CENTRAL BANK OF CYPRUS
EUROSYSTEM

WORKING PAPER SERIES

A Comprehensive Comparison of Alternative Tests for Jumps in Asset Prices

Marina Theodosiou

Filip Zikes

July 2011

Working Paper 2011-2

Central Bank of Cyprus Working Papers present work in progress by central bank staff and outside contributors. They are intended to stimulate discussion and critical comment. The opinions expressed in the papers do not necessarily reflect the views of the Central Bank of Cyprus or the Eurosystem.

Address
80 Kennedy Avenue
CY-1076 Nicosia, Cyprus

Postal Address
P. O. Box 25529
CY-1395 Nicosia, Cyprus

E-mail
publications@centralbank.gov.cy

Website
http://www.centralbank.gov.cy

Fax
+357 22 378153

Papers in the Working Paper Series may be downloaded from:
http://www.centralbank.gov.cy/nqcontent.cfm?a_id=5755

A Comprehensive Comparison of Alternative Tests for Jumps in Asset Prices

Marina Theodosiou* and Filip Zikes**

July 2011

Abstract

This paper presents a comprehensive comparison of the existing tests for the presence of jumps in prices of financial assets. The relative performance of the tests is examined in a Monte Carlo simulation, covering scenarios of both finite and infinite activity jumps, stochastic volatility models with continuous and discontinuous volatility sample paths, microstructure noise, infrequent trading and deterministic diurnal volatility. The simulation results reveal important differences in terms of size and power across the different data generating processes and sensitivity to the presence of zero returns and microstructure frictions in the data. An empirical application to assets from different classes complements the analysis.

Keywords: Quadratic variation, jumps, stochastic volatility, realized measures, high-frequency data.

JEL Classification: E31, C12, C14, G10.

---

* Central Bank of Cyprus, **Imperial College London. We are indebted to Walter Distaso, Nour Meddahi and Roel Oomen for helpful comments. We also thank Roel Oomen for sharing his Matlab code with us. Zikes gratefully acknowledges financial support from ESRC, grant code RES-062-23-0311. A web appendix with the full set of simulation results is available at http://www3.imperial.ac.uk/people/f.zikes06/research. All calculations reported in this paper were performed using routines written in the Ox language of Doornik (2002).

Correspondence: Marina Theodosiou, Central Bank of Cyprus, 80 Kennedy Avenue, CY-1076, Nicosia, Cyprus, Telephone: +357 22714496. Email: MarinaTheodosiou@centralbank.gov.cy, Filip Zikes, Imperial College London, Business School, Exhibition Road, London SW7 2AZ. Telephone: +44 (0)795791 3476. Fax: +44 (0)20 7823 7685. Email: fzikes@imperial.ac.uk.

1 Introduction

Since the seminal work of Merton (1976) on the application of jump processes in option pricing, the inclusion of such processes in financial modeling has gained a lot of attention amongst academics and practitioners. It is now well documented that price discontinuities constitute an important component of variability in financial asset prices and thereby contribute to market incompleteness. In addition, it is recognized that the presence of jumps in price sample paths carries important implications for financial risk management and portfolio allocation, as well as pricing and hedging of derivatives (see, among others, Das (2002), Johannes (2004) and Piazzesi (2005) for interest rate modeling, Eberlein and Raible (1999) for bond pricing, Bakshi, Cao, and Chen (1997), Bates (2000) and Pan (2002) for derivative pricing).

The theoretical developments in the asset pricing literature have inspired a new stream of research developing statistical techniques for detecting discontinuities from discretely observed prices. Aït-Sahalia (2002) was one of the first authors to attempt separating jumps from diffusion in a parametric context. Using options data and the properties of the transition density corresponding to discrete observations, the author finds that market option prices are inconsistent with a pure diffusion model driving the underlying price process. Carr and Wu (2003) developed a different methodology based on the behavior of short dated options across maturities and at fixed moneyness states and reached a similar conclusion.

With the greater availability of high-frequency data, recent literature has focused on detecting and testing for jumps in a nonparametric, model-free context. Mancini (2001) was the first to estimate jumps in a simple jump-diffusion framework. Following her work, a large number of formal tests have been developed for detecting discontinuities in intraday price processes, including Barndorff-Nielsen and Shephard (2004), Huang and Tauchen (2005), Lee and Mykland (2008), Jiang and Oomen (2008), Corsi, Pirino, and Renò (2008) and Podolskij and Ziggel (2008).

