage effect. The multiplicative periodic factor p_t in (2.1) is assumed to be a smooth function. The jump component κ_t in (2.1) is given by the product $\bar{\kappa}\sqrt{\theta} p_t U_t$ where $\bar{\kappa}$ determines the size of the jump relative to the long-run volatility and periodicity, and where U_t is a random variable which takes the values -1 and 1 with equal probability. To produce the process N_t in (2.1) we allocate via the homogeneous Poisson process at least one jump over a trading day.

We simulate sample paths for the price process with $X_0 = 0$ using the Euler discretization method at a time interval $t_{i+1} - t_i = 86400^{-1}$, corresponding to the one second frequency in a 24h market. Then we sample observations from the price path at different frequencies Δ; by increasing Δ the sampling becomes more sparse, which mitigates microstructure effects. The price increment $X_{t_i} - X_{t_{i-1}}$ is contaminated by microstructure noise as implied by assumption (iii) of Section 3.2.

Our Monte Carlo design is based on the following settings:

- For each experiment, we sample $M = 2500$ realizations of the prices process. Each realization is a time series of prices for a single 24h trading day. The length of the time series depends on sampling frequency $\Delta \in \{1sec, 5sec, 30sec, 1min, 5min\}$.

- The stochastic variance specification (5.1) is adopted from Heston (1993) and is given by the two variants

  i) Heston-A: $\psi = 1$, $\theta = 0.04$, $\gamma = 0.15$;

  ii) Heston-B: $\psi = 8$, $\theta = 0.04$, $\gamma = 0.4$.

where both variants satisfy the condition $2\psi\theta \geq \gamma^2$ such that the variance is strictly positive. The Heston-A specification has the smallest mean-revision parameter and the smallest value for γ. In both specifications we have $\theta = 0.04$ that corresponds to a long-run volatility level of 20%.

- The multiplicative periodic diurnal factor is given by $p^2_{t_i} = 1 - \phi\cos(2\pi i/n)$ with $\phi \in [0, 1)$. For $\phi = 0$, we have no diurnal shape in the return variance and in the microstructure noise variance.

- The jump size is taken from the set $\bar{\kappa} \in \{0, 0.05, 0.15, 0.25, 0.5\}$. For these values of $\bar{\kappa}$, the mean estimates of relative jump variation in (2.13) at $\Delta = 1sec$ have values $\widehat{\text{RJV}}_{[0,1]} \in \{0.22\%, 0.34\%, 2.24\%, 6.63\%, 23.26\%\}$ and $\widehat{\text{RJV}}_{[0,1]} \in \{0.22\%, 0.31\%, 2.21\%, 6.59\%, 23.19\%\}$ with $\varpi = 0.001$ and $\varpi = 0.004$ respectively. When the data is sampled more sparsely and/or microstructure noise increases, the impact of jumps decreases. In case $\bar{\kappa} \neq 0$, we enforce at least one jump in a trading day.

- The microstructure noise standard deviation is taken from the set $\varpi \in \{0, 0.001, 0.004\}$. For positive values of ϖ, the ratios of microstructure noise variance and integrated variance correspond to values found in equity data by Hansen and Lunde (2006, Table 3).

- When a Monte Carlo experiment requires the estimation of periodic factors, we repeat the simulations for $D = 100$ trading days. For all other experiments we set $D = 1$.

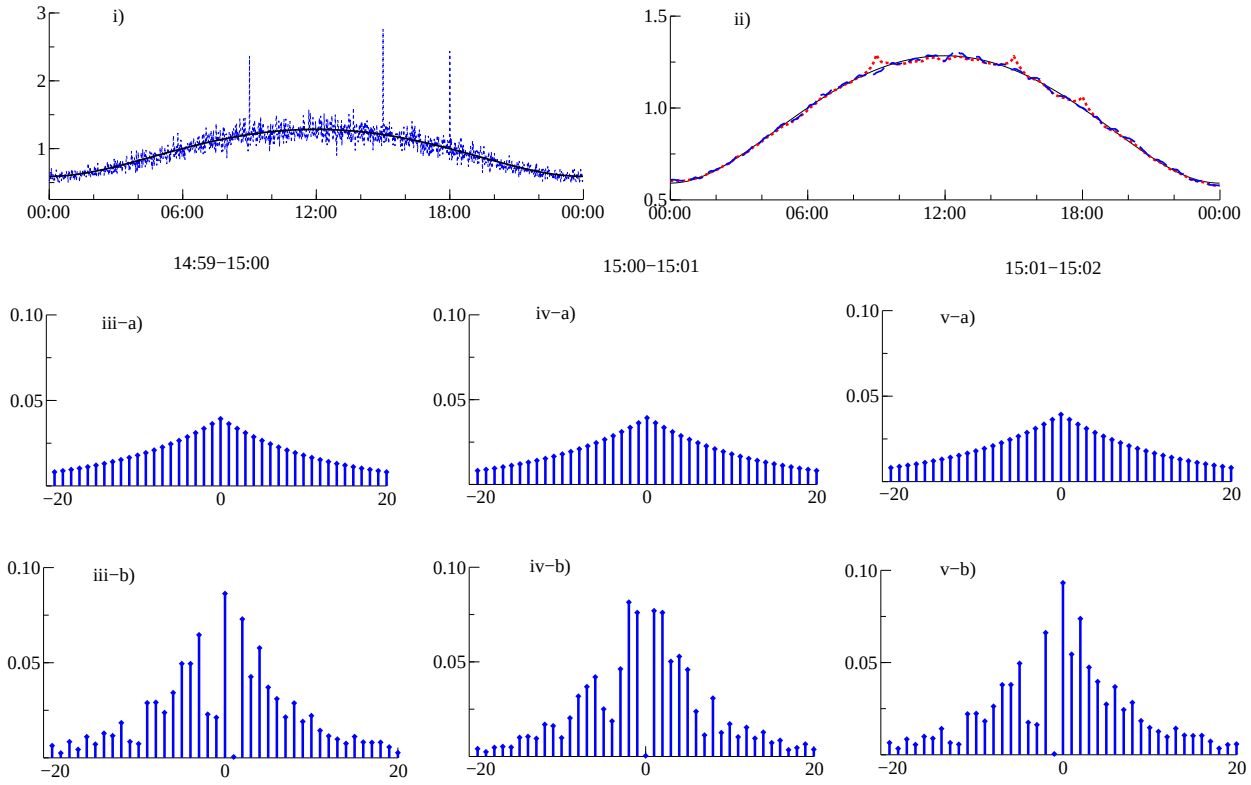
Each experiment depends on the choice of parameters: ϕ (periodicity), $\bar{\kappa}$ (jump size), ϖ (standard deviation of microstructure noise) and ρ (leverage effect). Their subsequent values are specified for each experiment and reported in the discussions, figures and tables.

5.2 Periodic patterns in spot variance

We first investigate whether the cross-sectional estimation of the periodic pattern (3.4) in the spot variance is robust to jumps. To measure the impact of jumps and its sensitivity to different weight adjustments in (3.6), we carry out the following experiment. Given the features of the empirical data in the next section, we consider jumps arriving at fixed time arrivals (news announcements at 9:00-9.01, 15.00-15.01 and 18.00-18.01) and at random time arrivals. The jumps themselves are drawn with $\bar{\kappa} = 0.15$. The remaining Monte Carlo parameters are given by $\phi = 0.65$, $\varpi = 0.001$ and $\rho = 0$.

Figure 1 presents the estimated periodic pattern. The benchmark estimate of the diurnal shape from (3.4) shown in panel *i)* of Figure 1 is affected by jumps. The scheduled news announcements have an adverse effect. The smoothing method in (3.6) is based on the model (3.7)-(3.8) with Gaussian or Student's t disturbances in (3.7). Observation based weighting method (OWM) in (3.6) with Gaussian (Student's t) observation disturbances is denoted as OWM-N (OWM-t). Both estimates are presented in panel *ii)* of Figure 1 and are more smooth relative to benchmark estimate. However, the symmetric weights ω_{ij} implied by the model with Gaussian observation disturbances do not adjust sufficiently at the jump location and its neighboring time intervals; see panels *iii-a)* – *v-a)* of Figure 1. The

Figure 1: *Periodic factor estimation. Heston-A specification with* $\phi = 0.65$, $\bar{\kappa} = 0.15$, $\varpi = 0.001$, $\rho = 0$, $\Delta = 1min$ *and* $D = 100$. *Panel i) presents square root of cross-sectional variances (dotted line) and true periodic factor (solid line); ii) presents estimates from OWM-N (dotted line), OWM-t (dashed line) and true periodic factor (solid line); panels iii), iv) and v) present* ω_{ij} *in (3.6) for 20 lags/leads for OWM-N iii-a)* $-$ *v-a), and OWM-t, iii-b)* $-$ *v-b), in time interval 14:59-15:00, 15:00-15:01 and 15:01-15:02 respectively.*



weights implied by Student's t observation disturbances are small or zero at jump locations; see panels *iii-b)* $-$ *v-b)* of Figure 1. Also, neighboring weights are not affected by the actual jumps. We can conclude that our last and preferred extraction procedure is robust to jumps.