The purpose of this paper is to compare the existing tests for jumps in a an extensive Monte Carlo simulation. We consider various models with finite and infinite-activity jumps processes to study the size and power of the various tests. We also study the impact of market microstructure noise, the presence of zero intraday returns typically found in intraday data even for liquid assets, and the effect of deterministic diurnal volatility, which is also characteristic of most high-frequency return series. We first examine the size of each test and data generating process and then turn to their ability to detect jumps. Here we calculate the proportion of correctly identified days when jumps occurred in the simulation, and also compare the tests in terms of their agreement about whether or not a jump occurred on a given day.

Overall, we find that substantial differences exist among the competing tests both in terms of size and power and that the differences vary across the data generating processes considered. First, the test that employ thresholds are very sensitive to a particular choice of the threshold parameters. Second, stochastic volatility that exhibit sudden erratic movements posses a serious challenge to some tests even at very high sampling frequencies. Third, the tests robust microstructure noise work well when the noise is independently and identically distributed and has moderate signal-to-noise ratio but some of the tests get in trouble when faced with high-variance or highly persistent noise. Fourth, the presence of zero returns has severe impact on almost all tests and results into substantial increase in the spurious detection of jumps. Finally, the deterministic U-shaped intraday volatility pattern induces very similar distortions as the highly erratic stochastic volatility case.

In the empirical part of the paper, we apply the tests for jumps to recent samples of high-frequency foreign exchange, individual stocks and equity index futures. Similar to the simulation results, we find that the tests yield different results regarding the identification of jumps and tend to disagree as to whether or not a jump occurred on a given day.

The paper is organized as follows. The theoretical framework underlying the various test is described in section 2, while in section 3, we provide a brief description of the different tests studied in this paper. Section 4 discusses the Monte Carlo simulation design and the simulation results. In section 5, we investigate the impact of market microstructure noise and in section 6 we look at the effect of the presence of zero intraday returns on the behavior of the tests. Section 7 is dedicated to an empirical application and, finally, section 8 concludes.

2 Theoretical Framework

Let X_t denote the logarithmic price process that belongs to the class of Brownian semi-martingales, which can be written as

$$X_t = \int_0^t a_u \mathrm{d}u + \int_0^t \sigma_u \mathrm{d}W_u + Z_t, \tag{1}$$

where a is the drift term, σ denotes the spot volatility process, W is a standard Brownian motion and Z is a jump process defined by:

$$Z_t = \sum_{j=1}^{N_t} \kappa_j,$$

where N is a simple counting process and κ_j are nonzero random variables. The counting process can be either finite or infinite for finite or infinite activity jumps.

Since the seminal work of Andersen and Bollerslev (1998), realized volatility, $RV_{t,M}$, obtained by summing M squared intraday returns has become the standard measure of the quadratic variation of the price process in (1). Formally,

$$RV_{t,M} = \sum_{i=1}^{M} r_{t_i}^2,$$

where r_{t_i} denotes the i-th intraday return on day t:

$$r_{t_i} = x_{t-1+i/M} - x_{t-1+(i-1)/M} \qquad \text{for } i = 1, 2, 3, \ldots, M.$$

RV can be used to approximate the variation of both the continuous and the discontinuous part of the price process since

$$\begin{aligned} \lim_{M\to\infty} RV_{t,M} &= \int_0^t \sigma_s^2 \mathrm{d}s + \sum_{j=1}^{N_t} \kappa_j^2, \\ &\equiv IV_t + JV_t. \end{aligned}$$

However, in empirical applications one may be concerned with the behavior of IV_t and JV_t in isolation, and it is therefore essential to decompose the two sources of variability of the price process.

3 The Tests

3.1 Tests based on multipower variation

The first formal test developed for detecting jumps in high frequency data was constructed by Barndorff-Nielsen and Shephard (2004, 2006) (henceforth BNS). Their work was later extended and further investigated in Andersen, Bollerslev, and Diebold (2004) and Huang and Tauchen (2005) using a variety of asymptotically equivalent statistics.

To consistently estimate integrated variance in presence of jumps BNS propose the realized bipower variation (BPV) defined by,

$$BPV_{t,M} = \sum_{i=2}^{M} |r_{t_{i-1}}||r_{t_i}|.$$

The idea underlying the bipower variation is that if the jumps are of finite activity, the probability of observing jumps in two consecutive returns approaches zero sufficiently fast as the sampling frequency increases. Consequently, the product of any two consecutive returns will be asymptotically driven by the diffusion component only thereby eliminating the contribution of jumps.

Since the realized volatility converges to the sum of integrated variance and jump variation, it follows that the difference between $RV_{t,M}$ and $BPV_{t,M}$ captures the jump part only, and this observation underlies the BNS test for jumps. Based on the joint Central Limit theorem (CLT) of RV and BPV, they propose the following test statistics for testing the null hypothesis of no jumps:

$$G_{t,M}^{bpv,qpq} = \frac{RV_{t,M} - BPV_{t,M}}{\sqrt{\theta_2 \frac{1}{M} QPQ_{t,M}}} \xrightarrow{L} \mathrm{N}(0,1),$$

where $QPQ_{t,M}$ denotes the realized quadpower quarticity given by,

$$QPQ_{t,M} = M \sum_{i=4}^{M} |r_{t_{i-3}}||r_{t_{i-2}}||r_{t_{i-1}}||r_{t_i}| \underset{M\to\infty}{\longrightarrow} \int_0^t \sigma_s^4 \mathrm{d}s.$$

The test can be generalized by replacing $BPV_{t,M}$ with $MPV_{t,M}$ and $QPQ_{t,M}$ with $MPQ_{t,M}$, i.e. the realized *multipower variation* and realized *multipower quarticity* respectively. These are defined by,

$$\begin{aligned} MPV_{t,M}(p) &= \mu_{2/p}^{-p} \tfrac{M}{M-p+1} \textstyle\sum_{i=p}^{M} \prod_{j=0}^{p-1} |r_{t_{i-j}}|^{2/p}, \\ MPQ_{t,M}(p) &= \mu_{4/p}^{-p} \tfrac{M}{M-p+1} \textstyle\sum_{i=p}^{M} \prod_{j=0}^{p-1} |r_{t_{i-j}}|^{4/p}, \end{aligned}$$

μ_p denotes the pth absolute moment of a variable $U \sim N(0,1)$ defined by,

$$\mathbb{E}(|U|^p) = \pi^{-1/2} 2^{p/2} \Gamma\left(\frac{p+1}{2}\right),$$

and $\theta_p = \mu_{2/p}^{-2p} \omega_p^2$, for $\omega_p^2 = \mu_{4/p}^{p} + (1-2p)\mu_{2/p}^{2p} + 2\sum_{j=1}^{p-1} \mu_{p/4}^{p-j} \mu_{2/p}^{2j}$. The bivariate limit theory for realized variance and realized multipower variation both in the presence and absence of jumps is studied by Veraart (2008).

Various alterations to the above test statistic have been suggested to improve the finite sample performance of the test. These include the logarithmic and ratio tests and some finite-sample corrections in the denominator of the test statistic (Andersen, Bollerslev, and Diebold, 2004, Huang and Tauchen, 2005, andBarndorff-Nielsen and Shephard, 2006). Throughout the analysis, we will be using the adjusted jump ratio statistic:

$$J_{t,M}^{mpv,mpq} = \frac{\left(1 - \frac{MPV_{t,M}}{RV_{t,M}}\right)}{\sqrt{\theta_p \frac{1}{M} \max(1, MPQ_{t,M}/MPV_{t,M}^2)}} \xrightarrow{L} \mathrm{N}(0,1),$$

since this has been shown to be the best option amongst the three alternatives (Huang and Tauchen, 2005) in term of finite-sample performance. We will also explore various combinations of the bipower, tri-power and quad-power variation in the nominator and the tri-power and quad-power quarticity in the denominator of the test statistic.

3.2 Tests based on threshold multipower variation

A test that combines the idea of the threshold estimators of Mancini (2001) and the multi-power variation estimation of BNS was proposed by Corsi, Pirino, and Renò (2008) (henceforth CPR). The authors argue that truncating large absolute returns alleviates the bias associated with multipower variation in the presence of jumps.