5.3 Spot volatility estimation

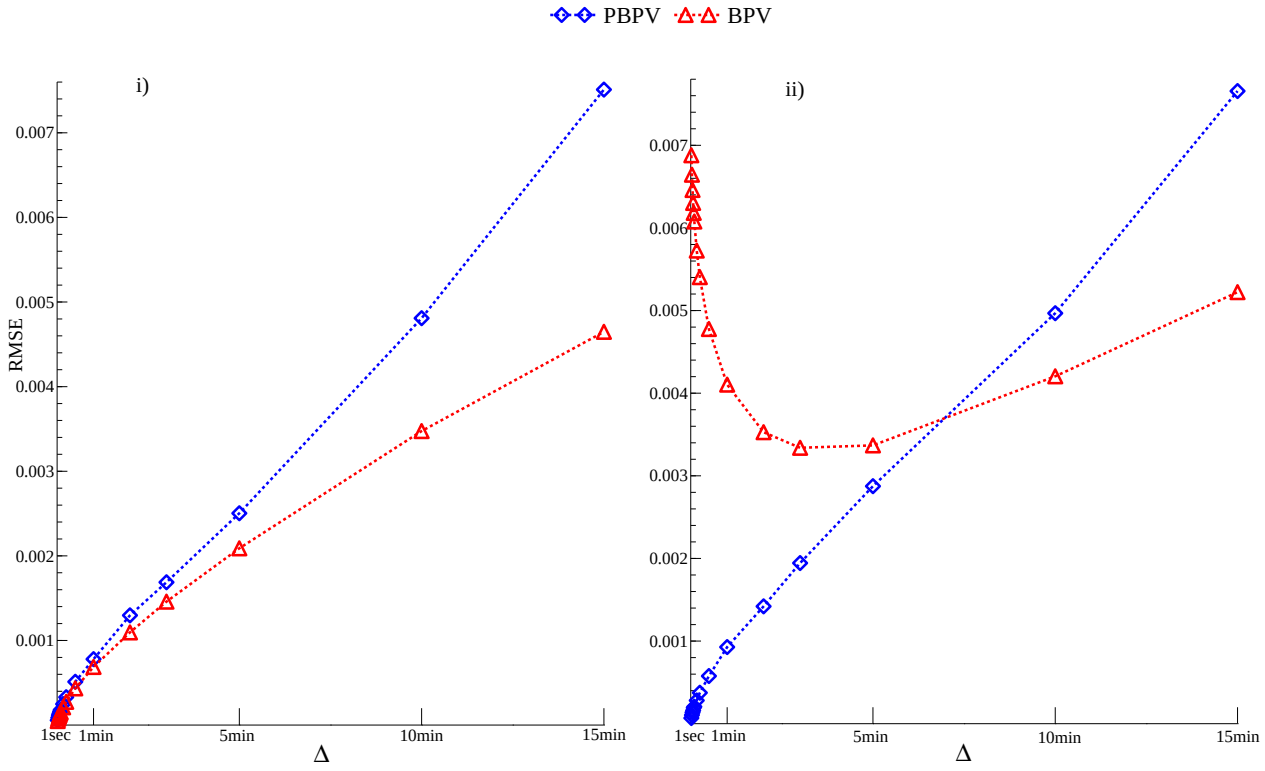
This section considers the spot volatility estimation with regards to microstructure noise, periodic pattern, window width and leverage effect.

5.3.1 Presence of microstructure noise

We focus on the effect of microstructure noise and compare performance of σ_t^2 estimator based on bipower variation (2.16) and pre-averaged bipower variation (2.17). The Monte Carlo parameters are given by $\phi = 0$, $\varpi = 0.001$ (or $\varpi = 0$ for a comparison) and $\rho = 0$.

Figure 2 presents the root mean squared error (RMSE) against the sampling frequency. Without microstructure noise, the best estimation performance for both estimators is obtained at the finest

Figure 2: *Root mean squared error (RMSE) against sampling frequency. Panel i) with no microstructure noise $\varpi = 0$; ii) with microstructure noise $\varpi = 0.001$. Triangles correspond to spot variance estimator based on bipower variation (2.16), diamonds to pre-averaged based estimator (2.17). Heston-B specification with $h_n = n^{0.8}$, $\phi = 0$, $\rho = 0$, $\bar{\kappa} = 0.25$.*



frequency as shown in panel *i)* of Figure 2. The bipower variation based spot variance estimator (2.16) has lower RMSE than the pre-averaged based estimator (2.17) when the sampling frequency increases. With microstructure noise, the (2.16) estimator is not optimal for the finest sampling frequency; see panel *ii)* of Figure 2. The RMSE increases when $\Delta \to 0$ for the bipower variation estimator (2.16). On the other hand, the pre-averaged based estimator (2.17) has the lowest RMSE when data is sampled at the finest frequency. It shows that the correction for bias due to microstructure noise is effective. There is no need to search for an optimal sampling frequency as in the case of bipower variation based estimator (2.16).

5.3.2 Effect of periodic patterns

Here we show the importance of having separate estimation procedures for the spot variance components σ_t and p_t in (2.1). We focus on the impact of periodicity on the joint estimation of $\sigma_t p_t$ based on pre-averaged estimator. The Monte Carlo parameters are given by $\varpi = 0.001$, $\rho = 0$ and $\bar{\kappa} = 0.25$.

Figure 3 presents estimates of $\sigma_t p_t$ obtained by using different window widths determined by β, i.e. $h_n = n^{\beta}$ for $\beta \in (1/2, 1]$. Four different intraday points of time are shown. The top panels of

Figure 3: *Estimates of spot volatility as a function of* β*;* $h_n = n^{\beta}$ *with* $\beta \in (0.5; 1]$*. Crosses correspond to* 2.5%*,* 50% *and* 97.5% *quantiles based on* $M = 2500$*, solid line is the true value. Panels correspond to four different points of time: i) 00:00; ii) 03:00; iii) 09:00; iv) 12:00. Top panels with* $\phi = 0$ *(no diurnal pattern) and bottom panels with* $\phi = 0.95$ *(diurnal pattern). For both,* $\rho = 0$*,* $\bar{\kappa} = 0.25$*,* $\varpi = 0.001$ *and* $\Delta = 1sec$*.*

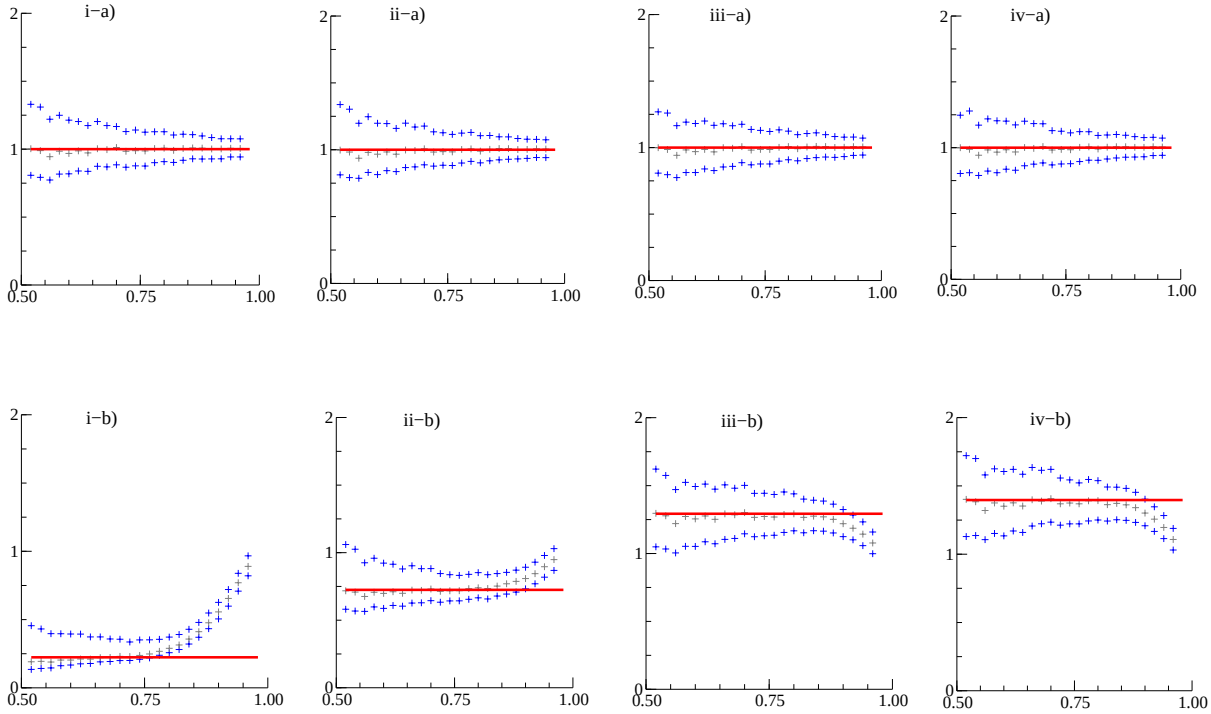


Figure 3 correspond to $\phi = 0$ (no periodicity) and the bottom panels have $\phi = 0.95$. Even though the periodic variance is smooth and data is sampled at the highest frequency $\Delta = 1sec$, presence of diurnal pattern induces a systematic bias and increases when the width of the local window gets larger. Periodic patterns have therefore an adverse effect on the joint estimation of the spot variance. The systematic bias increases when the data is sampled more sparsely and with more pronounced periodic patterns.

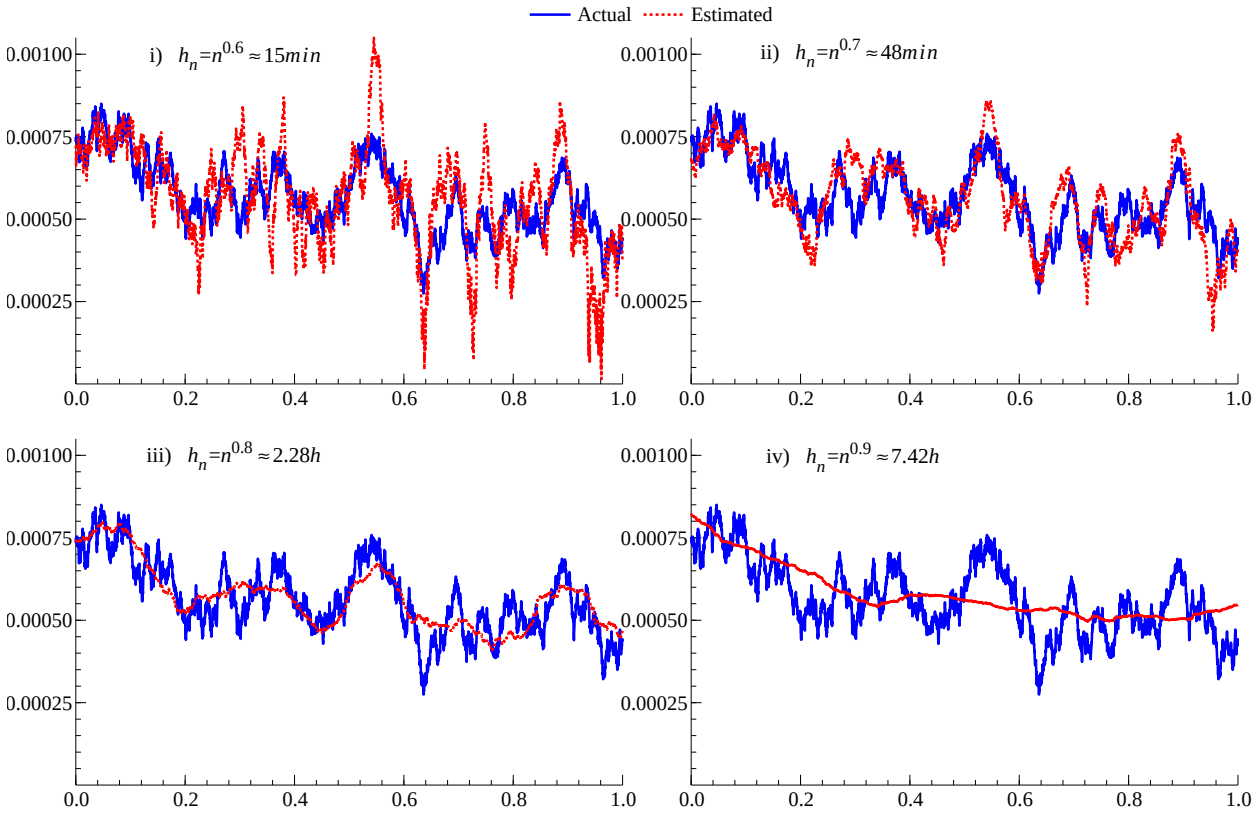
5.3.3 Selection of window width

Next we turn our attention to the estimation of the stochastic component in spot volatility and the choice of the window length in (2.17). We consider $h_n = n^{\beta}$ with $\beta \in \{0.6, 0.7, 0.8, 0.9\}$.