Their test statistic is therefore based on the *realized threshold multipower* variation defined by

$$TMPV_{t,M}(p) = \mu_{2/p}^{-p} \frac{M}{M-p+1} \sum_{i=p}^{M} \prod_{j=0}^{p-1} |r_{t_{i-j}}|^{2/p} I_{\left\{|r_{t_{i-j}}| \le \vartheta_{t_{i-j}}\right\}}.$$

The threshold ϑ_{t-1+j} is defined as a multiple of the local variance, which is approximated by a local linear filter of length $2L+1$, adjusted iteratively for the presence of jumps:

$$\vartheta_t = c_\theta^2 \hat{V}_t^Z,$$

where c_θ is a constant, and $\hat{V}_t^Z$ denotes an estimator of local variance. The latter is given by:

$$\hat{V}_t^Z = \frac{\sum_{i=-L, i\neq -1,0,1}^{L} K\left(\frac{i}{L}\right)(r_{t_i})^2 I_{\left\{(r_{t_i})^2 \leq c_V^2 \hat{V}_{t_i}^{Z-1}\right\}}}{\sum_{i=-L, i\neq -1,0,1}^{L} K\left(\frac{i}{L}\right) I_{\left\{(r_{t_i})^2 \leq c_V^2 \hat{V}_{t_i}^{Z-1}\right\}}}.$$

Z denotes the iteration number with starting value $\hat{V}^0 = +\infty$, which corresponds to using all observations. c_V is a constant and $K(.)$ denotes the Gaussian kernel:

$$K(y) = \left(1/\sqrt{2\pi}\right)\exp(-y^2/2).$$

In order to avoid a negative bias associated with introducing zero returns by truncations, the authors correct the realized threshold multipower variation by replacing the absolute squared returns that exceed the threshold with their expected value under the null hypothesis of no jumps. Thus, the corrected estimator is given by

$$cTMPV_{t,M}(p) = \mu_{2/p}^{-p} \frac{M}{M-p+1} \sum_{i=p}^{M} \prod_{j=0}^{p-1} Z(r_{t_{i-j}}, \vartheta_{t_{i-j}}).$$

$Z(x,y)$ is defined as:

$$Z(x,y) = \begin{cases} |x|^{2/p}, & \text{if } x^2 \leq y \\ \frac{1}{2M(-c_\vartheta)\sqrt{\pi}}\left(\frac{2}{c_\vartheta^2} y\right)^{1/p} \Gamma\left(\frac{2/p+1}{2}, \frac{c_\vartheta^2}{2}\right), & \text{if } x^2 > y \end{cases},$$

where $\Gamma(\alpha, x)$ is the upper incomplete gamma function.

The test statistics for jumps is then based on the corrected realized threshold multipower variation and is given by:

$$J_{t,M}^{tbv,ttpq} = \frac{\left(1 - \frac{cTBPV_{t,M}}{RV_{t,M}}\right)}{\sqrt{\theta_2 \ \frac{1}{M} \ \max(1, cTTPQ_{t,M}/cTBPV_{t,M}^2)}} \xrightarrow{L} \mathrm{N}(0,1).$$

A disadvantage of the CPR test is the need to choose the threshold parameters c_θ and c_V. The simulation results reported later in the paper indeed reveal important differences in terms of size and power of the test across different values of these constants.

3.3 Tests based on median realized volatility

Andersen, Dobrev, and Schaumburg (2009) (henceforth ADS) have proposed a new set of estimators for integrated variance in the presence of jumps. They are based on the minimum and median of a number of consecutive absolute intraday returns:

$$
\begin{aligned}
MinRV_{t,M} &= \tfrac{\pi}{\pi-2}\left(\tfrac{M}{M-1}\right)\sum_{i=1}^{M-1}\min(|r_{t_i}|,|r_{t_{i+1}}|)^2,\\
MedRV_{t,M} &= \tfrac{\pi}{6-4\sqrt{3}+\pi}\left(\tfrac{M}{M-2}\right)\sum_{i=2}^{M-1}\text{med}(|r_{t_{i-1}}|,|r_{t_i}|,|r_{t_{i+1}}|)^2.
\end{aligned}
$$

These estimators are more robust to jumps than the multipower variations since large absolute returns associated with jumps tend to be eliminated from the calculation by the minimum and median operators. Furthermore, the $MedRV$ estimator enjoys robustness against the presence of occasional zero intraday returns induced by calendar-time sampling, unlike the multipower variation that becomes downward biased.

In this paper, we exploit the joint central limit theorem for $RV_{t,M}$ and $MedRV_{t,M}$ derived by ADS to construct a test for jumps in the same way as BNS and CPR do. The test statistics read:

$$
\begin{aligned}
J_{t,M}^{medrv,minrq} &= \frac{\left(1-\frac{MedRV_{t,M}}{RV_{t,M}}\right)}{\sqrt{0.96\frac{1}{M}\max(1,MinRQ_{t,M}/MedRV_{t,M}^2)}} \xrightarrow{L} \text{N}(0,1),\\
J_{t,M}^{medrv,medrq} &= \frac{\left(1-\frac{MedRV_{t,M}}{RV_{t,M}}\right)}{\sqrt{0.96\ \frac{1}{M}\max(1,MedRQ_{t,M}/MedRV_{t,M}^2)}} \xrightarrow{L} \text{N}(0,1),
\end{aligned}
$$

where $MinRQ_{t,M}$ and $MedRQ_{t,M}$, given by

$$
\begin{aligned}
MinRQ_{t,M} &= M\frac{\pi}{3\pi-8}\left(\frac{M}{M-1}\right)\sum_{i=1}^{M-1}\text{med}(|r_{t_i}|,|r_{t_{i+1}}|)^4,\\
MedRQ_{t,M} &= M\frac{3\pi}{9\pi+72-52\sqrt{3}}\left(\frac{M}{M-2}\right)\sum_{i=2}^{M-1}\text{med}(|r_{t_{i-1}}|,|r_{t_i}|,|r_{t_{i+1}}|)^4,
\end{aligned}
$$

are consistent estimators of the integrated quarticity. Due to the nice properties of $MedRV$ discussed above, we expect these tests to be more powerful than their BNS counterparts.

3.4 Tests based on truncated power variation

Podolskij and Ziggel (2008)(henceforth PZ) build further on the threshold idea of Mancini (2001) and suggest to construct a test statistics for jumps based on the difference between power variation and the truncated version thereof, since the difference between the two captures the contribution of jumps.

The test statistic is defined by

$$S_{t,M}(p) = \frac{T(X,p)_{t,M}}{\rho^2(p)_{t,M}},$$

where $T(X,p)_{t,M}$ denotes the difference between the realized power variation and the truncated realized power variation, and $\rho^2(p)_{t,M}$ is a standardizing term:

$$T(X,p)_{t,M} = 1/M^{1-p/2} \sum_{i=1}^{M} |r_{t_i}|^p (1 - \eta_i I_{\{|r_{t_i}| \leq \alpha(1/M)^{\varpi}\}}),$$

$$\rho^2(p)_{t,M} = Var[\eta_i] MPV_{t,M}(2p),$$

where η_i are positive $i.i.d$ random variables with $\mathbb{E}[\eta_i] = 1$ and $\mathbb{E}[|\eta_i|^2] < \infty$. The test statistics is thus based on the truncated power variation constructed from randomly perturbed intraday returns as opposed to the usual threshold power variation studied by Mancini (2001). This is required to obtain limit theory for the test (see the original paper by the authors for details).

Podolskij and Ziggel (2008) suggest that η_i can be sampled from the distribution:

$$\mathbb{P}^{\eta} = \frac{1}{2}(\delta_{1-\tau} + \delta_{1+\tau}),$$

where δ is the Dirac measure. They also suggest to take $\alpha = c\sqrt{BPV_{t,M}}$, $\varpi = -0.4$ and $\tau = 0.1$ or 0.05. The threshold is therefore proportional to an initial estimate of integrated variance, while the choice of the other two constants is rather arbitrary.

Similar to the CPR test, the PZ test requires the specification of the threshold constant, which in turn affect the performance of the test. In addition, one has to specify the distribution of η_i. In this paper, we only experiment with different values of the threshold constant c and follow Podolskij and Ziggel (2008) regarding the choice of the other free parameters.

3.5 Swap variance tests

Inspired by the replication strategy of Neuberger (1994) for hedging variance swap contracts[1], Jiang and Oomen (2008)(henceforth JO) propose a new test for jumps which is based on the difference between the simple and logarithmic returns. Their idea compares and contrasts to that of BNS in that they use a jump-sensitive measure to be compared with the realized volatility rather than a jump-robust measure, as in BNS.