Figure 4 presents estimated paths over a trading day for the Heston-A variance specification. Estimation based on the small window $n^{0.6}$ leads to high variation of the spot variance estimator; see panel *i)* of Figure 4. When the long window $n^{0.9}$ is taken, a highly smooth estimate is obtained, see panel *iv)* of Figure 4.

Table 1a-b presents simulation results for the selection of the optimal bandwidth for Heston-A

Figure 4: *An example of actual (solid line) vs. estimated (dashed line) spot volatility path based on* $h_n = n^{\beta}$. *Heston-A specification with* $\phi = 0$, $\rho = 0$, $\bar{\kappa} = 0.25$, $\varpi = 0.001$, $\Delta = 1sec$.



and Heston-B specifications. We obtain the following findings. For both specifications, the mean squared error (MSE) increases when the sampling is more sparse. For all sampling frequencies, the measures of fit for Heston-A are more satisfactory than the one for Heston-B. The lowest MSEs are obtained for $h_n = n^{0.8}$ (Heston-A) and $h_n = n^{0.7}$ (Heston-B). This result is robust up to the sampling frequency $\Delta = 5min$.

When we allow for a leverage effect with $\rho = -0.75$, we do not find any adverse effect on the estimation performance; see the lowest panels of Tables 1a-b, where we compare the resulting MSEs relative to the case with $\rho = 0$. The obtained MSEs are very much the same with ratios oscillating around unity.

5.4 Intraday jump testing

In this section we investigate the distribution of the test statistic under the null of no jumps in the price process, also, effective power and size are studied.

5.4.1 Null distribution of the jump test statistic

We study the distributions of the jump test statistic based on empirical densities of the maximums of the test statistic under the absence of jumps and compare those to the theoretical standard Gumbel

Table 1a: *The bandwidth selection for the spot variance estimator (2.17); Heston-A specification.*

	$\Delta = 1sec$	$\Delta = 5sec$	$\Delta = 30sec$	$\Delta = 1min$	$\Delta = 5min$
$\rho = 0$					
$\beta = 0.6$	2.066	1.614	2.787	2.046	5.695
$\beta = 0.7$	1.186	1.062	1.522	1.128	1.897
$\beta = 0.8$	**1.000** [0.001]	**1.000** [0.011]	**1.000** [0.113]	**1.000** [0.283]	1.328
$\beta = 0.9$	1.778	1.284	1.161	1.098	**1.000** [2.403]
$\rho = -0.75$					
$\beta = 0.6$	2.107	1.600	2.823	2.052	5.552
$\beta = 0.7$	1.207	1.051	1.535	1.128	1.846
$\beta = 0.8$	**1.000** [0.001]	**1.000** [0.011]	**1.000** [0.112]	**1.000** [0.280]	1.333
$\beta = 0.9$	1.788	1.300	1.174	1.005	**1.000** [2.407]
$\rho = -0.75/\rho = 0$					
$\beta = 0.6$	1.019	0.982	1.001	0.993	0.977
$\beta = 0.7$	1.017	0.988	0.997	0.990	0.975
$\beta = 0.8$	0.999	0.991	0.988	0.990	1.006
$\beta = 0.9$	1.004	1.003	1.001	0.996	1.002

Note: The entries present the mean-squared errors (MSE) relative to the smallest one (between brackets $\times 10^6$) within the model setup, e.g. $\{\Delta = 1sec, \rho = 0\}$. Panel $\rho = -0.75/\rho = 0$ compares MSEs with and without leverage effect. It holds, $\phi = 0$, $\varpi = 0.001$, $\bar{\kappa} = 0.25$. Estimation done over a local window $h_n = n^\beta$ with $\beta = \{0.6; 0.7; 0.8; 0.9\}$.

distribution. We compute empirical density of test statistic (4.1) based on the original $\mathcal{L}(t_i)$ in (4.3), and the extended test statistic based on $\widetilde{\mathcal{L}}(t_i)$ in (4.4).

Figure 5 presents empirical densities of simulated test statistics. To see the impact of microstructure noise, periodicity and leverage effect, we study the empirical distributions under the null separately for each of these features shown in panels *i)-iii)* of Figure 5.

The left-hand panels show the impact of microstructure noise with $\varpi = 0.004$. The empirical distribution of the original LMJ test based on $\mathcal{L}(t_i)$ is affected, as the mass of the density shifts to the left as seen in *i-a)* relative to the theoretical density. Thus, instead of bipower variation (2.16) we apply pre-averaged bipower variation (2.17). Further, the variance of microstructure noise has to be accounted for, because when $\Delta \to 0$, then NSR in (3.1) approaches unity and a larger proportion of return variation is due to microstructure effects. The empirical density of the extended LMJ test based on $\widetilde{\mathcal{L}}(t_i)$ coincides with the theoretical Gumbel density; see *i-b)*.

The impact of diurnal pattern with $\phi = 0.95$ is shown in the middle panels. Intraday periodicity has an adverse effect on the LMJ test as seen in *ii-a)*, where it overrejects the null of no jump. This

Table 1b: *Continued. Heston-B specification.*

	$\Delta = 1sec$	$\Delta = 5sec$	$\Delta = 30sec$	$\Delta = 1min$	$\Delta = 5min$
$\rho = 0$					
$\beta = 0.6$	1.134	1.029	1.382	1.241	2.592
$\beta = 0.7$	**1.000** [0.003]	**1.000** [0.021]	**1.000** [0.256]	**1.000** [0.546]	1.119
$\beta = 0.8$	1.777	1.644	1.177	1.311	**1.000** [5.743]
$\beta = 0.9$	4.376	3.125	1.834	1.891	1.107
$\rho = -0.75$					
$\beta = 0.6$	1.139	1.025	1.389	1.237	2.588
$\beta = 0.7$	**1.000** [0.003]	**1.000** [0.021]	**1.000** [0.259]	**1.000** [0.550]	1.120
$\beta = 0.8$	1.768	1.649	1.176	1.320	**1.000** [5.732]
$\beta = 0.9$	4.371	3.122	1.849	1.903	1.122
$\rho = -0.75/\rho = 0$					
$\beta = 0.6$	1.008	0.997	1.009	1.005	0.997
$\beta = 0.7$	1.004	1.000	0.997	1.009	0.999
$\beta = 0.8$	0.998	1.004	0.996	1.016	0.998
$\beta = 0.9$	1.002	0.999	1.004	1.015	1.011

Note: See Table 1a.

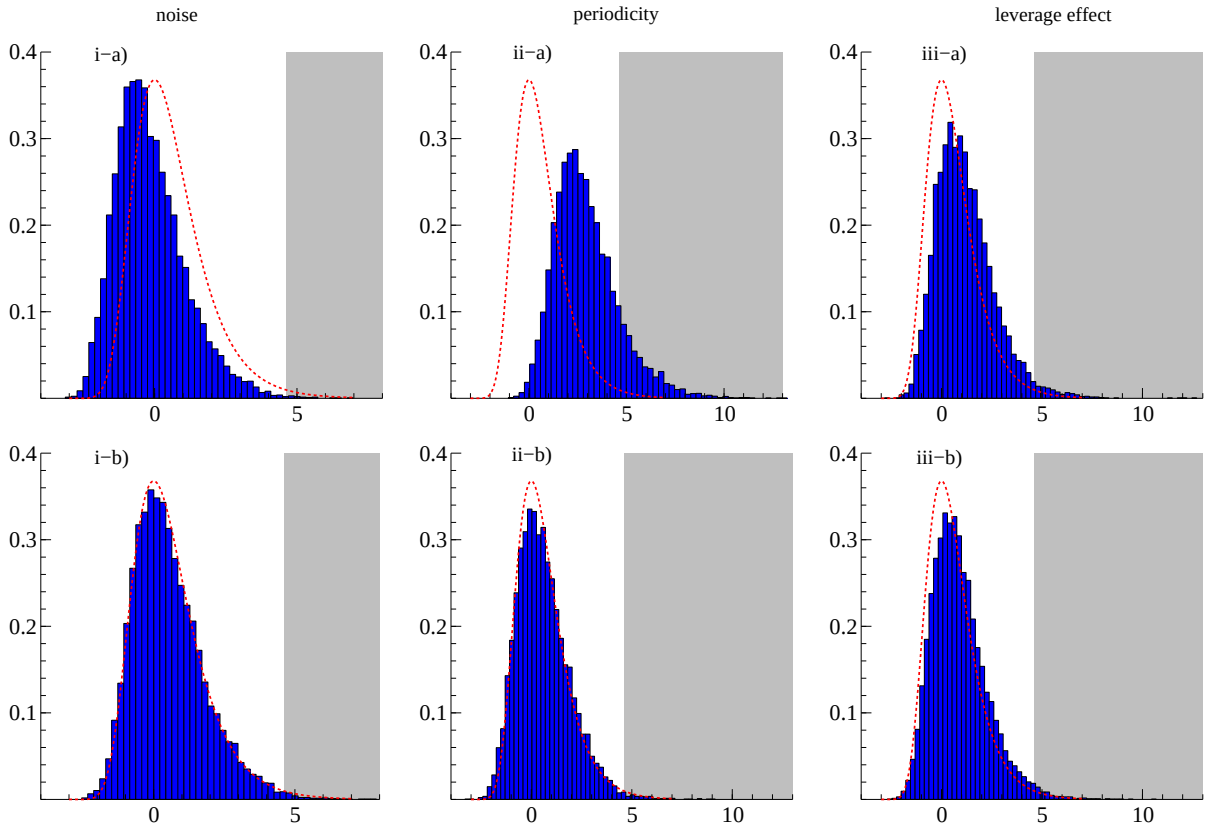
effect was noticed before by Boudt et al. (2008). Pronounced diurnal patterns lead to systematic bias in the spot variance estimator, unless a very small local window is selected at the cost of lack of robustness against jumps and increased variation of the estimator; see the discussion in Section 5.3.2. We estimate the spot variance components separately and panel *ii-b)* shows the appropriate empirical density of extended test based on $\widetilde{\mathcal{L}}(t_i)$.