The underlying idea behind the variance swap replication strategy is that in the absence of jumps, the accumulated difference between the simple return and the log return captures one half of the integrated variance. Thus

$$SwV_{t,M} = 2\sum_{i=1}^{M}(R_{t_i} - r_{t_i}) \xrightarrow{p} \int_{t-1}^{t} \sigma_u^2 \mathrm{d}u,$$

[1]A variance swap is a contract whose payoff is equal to the difference between the square of annualized realized volatility of the underlying price over a given time period, and a strike price fixed at the inception of the contract

where for the series of log prices X_t and $i = 1, 2, 3, \ldots, M$,

$$R_{t_i} = \frac{\exp(x_{t-1+i/M}) - \exp(x_{t-1+(i-1)/M})}{\exp(x_{t-1+(i-1)/M})},$$

and r_{t_i} denotes the continuously compounded returns defined in (4). Thus in the absence of jumps, the difference between $SwV_{t,M}$ and $RV_{t,M}$ converges to zero. If jumps are present, however, the limit reads

$$SwV_{t,M} - RV_{t,M} \xrightarrow{p} 2 \sum_{t_j \in [t-1,t]} (\exp(\kappa_j) - \kappa_j - 1) - \sum_{t_j \in [t-1,t]} \kappa_j^2,$$

Building on this insight, JO define the test statistics for jumps as follows:

$$JO_{t,M} = \frac{M}{\sqrt{\Omega_{SwV}}}(SwV_{t,M} - RV_{t,M}) \xrightarrow{L} \mathrm{N}(0,1),$$

where

$$\Omega_{SwV} = \frac{\mu_6}{9} \frac{M^3 \mu_{6/p}^{-p}}{M-p-1} \sum_{i=0}^{M-p} \prod_{k=1}^{p} |r_{t_i}|^{6/p}$$

is an estimator of integrated sixticity, $\int \sigma_u^6 du$. The authors suggest using $p = 4$ and $p = 6$.

Similarly to BNS, CPR and ADS, JO find that a test based on the ratio of $SwV_{t,M}$ and $RV_{t,M}$ exhibits better finite-sample properties than the difference test in equation (3.5). The ratio test statistics is given by,

$$JO_{t,M} = \frac{M \cdot BPV_{t,M}}{\sqrt{\Omega_{SwV}}} \left(1 - \frac{RV_{t,M}}{SwV_{t,M}}\right) \xrightarrow{L} \mathrm{N}(0,1),$$

and this is the version of the test we employ in this paper.

3.6 Tests based on two-time scales power variation

Using the convergence properties of power variation and its dependence on the time scale on which it is measured, Aït-Sahalia and Jacod (2009) (henceforth ASJ) define a new variable which converges to 1 in the presence of jumps in the underlying return series, or to another deterministic and known number in the absence of jumps. This quantity is defined as the ratio of power variations calculated under two different time scales ($1/M$ and k/M):

$$\hat{S}(p, k, 1/M)_t = \frac{\hat{B}(p, k/M)_t}{\hat{B}(p, 1/M)_t}$$

where

$$\hat{B}(p, 1/M)_t = \sum_{i=1}^{M} |r_{t_i}|^p \qquad p > 2$$

denotes the usual power variation. Under the null hypothesis of no jumps and with $p > 2$, $\hat{S}(p, k, 1/M)_t$ converges to $k^{p/2-1}$, while under the alternative the limit is equal to one.

Building on these insights, the ASJ test statistics for the null hypothesis of no jumps is defined as

$$\frac{\hat{S}(p,k,1/M)_t - k^{p/2-1}}{\sqrt{\hat{V}^c_{t,M}}}$$

where $\hat{V}^c_t$ denotes the asymptotic variance of $\hat{S}(p,k,1/M)_t$ and is given by,

$$\hat{V}^c_t = \frac{1/M \ N(p,k) \ \hat{A}(2p,1/M)_t}{\hat{A}(p,1/M)^2_t},$$

where

$$\begin{aligned}
\hat{A}(p,1/M)_t &= \frac{1/M^{1-p/2}}{\mu_p}\sum_{i=1}^{M}|r_{t_i}|^p I_{\{|r_{t_i}|\leq\alpha(1/M)^{\varpi}\}} \\
N(p,k) &= \frac{1}{\mu_p^2}(k^{p-2}(1+k)\mu_{2p} + k^{p-2}(k-1)\mu_p^2 - 2k^{p/2-1}\mu_{k,p} \\
\mu_{k,p} &= \mathbb{E}(|U|^p|U+\sqrt{k-1}V|^p)
\end{aligned}$$

for U, V independent standard normal random variables.

The ASJ test requires the choice of four parameters, namely p, k, α and ϑ. In this paper, we follow ASJ in using $p = 4$ and $k = 2$ and experiment with different values of the threshold parameters.

3.7 Tests based on local volatility

The last test for jumps we consider in this paper is the one developed by Lee and Mykland (2008)(henceforth LM). The intuition behind their approach is that the magnitude of price changes depends on the local volatility conditions and that a 'large' price change does not necessarily imply a jump in the return process without conditioning on the current variability. An important advantage of their tests lies in the fact that one can draw conclusions not only about the presence of jumps in a given time period, but also about the number and location of jumps within this period.

For every intraday period t_i, LM propose to calculate the ratio between the intraday return, r_{t_i}, and the instantaneous volatility, σ_{t_i}, which they approximate using bipower variation, i.e.

$$\mathcal{L}(i) = \frac{r_{t_i}}{\widehat{\sigma_{t_i}}}$$

where

$$\widehat{\sigma_{t_i}}^2 = \frac{1}{K-2}\sum_{j=i-K+2}^{i-1}|r_{t_i}||r_{t_{i-1}}|$$

and K denotes the window size or bandwidth used for the estimation of the instantaneous volatility. If a jump occurred in a given period of time, this ratio should be large in absolute value and vice versa. This idea underlies the test statistic for jumps in the intra-day period t_i:

$$\frac{|\mathcal{L}(i)| - C_M}{S_M},$$

where

$$C_M = \frac{(2\log(M))^{1/2}}{\mu_1} - \frac{\log(\pi)+\log(\log(M))}{2\mu_1(2\log(M))^{1/2}} \qquad \text{and} \qquad S_M = \frac{1}{\mu_1(2\log(M))^{1/2}}$$

represent the centering and normalizing terms.

To select a rejection region for the test, LM derive the limiting distribution of the maximum of $|\mathcal{L}(i)|$ over all $i = 1, ..., M$ and show that the limiting distribution implies that for a given significance level α, the relevant threshold for $\frac{|\mathcal{L}(i)|-C_M}{S_M}$ is given by $\beta = -\log(-\log(1-\alpha))$. Thus if $\frac{|\mathcal{L}(i)|-C_M}{S_M} > \beta$ the null hypothesis of no jump at time t_i is rejected. The choice of the bandwith parameter K is guided by asymptotic theory and the authors recommend using a value of $\sqrt{252 \times M}$.

4 Monte Carlo Simulation

4.1 Simulation Design

We consider three different data generating processes (DGPs) to investigate the size and power properties of the various tests for jumps described above. The first two are the one and two-factor log-linear stochastic volatility (SV) models studied by Chernov, Gallant, Ghysels, and Tauchen (2003), and employed by Barndorff-Nielsen and Shephard (2004) and Huang and Tauchen (2005) in a simulation study of the behavior of the bipower variation based tests. These are defined by:

LL1F: one-factor log-linear SV

$$\begin{aligned}
\mathrm{d}p(t) &= \mu \mathrm{d}t + \exp[\beta_0 + \beta_1 v(t)]\mathrm{d}W_p(t), \\
\mathrm{d}v(t) &= \alpha_v v(t)\mathrm{d}t + \mathrm{d}W_v(t),
\end{aligned}$$

LL2F: two-factor log-linear SV

$$\begin{aligned}
\mathrm{d}p(t) &= \mu \mathrm{d}t + \mathrm{sexp}[\beta_0 + \beta_1 v_1(t) + \beta_2 v_2(t)]\mathrm{d}W_p(t), \\
\mathrm{d}v_1(t) &= \alpha_{v1} v_1(t)\mathrm{d}t + \mathrm{d}W_{v1}(t), \\
\mathrm{d}v_2(t) &= \alpha_{v2} v_2(t)\mathrm{d}t + [1 + \beta_{v2} v_2(t)]\mathrm{d}W_{v2}(t),
\end{aligned}$$

where W_p, W_v, W_{v1}, and W_{v2} are standard Brownian motions with leverage correlations $\mathrm{Corr}(\mathrm{d}W_p(t), \mathrm{d}W_v(t)) = \rho \mathrm{d}t$, $\mathrm{Corr}(\mathrm{d}W_p(t), \mathrm{d}W_{v1}(t)) = \rho_1 \mathrm{d}t$, and $\mathrm{Corr}(\mathrm{d}W_p(t), \mathrm{d}W_{v2}(t)) = \rho_2 \mathrm{d}t$, and $v(t)$, $v_1(t)$ and $v_2(t)$ are stochastic volatility factors. The process $v_1(t)$ is a standard Gaussian process, while $v_2(t)$ exhibits a feedback term in the diffusion function. The spliced exponential function sexp ensures a solution to LL2F exists (see Chernov, Gallant, Ghysels, and Tauchen, 2003, for details).