We allow for the leverage effect with $\rho = -0.75$ and present the resulting empirical densities in the right-hand panels. The negative correlation in price-variance innovations slightly shifts the mass of the empirical densities to the right. It is more apparent in the original LMJ test based on $\mathcal{L}(t_i)$ shown in panel *iii-a)*. Both tests can possibly overreject the null of no jump due to the leverage.

5.4.2 Effective power and size

This section investigates effective power and size of the extended LMJ statistic (4.1) based on $\widetilde{\mathcal{L}}(t_i)$ in (4.4). The effective power shows how many jumps are correctly identified, that is 1 minus the frequency of failure to detect an actual jump. On the contrary, the effective size is interpreted as the probability at each t_i of a spurious jump detection that asymptotically equals $\alpha\% \times \Delta$. The more volatile the volatility process is, the more difficult jump detection becomes, see Lee and Mykland (2008). Our

Figure 5: *Theoretical density of the jump test statistic (dotted line) versus empirical density under the null of no jump. Panels a) for LMJ test (4.1) based on original $\mathcal{L}(t_i)$ in (4.3); panels b) for extended LMJ test based on $\widetilde{\mathcal{L}}(t_i)$ in (4.4). Heston-A specification with $\Delta = 1sec$ and $M = 2500$. Left panels present the effect of noise with $\varpi = 0.004$ ($\phi = 0$, $\rho = 0$); middle panels present the effect of periodicity with $\phi = 0.95$ ($\varpi = 0$, $\rho = 0$); right panels present the effect of leverage with $\rho = -0.75$ ($\varpi = 0$, $\phi = 0$). Grey area is the rejection region at the $\alpha = 1\%$ significance level; $\alpha^* = 4.6001$.*



focus here centers on how sparse sampling, microstructure noise and leverage affect detection of jumps. For brevity, we confine ourselves to the Heston-A specification.

Table 2a-b presents simulation results for the effective power of applying the test to each individual price increment. As price observations are sampled more sparsely and Δ increases, the power to detect smaller-sized jumps decreases rapidly. For instance, for $\varpi = 0.001$ and $\rho = 0$, the extended test detects 74.38% of all jumps with size $\bar{\kappa} = 0.05$ at $\Delta = 1sec$, while only 0.06% at $\Delta = 5min$. The jumps from $\bar{\kappa} = 0.5$ are well detected at $\Delta = 5min$. This is not surprising given that such jumps contribute around 11% to total price variation. Additional microstructure noise has an adverse effect on detecting small jumps. When ϖ increases from 0.001 to 0.004, only 0.02% of all jumps are detected instead of the original 74.38%. In case $\varpi = 0.004$, the smallest jumps are not distinguishable from the innovations coming from the continuous part of the process. For the smaller jumps from $\bar{\kappa} = 0.05$ and $\bar{\kappa} = 0.15$, and large noise power increases initially with increase of Δ as microstructure noise is mitigated. Later it decreases again due to sparse sampling.

Table 2a: *Effective power and daily effective size of extended LMJ test. Heston-A;* $\varpi = 0.001$.

	$\Delta = 1sec$	$\Delta = 5sec$	$\Delta = 30sec$	$\Delta = 1min$	$\Delta = 5min$
$\rho = 0$					
$\bar{\kappa} = 0.05$	0.7438	0.3651	0.0209	0.0044	0.0006
	[0.009]	[0.009]	[0.014]	[0.012]	[0.038]
$\bar{\kappa} = 0.15$	1.0000	1.0000	0.9761	0.7273	0.0761
	[0.006]	[0.009]	[0.013]	[0.013]	[0.028]
$\bar{\kappa} = 0.25$	1.0000	1.0000	0.9998	0.9993	0.4857
	[0.007]	[0.009]	[0.009]	[0.014]	[0.027]
$\bar{\kappa} = 0.5$	1.0000	1.0000	0.9998	0.9998	0.9951
	[0.010]	[0.007]	[0.014]	[0.015]	[0.020]
$\rho = -0.75$					
$\bar{\kappa} = 0.05$	0.7463	0.3647	0.0232	0.0053	0.0004
	[0.009]	[0.009]	[0.016]	[0.014]	[0.038]
$\bar{\kappa} = 0.15$	1.0000	1.0000	0.9759	0.7263	0.0774
	[0.006]	[0.009]	[0.014]	[0.014]	[0.028]
$\bar{\kappa} = 0.25$	1.0000	1.0000	0.9998	0.9993	0.4951
	[0.006]	[0.008]	[0.009]	[0.015]	[0.025]
$\bar{\kappa} = 0.5$	1.0000	1.0000	0.9998	0.9998	0.9960
	[0.012]	[0.006]	[0.015]	[0.014]	[0.019]

Note: Significance level $\alpha = 1\%$, thus $\alpha^* = 4.6001$. Effective power equals 1 minus frequency of failure to detect an actual simulated jump. Effective daily size between square brackets equals probability of detecting a spurious jump times number of observations per day at the sampling frequency. We set $h_n = n^{0.8}$, $\phi = 0.95$ and $M = 2500$.

Table 2a-b also reports the daily effective size, defined as frequency of spuriously detected jumps divided by Δ. As long as the data is sampled at fine intervals, the increase of microstructure noise does not affect the effective size of extended test statistic. When we lower the sampling frequency, the estimate of the stochastic component in spot volatility becomes too smooth to capture high frequency volatility movements. This leads to both power and size distortions.

We do not find a noticeable impact of leverage on the effective power and size of the extended test statistic based on $\widetilde{\mathcal{L}}(t_i)$ for data sampled at high frequencies. When Δ increases, the size is however more distorted. For instance, for $\Delta = 30sec$ and $\rho = 0$, the empirical daily size of a test equals 0.014 instead of the theoretical $\alpha = 0.01$, while it increases to 0.016 when $\rho = -0.75$. In terms of the effective size, it means an increase of the probability of a spurious jump detection at each tested time t_i from 4.86 $\times 10^{-6}$ to 5.56 $\times 10^{-6}$, instead of the theoretical $\alpha \times \Delta$ of 3.47 $\times 10^{-6}$.

Table 2b: *Continued. Heston-A;* $\varpi = 0.004$.

	$\Delta = 1sec$	$\Delta = 5sec$	$\Delta = 30sec$	$\Delta = 1min$	$\Delta = 5min$
$\rho = 0$					
$\bar{\kappa} = 0.05$	0.0002 [0.0104]	0.0001 [0.006]	0.0014 [0.006]	0.0004 [0.007]	0.0006 [0.064]
$\bar{\kappa} = 0.15$	0.4190 [0.005]	0.4569 [0.006]	0.3179 [0.006]	0.1890 [0.009]	0.0581 [0.052]
$\bar{\kappa} = 0.25$	0.9779 [0.008]	0.9821 [0.006]	0.9636 [0.005]	0.9156 [0.005]	0.4020 [0.047]
$\bar{\kappa} = 0.5$	1.0000 [0.009]	1.0000 [0.006]	0.9998 [0.006]	0.9998 [0.004]	0.9907 [0.045]
$\rho = -0.75$					
$\bar{\kappa} = 0.05$	0.0002 [0.0104]	0.0001 [0.006]	0.0014 [0.006]	0.0003 [0.009]	0.0005 [0.069]
$\bar{\kappa} = 0.15$	0.4074 [0.005]	0.4604 [0.006]	0.3198 [0.005]	0.1896 [0.008]	0.0522 [0.060]
$\bar{\kappa} = 0.25$	0.9705 [0.007]	0.9843 [0.007]	0.9666 [0.004]	0.9130 [0.005]	0.4092 [0.054]
$\bar{\kappa} = 0.5$	1.0000 [0.009]	1.0000 [0.006]	0.9998 [0.006]	0.9998 [0.004]	0.9924 [0.054]

Note: See Table 2a.

6 Testing for Jumps Empirically

6.1 EUR/USD exchange rate data

In our empirical study we analyze the EUR/USD exchange rate data from July 2 to December 31, 2007.[1] The data time span is six months with $D = 131$ working days. The transaction price data at the sampling frequencies of 1, 5, and 15 minutes is provided by Disk Trading.[2] The percentage returns of the three price series are computed by taking their first differences in logs and multiplying by 100.

Table 3: *Descriptive statistics of the EUR/USD returns.*

Δ	mean	std.dev.	skewness	kurtosis	ACF_1	#obs
$1min$	3.60e-05	0.014	-0.083	16.954	-0.176 [±0.0046]	187816
$5min$	1.84e-04	0.027	0.212	22.045	-0.081 [±0.0103]	37670
$15min$	5.44e-04	0.044	0.234	14.366	-0.043 [±0.0178]	12571

Note: ACF_1 is the first order autocorrelation of raw returns; associated critical range is between square brackets.

[1] Our empirical study was also partially conducted on GBP/USD, JPY/USD and CHF/USD exchange rates. The results are relatively similar, thus for brevity only results for most heavily traded EUR/USD rate are presented.

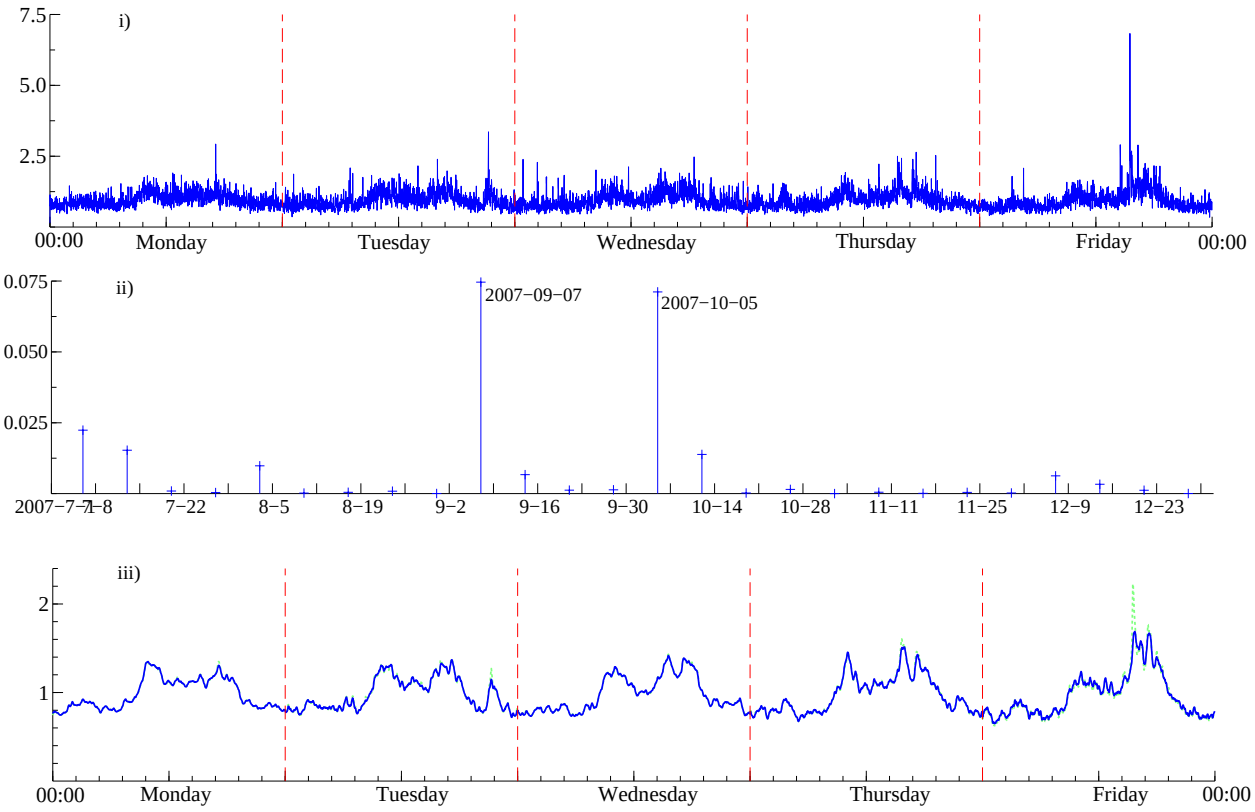
[2] Source: `http://disktrading.is99.com`

Table 3 presents the descriptive statistics of computed returns. The excess kurtosis of the return series indicates the presence of time-varying volatility or jumps. The first order autocorrelation of raw returns is negative and significant at all frequencies. This finding is a typical feature of price series contaminated with microstructure noise.

6.2 Estimation of diurnal volatility patterns

Figure 6 presents diurnal shapes estimated with the benchmark method (Section 3.3) and our robust procedure (Section 3.4). Panel *i)* of Figure 6 presents the benchmark $\widehat{p}_{t_i}$ from (3.4) at $\Delta = 1min$ sampling frequency. Diurnal shapes are relatively similar for different days of the week. The main differences arise in the time period of 14.00-17.00, when markets in the US and Europe overlap. The numerous peaks in the return series correspond to various scheduled news announcements. European interest rate reports and macroeconomic releases appear on Tuesday and Wednesday whereas US macroeconomic news and interest rates decisions are released on Tuesday, Thursday and Friday.

Figure 6: *Periodicity extraction for the EUR/USD rate; returns sampled at* $\Delta = 1min$ *frequency. Panel i) presents square root of cross-sectional variance (3.4); ii) presents absolute price increments of consecutive Fridays in the sample in time interval 15:00-15.31; iii) presents estimates of square root of a periodic factor (3.6) based on OWM-N (dotted line) and OWM-t (solid line).*



There are numerous outliers in the benchmark $\widehat{p}_{t_i}$. To study whether these are caused by single distortions or are a part of the diurnal shape, panel *ii)* of Figure 6 presents the absolute price increments

in the time interval 15.30-15.31 of all Fridays. Two large return values appear on September, 7 and October, 5, which are the first Fridays of these months. Both days correspond to news surprises on payroll data from the US labor market survey.[3] At this particular intraday time, $\widehat{p}_{t_i}$ is affected by those two large price increments. To be robust against such news surprises, the estimation method of Section 3.4 is used. Panel *iii)* of Figure 6 presents the estimated diurnal patterns based on weighting as in (3.6) with the use of observation disturbances from Gaussian (OWM-N) and Student's t (OWN-t) densities. The resulting patterns are much smoother and clearly display repetitive low and high intraday volatility periods.

Table 4: *Maximum likelihood estimates of the observation weight method.*

	OWM-N		OWM-t		
	σ_ζ^2	σ_η^2	σ_ζ^2	σ_η^2	ν
Monday	0.0376	0.0003	0.0377	0.0003	14.1876
	(0.0015)	(0.0001)	(0.0017)	(0.0001)	(5.7772)
Tuesday	0.0398	0.0007	0.0404	0.0005	7.4998
	(0.0017)	(0.0002)	(0.0022)	(0.0002)	(1.4718)
Wednesday	0.0403	0.0004	0.0403	0.0004	12.2512
	(0.0016)	(0.0001)	(0.0018)	(0.0001)	(3.6748)
Thursday	0.0378	0.0007	0.0377	0.0006	6.5195
	(0.0015)	(0.0002)	(0.0021)	(0.0001)	(1.0686)
Friday	0.0524	0.0021	0.0437	0.0008	4.2041
	(0.0023)	(0.0005)	(0.0036)	(0.0002)	(0.5175)

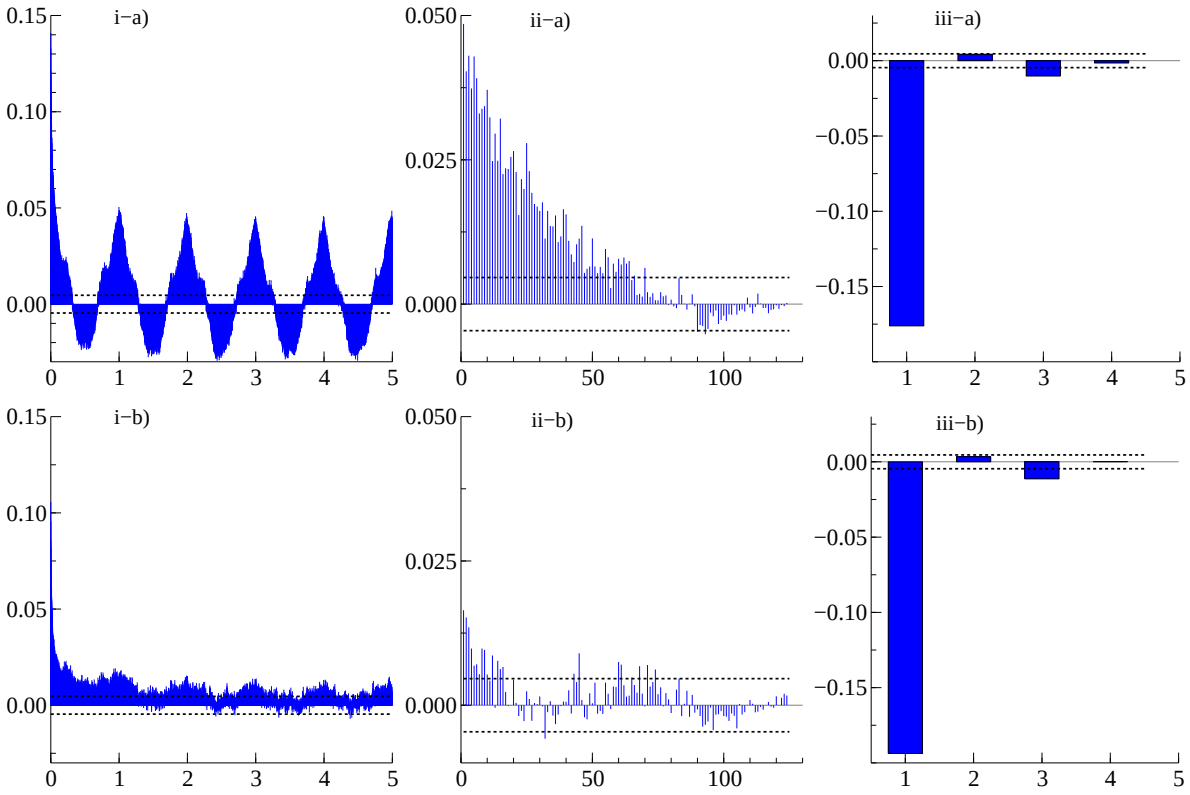
Note: Estimates of parameters of the periodicity model (3.7)-(3.8) for the EUR/USD rate at $\Delta = 1min$ sampling frequency. Standard errors between the brackets.

Table 4 presents the likelihood-based estimation results of the parameters in the periodicity model (3.7)-(3.8). The largest outliers in this sample are on Thursday and Friday, which is reflected by low values for the degree-of-freedom parameter in the Student's t distribution ($\nu = 6.52$ and $\nu = 4.20$ respectively). Also, the excess kurtosis, $\frac{6}{\nu-4}$, is remarkably high for these days, especially on Friday (≈ 29.4). Given the low value of the estimates of ν and numerous outliers in the benchmark $\widehat{p}_{t_i}$, the robust method with Student's t observation disturbances is preferred.

Panels *i-a)* $-$ *iii-a)* in Figure 7 present the ACFs for absolute and raw returns at $\Delta = 1min$. The U-shape pattern recurring at a daily frequency in the autocorrelations of absolute returns can be seen in panel *i-a)*. Due to periodicity, the ACF of absolute returns is persistent from day to day as shown in *ii-a)*. Panels *i-b)* $-$ *iii-b)* present the ACFs after the diurnal pattern (robustly estimated) is

[3]The non-farm payrolls, released by the Bureau of Labor Statistics of the US Department of Labor, measure the number of people on the payrolls of all non-agricultural businesses. At the FX market a high reading is seen as positive for the US dollar, while a low reading is seen as negative. On September 7, the actual value of nonfarm payrolls (for August) was 93k, way below expected by market averaged at 108k. On October 5, the actual value of nonfarm payrolls (for September) was 96k, below expected by market averaged at 100k. According to the economic calendar data at `http://www.fxstreet.com/` these came as market surprises.

Figure 7: *Autocorrelation functions of EUR/USD returns. Panels i) shows ACF of absolute returns displayed for 5 days; ii) ACF of absolute returns as in panel i) now showing only values at daily lags; iii) ACF of raw returns for 4 lags. Panels a) raw returns; b) returns adjusted for variance diurnal shape using OWM-t.*



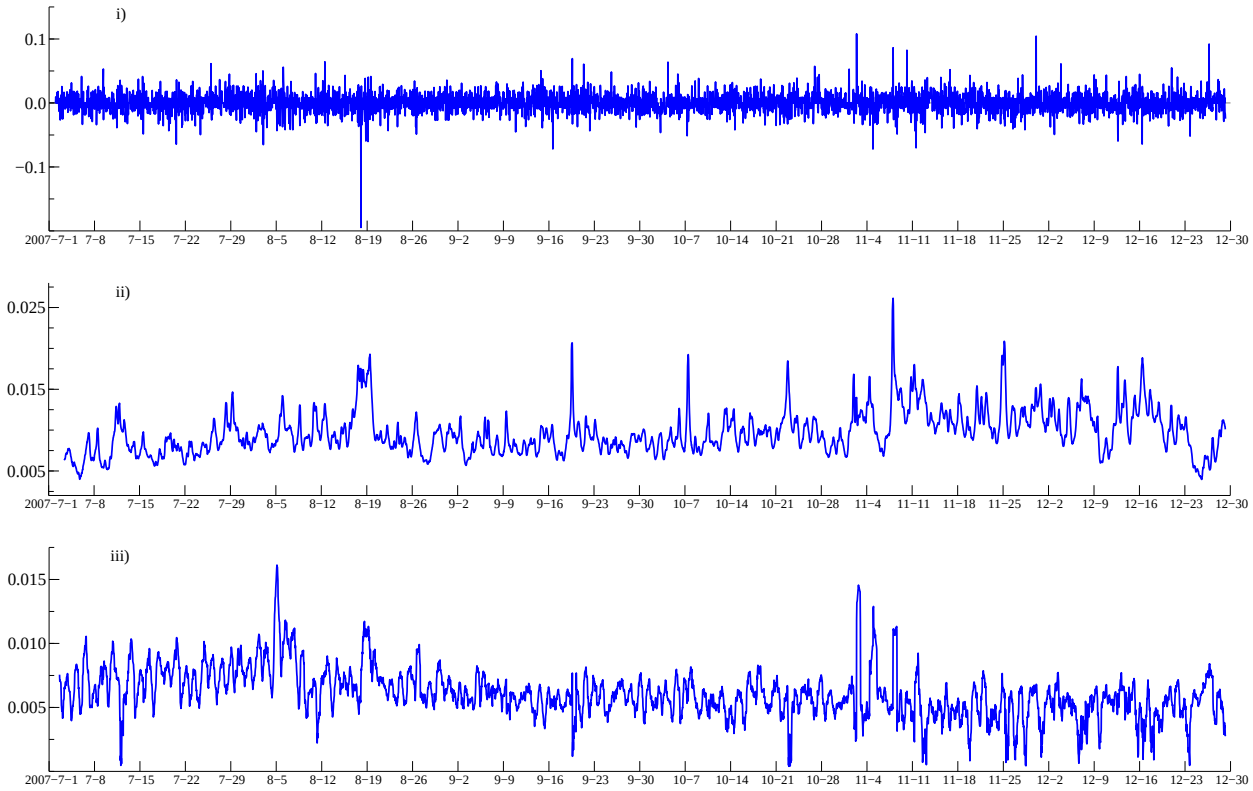
removed. Periodicity is not perfectly extracted, but the ACF of absolute returns shown in *i-b)* has no clear recurring U-shaped cycles as comparing *i-b)* to *i-a)*. Moreover, the ACF is not that persistent, comparing *ii-b)* to *ii-a)*. The significant negative first order serial correlation seen in *iii-a)* and *iii-b)* is due to microstructure noise, and is not affected by removing the diurnal shape.

6.3 Spot measures

Figure 8 presents the returns (after robust adjustment for periodicity) and the estimated path of the stochastic component in the spot volatility based on the pre-averaged estimator (2.17) at $\Delta = 1min$. Estimates of σ_t^2 are based on averaging over a set of local windows with $h_n = n^\beta$ with $\beta \in [0.52; 0.54, \ldots, 1]$. In panel *i)* of Figure 8, several return clusters (e.g., in August or November) can be recognized. These clusters correspond to the increased level of volatility $\widehat{\sigma}_{d+t_i}$; see panel *ii)* of Figure 8. Overall, the level of spot volatility increases and is possibly a sign of the start of the financial crisis in 2008. Estimated spot volatility reaches its maximum at the beginning of November 2008.

The microstructure noise variance estimator of (2.11) is based on returns over one day. Panel *iii)* of Figure 8 displays 24-hour rolling window estimates of $\widehat{\varpi}_{d+t_i}$. It is found that the microstructure

Figure 8: *Estimated spot measures. Panel i) EUR/USD periodically adjusted returns at* $\Delta = 1min$ *frequency; ii) estimated stochastic component* $\widehat{\sigma}_{d+t_i}$ *(times* $\sqrt{\Delta}$*) of the spot volatility path; iii) estimated time-varying volatility of microstructure noise* $\widehat{\varpi}_{d+t_i}$*.*



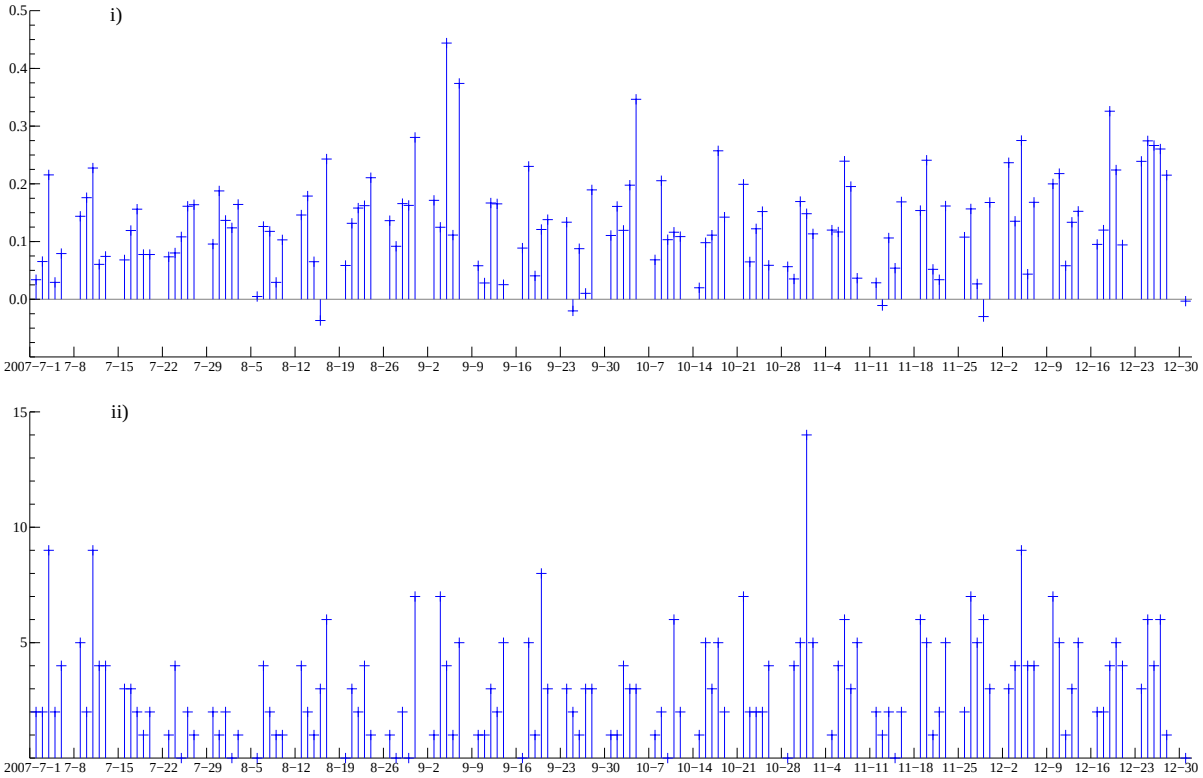
noise standard deviation is relatively high at the beginning of the sample, while lower towards the end of the sample. As a result, the estimated noise-to-signal ratio in (3.1) also varies considerably over time. Its value (not displayed) starts from around 80% at the beginning of the sample period, decreasing to 40-50% at the end of the sample.

6.4 Aggregate evidence of jumps

We first report daily estimates $\widehat{\text{PBPV}}_d(2,0)$ and $\widehat{\text{PBPV}}_d(1,1)$ for $d = 1, \ldots, 131$. The relative jump variation estimate $\widehat{\text{RJV}}_d$ indicates the presence of a jump component in the price process. Given the sampling frequency $\Delta = 1min$, the average of $\widehat{\text{RJV}}_d$ series equals 13% (16% and 22% at $\Delta = 5min$ and $\Delta = 15min$, respectively). Mean estimates of RJV increase with an increasing Δ, but the accuracies of IV and QV estimates decrease. Using realized measures that are not robust to microstructure noise, Huang and Tauchen (2005) find that jumps contribute in around 7% in the S&P index at $\Delta = 5min$. Barndorff-Nielsen and Shephard (2006) find that jumps contribute between 5% and 22% to total price variation dependently on the exchange rate at frequencies ranging from $\Delta = 5min$ to $\Delta = 1h$.

Figure 9 presents the sequence of RJV estimates and the number of significant jumps over the sample period at $\Delta = 1min$. There are numerous days when $\widehat{\text{RJV}}_d$ exceeds 25-30% as well as when

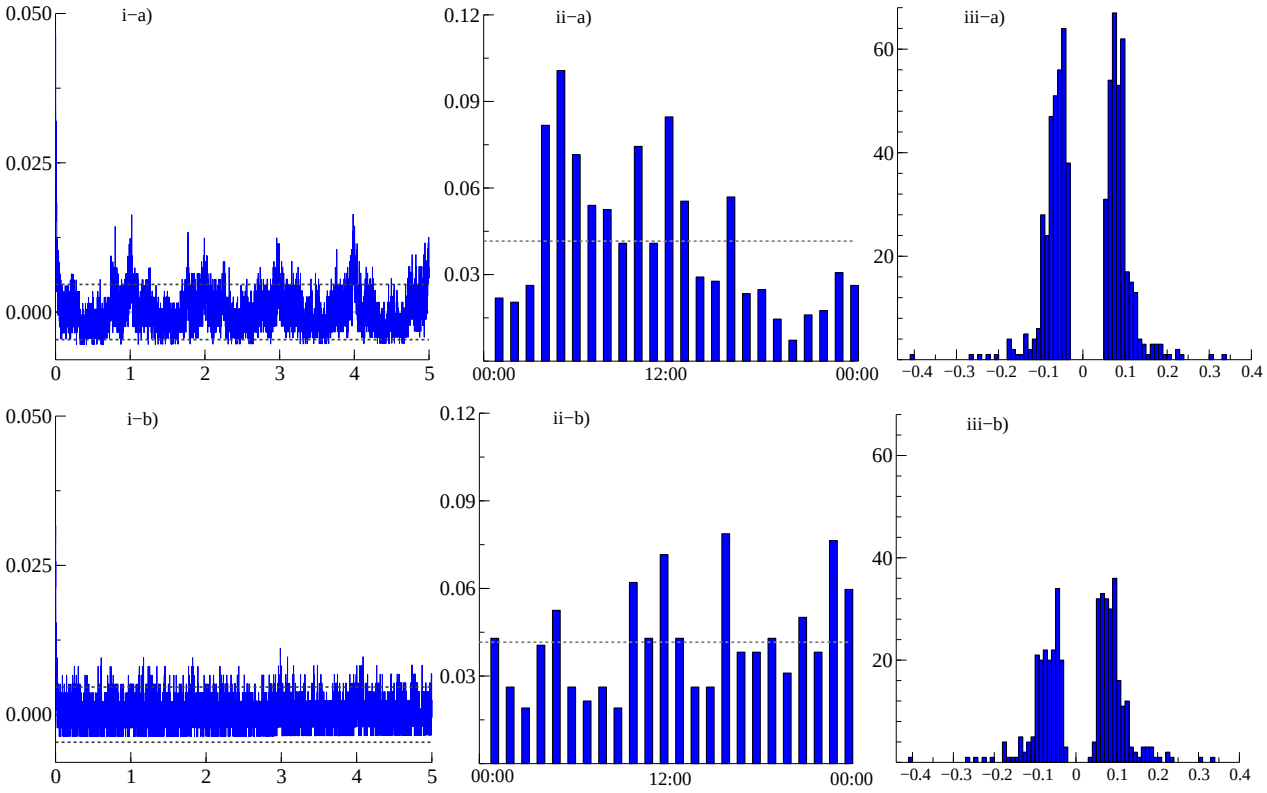
Figure 9: *Relative jump variation (2.13) in panel i) and the number of significant jumps in panel ii) at* $\Delta = 1min$*. Significant jumps based on the LMJ test (4.1) with* $\widetilde{\mathcal{L}}(t_i)$ *in (4.4) at the* $\alpha = 1\%$ *significance level.*



it is roughly zero. Using the extended LMJ statistic we find more jumps at the higher frequency. At $\Delta = 1min$, in total 419 jumps are detected at $\alpha = 1\%$, so that 3 jumps occur on average per trading day. When $\Delta = 5min$ 132 jumps are detected, which means that 1 jump occurs on average per day. Finally, at $\Delta = 15min$ sampling frequency only 45 jumps are detected, which results in 1 jump on average every 3 days. We find a maximum of 2 jumps (July 18, September 5, October 5 and December 27) at $\Delta = 15min$, while we find a maximum of 14 jumps (November 1) at $\Delta = 1min$. The significant jumps at $\Delta = 1min$ have an average absolute size of 8.4 basis points (bps), while at $\Delta = 5min$ ($\Delta = 15min$) the average size is of 15.6bps (24.4bps).

Figure 10 presents the dependence structure, timing and histogram of the detected jumps. We compare the LMJ test statistic based on (4.3) and on the extended statistic (4.4). Since the original LMJ test does not account for diurnal patterns, the autocorrelation structure indicates the periodical dependence of the detected jumps; see the autocorrelation clusters at daily time intervals in *i-a)* of Figure 10. The LMJ statistic detects more (less) jumps when deterministic volatility is higher (lower); see *ii-a)* of Figure 10. It causes spurious time dependence of jumps. Jumps detected with the extended statistic are not periodically dependent and are more equally distributed over intraday time as seen in *i-b)* and *ii-b)*, respectively, of Figure 10. Over this sample period, there were more positive jumps than negative as seen in *iii-b)* of Figure 10. Given that the US dollar was depreciating in those months, this

Figure 10: *Autocorrelation function of jumps indicators in panel i) for 5 days; timing aggregated to 1 hour in panel ii); histogram of significant jumps in panel iii), at* $\Delta = 1min$. *Panels a) for original LMJ test based on* $\mathcal{L}(t_i)$ *in (4.3); b) for the extended LMJ test based on* $\widetilde{\mathcal{L}}(t_i)$ *in (4.4).*



result is not surprising. The extended test statistic detects more smaller-sized jumps than the original test; see the histograms in panel *iii-b)* to *iii-a)* of Figure 10. The original test detects more jumps but in the Monte Carlo study it is shown that the presence of diurnal patterns leads to an overrejection of the null hypothesis of no jumps. Too many jumps are spuriously detected due to the periodic increase of return volatility that is not captured when we estimate the two spot variance components jointly.

6.5 Jumps at specific days

The LMJ jump test statistic yields the timing, size and direction of the detected jumps. We study in some detail the outcome of applying the extended test for two specific days. On the one hand, we select Wednesday, September 5, when $\widehat{\text{RJV}}_{2007/09/05}$ reaches its maximum of 44.4% at $\Delta = 1min$, see Figure 9. On the other hand, we choose Thursday, November 1, when a maximum number of 14 jumps is found at $\Delta = 1min$, while $\widehat{\text{RJV}}_{2007/11/01}$ equals 14.8% at this frequency. Figures 11a and 11b zoom in on both days.

Selected days in this case study correspond to scheduled macroeconomic announcements relevant for the EUR/USD exchange rate, these are listed in Table 5. On September 5 there are 4 jumps detected at $\Delta = 1min$ frequency at 00:44, 07:38, 15:18 and 17:12. At the 5% significance level, three

Figure 11a: *Case study for September 5, 2007. Panels i) present the EUR/USD exchange rate in level; ii) EUR/USD returns and significant jumps based on the extended LMJ test with* $\widetilde{\mathcal{L}}(t_i)$ *in (4.4) (at* $\alpha = 1\%$ *marked with circles; at* $\alpha = 5\%$ *marked with diamonds).*

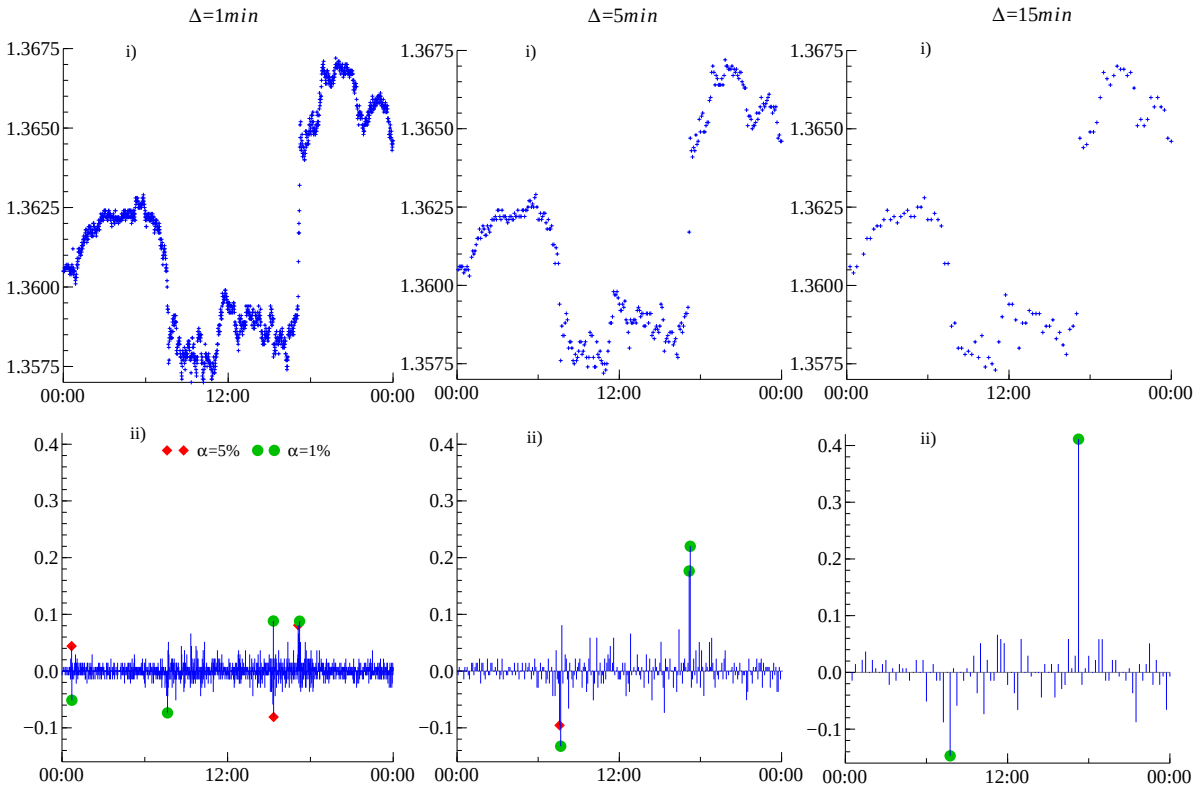
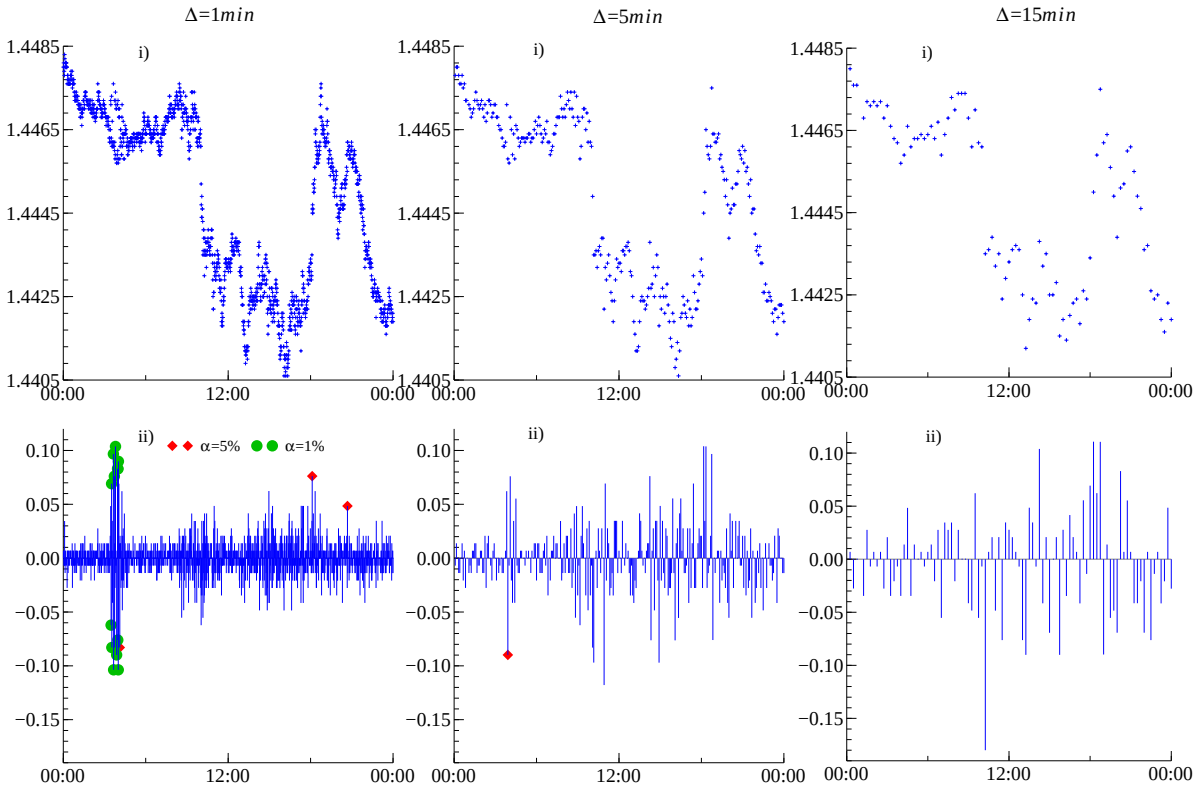


Figure 11b: *Continued. November 1, 2007.*



additional jumps are found at 00:43, 15:19 and 17:10; see panel *ii)* of Figure 11a. These time intervals correspond to high intraday volatility time periods, when markets open or overlap. The morning European news items were positive for the EUR/USD rate on that day. It is seen from Figure 11a that the EUR/USD rate started to appreciate dramatically in the afternoon when the US readings, presented in Table 5, were much worse than expected. The sizes (in absolute terms) of the detected jumps were from 7 to 10bps (or from 6bps at $\alpha = 5\%$), which corresponds to around 0.07-0.1% change of the fundamental within a $1min$ interval. When lowering the sampling frequency to $\Delta = 15min$ only 2 jumps are found to be significant, at 15:18 (-15bps) and at 17:12 (42bps) contributing around 51.6% to the total price variation as measured by the estimate of RJV.

Table 5: *News releases.*

September 5, 2007						
Time	Area	Event	Actual	Consensus	Previous	+/-
08:55	Germany	Purchasing Manager Index Services (Aug)	59.8	58.1	58.5	+
09:00	EMU	Purchasing Manager Index Services (Aug)	58.0	57.9	58.3	
10:00	EMU	Retail Sales (MoM) (Jul)	0.4%	0.3%	0.6%	+
10:00	EMU	Retail Sales (YoY) (Jul)	1.3%	1.1%	1.0%	+
13:15	US	ADP Employment Change (Aug)	27k	80k	48k	+
15:00	US	Pending Home Sales (MoM) (Jul)	-12.2%	-2.0%	5.0%	+
19:00	US	Fed's Beige Book				

November 1, 2007						
Time	Area	Event	Actual	Consensus	Previous	+/-
13:30	US	Core PCE - Prices Index (MoM) (Sep)	0.2%	0.2%	0.1%	
13:30	US	Core PCE - Prices Index (YoY) (Sep)	1.9%	1.8%	1.8%	-
13:30	US	Initial Jobless Claims (Oct 27)	330k	330k	331k	
13:30	US	Personal Income (MoM) (Sep)	0.4%	0.4%	0.4%	
13:30	US	Personal Spending (Sep)	0.3	0.4	0.6	
15:00	US	ISM Manufacturing (Oct)	50.9	51.4	52.0	+
15:00	US	ISM Prices Paid (Oct)	63	64	59	

Note: Important news releases on September 5, and November 1, 2007. The + (-) denotes a positive (negative) reading for the EUR/USD exchange rate. EMU - European Monetary Union, PCE - Personal Consumption Expenditure.
Source: `http://www.fxstreet.com/`.

On November 1 there are 14 (17) jumps found at $\Delta = 1min$ frequency at the 1% (5%) significance level. Interestingly, 12 of these jumps cluster in the time period between 03:28-04:06 when only Asian markets operate as seen in panel *ii)* of Figure 11b. These jumps have an absolute size of around 10bps. We have no detailed information what happened in this time interval, but on this particular Thursday many US news announcements were to be released, see Table 5, which possibly caused some additional

arrival of orders. At $\Delta = 1min$, $\widehat{\text{RJV}}_{2007/11/01}$ equals 14.8%, which indicates the presence of a jump component. When lowering the sampling frequency to $\Delta = 15min$, we find no significant jumps even at the $\alpha = 5\%$ significance level, although the estimate of RJV equals 17.7% at this frequency and the largest absolute price increment is of 18bps. It indicates that even when estimates of QV and IV differ considerably (and hence indicate the occurence of jumps) it may not be possible to point out which price increments are caused by jumps when the sampling frequency is low.

7 Summary and Conclusions

In this paper we introduce a new approach to the estimation of spot variance from high frequency data. We decompose spot variance into stochastic and deterministic components. Estimation of the stochastic variance component is based on adopting a pre-averaged bipower variation measure. The pre-averaging approach allows us to estimate the stochastic component in the jump-diffusion framework with microstructure noise. As a result, we can apply the estimator without the need for mitigating noise effects through sparse sampling. Estimation of the deterministic variance component is based on the newly developed semi-parametric approach. We smooth the cross-sectional variances using a robust weighting scheme. In a finite sample Monte Carlo study we show that the underlying spot variance path can be estimated accurately as long as the price observations are sampled at a sufficiently high frequency, indeed even in practical situations where microstructure noise and jumps are present or leverage effects confound the variance process.

Particular attention is given to the testing for jumps since our framework with microstructure noise offers an extension to the jump test statistic introduced by Lee and Mykland (2008) and modified later by Boudt et al. (2008). The Monte Carlo study has shown that the distribution of our extended test statistic is not distorted by periodic patterns, leverage effects or microstructure noise. Also, our extended jump test improves on the detection of smaller-sized jumps, which, however, also depends on the magnitude of microstructure noise contamination. With large noise, the smallest jumps cannot be distinguished from microstructure frictions, which is a theoretical result and not due to a failure of the testing method. The power of our extended test decreases when the sampling is less frequent but we should emphasize that our framework alleviates the need for sparse sampling.

The empirical illustration focusses on spot volatility and jumps in the EUR/USD exchange rate series for a period of 2007 when the subprime crisis became apparent. In this period the spot volatility has increased slowly while the microstructure noise has become less important in the later months. We provide clear evidence of discontinuities in the exchange rate series. More jumps are detected during days with important macro news announcements and from a dataset that is sampled at a higher frequency. When the periodic volatility factor is extracted and removed, we do not find periodic time dependence in the realized jump series and the number of jumps is also reduced. We therefore conclude that our modifications for estimating spot volatility and jump testing are effective and can be successfully applied in studies of financial markets based on intraday high frequency data.

References

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

Alvarez, A., F. Panloup, M. Pontier, and N. Savy (2008). Estimation of the instantaneous volatility and detection of volatility jumps. http://www.citebase.org/abstract?id=oai:arXiv.org:0812.3538.

Andersen, T. G. (2004). Discussion. *Journal of Financial Econometrics 2*, 37–48.

Andersen, T. G. and L. Benzoni (2009). Realized Volatility. In T. G. Andersen, R. A. Davis, J.-P. Kreiß, and T. Mikosch (Eds.), *Handbook of Financial Time Series*, pp. 555–575. Springer-Verlag, Berlin.

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

Andersen, T. G. and T. Bollerslev (1998). Deutsche mark-dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies. *Journal of Finance 53*, 219–265.

Andersen, T. G., T. Bollerslev, F. X. Diebold, and P. Labys (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association 96*(453), 42–55.

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

Andreou, E. and E. Ghysels (2002). Rolling-sample volatility estimators: Some new theoretical, simulation and empirical results. *American Statistical Association 20*(3), 363–76.

Bandi, F. M. and R. Reno (2009). Nonparametric stochastic volatility. Global COE Hi-Stat Discussion Paper Series gd08-035, Institute of Economic Research, Hitotsubashi University.

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

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

Barndorff-Nielsen, O. E. and N. Shephard (2004). Power and bipower variation with stochastic volatility and jumps (with discussion). *Journal of Financial Econometrics 2*(1), 1–48.

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

Bos, C. S. (2008). Model-based estimation of high frequency jump diffusions with microstructure noise and stochastic volatility. Tinbergen Institute Discussion Papers 2008-011/4, Tinbergen Institute.

Boudt, K., C. Croux, and S. Laurent (2008). Robust estimation of intraweek periodicity in volatility and jump detection. Working Paper, SSRN-id1297371.

Doornik, J. (2006). *Ox: An Object-Oriented Matrix Language*. London: Timberlake Consultants Press.

Duan, J. C. and A. Fülöp (2007). How frequently does the stock price jump? An analysis of high-frequency data with microstructure noises. MNB Working Papers No 2007/4, Magyar Nemzeti Bank.

Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. New York: Oxford University Press.

Fan, J. and Y. Wang (2008). Spot volatility estimation for high-frequency data. *Statistics and Its Interface 1*, 279–288.

Foster, D. P. and D. B. Nelson (1996). Continuous record asymptotics for rolling sample variance estimators. *Econometrica 64*, 139–174.

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

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

Huang, X. and G. Tauchen (2005). The Relative Contribution of Jumps to Total Price Variance. *Journal of Financial Econometrics 3*, 456–499.

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

Jiang, G. J. and R. C. A. Oomen (2008). Testing for jumps when asset prices are observed with noise-a "swap variance" approach. *Journal of Econometrics 127*, 352–370.

Kalnina, I. and O. Linton (2008). Estimating quadratic variation consistently in the presence of endogenous and diurnal measurement error. *Journal of Econometrics 147*(1), 47–59.

Kinnebrock, S. (2008). *Asymptotic results for semimartingales and related processes with econometric applications*. Ph. D. thesis, University of Oxford.

Koopman, S. J., N. Shephard, and J. A. Doornik (1999). Statistical algorithms for models in state space using ssfpack 2.2. *Economic Journal 2*, 107–160.

Kristensen, D. (2009). Nonparametric filtering of the realised spot volatility: A kernel-based approach. Forthcoming in the *Econometric Theory*.

Lahaye, J., S. Laurent, and Neely (2009). Jumps, cojumps and macro announcements. *Federal Reserve Bank of St. Louis*. Working Paper.

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

Mykland, P. A. and L. Zhang (2008). Inference for volatility-type objects and implications for hedging. *Statistics and Its Interface 1*, 255–278.

Ogawa, S. and S. Sanfelici (2008). An improved two-step regularization scheme for spot volatility estimation. Economics department working papers, Department of Economics, Parma University (Italy).

Oomen, R. C. A. (2006). Properties of Realized Variance under Alternative Sampling Schemes. *Journal of Business and Economic Statistics 24*, 219–237.

Podolskij, M. and M. Vetter (2009a). Bipower-type estimation in a noisy diffusion setting. Forthcoming in the *Stochastic Processes and their Applications.*

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

Shephard, N. and K. Sheppard (2009). Realising the future: forecasting with high frequency based volatility (heavy) models. Economics Series Working Papers 438, University of Oxford, Department of Economics.

Tauchen, G. E. and H. Zhou (2006). Realized jumps on financial markets and predicting credit spreads. Finance and Economics Discussion Series 2006-35, Board of Governors of Federal Reserve System.