The third DGP is a log-linear stochastic volatility model in which the volatility factor follows an infinite-activity pure-jump process recently considered by Todorov and Tauchen (2008):

LLIA: infinite-activity pure-jump SV

$$\begin{aligned} \mathrm{d}p(t) &= \mu \mathrm{d}t + \exp[\beta_0 + \beta_1 v(t)] \mathrm{d}W_p(t), \\ \mathrm{d}v(t) &= \alpha_v v(t) \mathrm{d}t + \mathrm{d}L_v(t) \end{aligned}$$

where L_v is a symmetric tempered stable process with Lévy density given by $\nu(x) = c\frac{\mathrm{e}^{-\lambda|x|}}{|x|^{1+\alpha}}$, $\alpha \in (0,2)$. The parameter α_v measures the degree of activity of jumps, while λ governs the tail behavior of the Lévy density.

We use the same parametrization for LL1F and LL2F as in Huang and Tauchen (2005) (see Table 1 & 2). For the LLIA, we fix λ at 2.5 as Todorov and Tauchen (2008) and vary c and α such that the variance of $L_v(1)$ remains constant at 1, (see Table 3). Thus the first two moments of the increments of the volatility factor v_t are identical under LL1F and LLIA, but these have fatter tails under LLIA. The sample paths are, of course, dramatically different with the former being continuous while the latter purely discontinuous.

To simulate sample paths of the log-price under LL1F and LL2F we use the Euler discretization scheme with the increment of the Euler clock set to 1 second. We generate 55,000 trading days, each 6.5 hours long, which corresponds to typical trading hours on major equity exchanges. We discard the first 5000 days to avoid distortions induces by initial conditions. For each day, we calculate the test statistics for jumps at different sampling frequencies ranging from 30 seconds to 15 minutes.

The simulation of the tempered stable process in LLIA is based on the series representation of tempered stable processes derived by Rosiński (2001), and outlined in Todorov (2007). For each 6.5-hour day, we generate 2,340 intraday observations of L_v corresponding to 10-second sampling. We truncate the infinite series expansion such that we simulate on average 10,000 jumps in L_v per day.

To study the power properties of the various tests for jumps, we first augment the LL1F model by a pure jump component of finite activity:

LL1F-FAJ: one-factor log-linear SV with finite-activity jumps

$$\begin{aligned} \mathrm{d}p(t) &= \mu \mathrm{d}t + \exp[\beta_0 + \beta_1 v(t)] \mathrm{d}W_p(t) + \mathrm{d}J_t, \\ \mathrm{d}v(t) &= \alpha_v v(t) \mathrm{d}t + \mathrm{d}W_v(t), \end{aligned}$$

where J_t is a compound Poisson process with normally distributed jumps with variance σ^2 and constant jump intensity λ. We experiment with various combinations of σ^2 and λ, ranging from large infrequent jumps to small frequent ones similar to Huang and Tauchen (2005) (see Table 4).

Next, we explore power against alternatives that entail infinite-activity jump processes:

LL1F-IAJ: one-factor log-linear SV with infinite-activity jumps

$$\begin{aligned} \mathrm{d}p(t) &= \mu \mathrm{d}t + \exp[\beta_0 + \beta_1 v(t)] \mathrm{d}W_p(t) + k\mathrm{d}L_t, \\ \mathrm{d}v(t) &= \alpha_v v(t) \mathrm{d}t + \mathrm{d}W_v(t), \end{aligned}$$

where k is a constant and L_t is a symmetric tempered stable process with Lévy density given by $\nu(x) = c\frac{e^{-\lambda|x|}}{|x|^{1+\alpha}}$, $\alpha \in (0,2)$. We use the same parameter values for the jump process as Todorov (2007) (Table 5). The parameters are calibrated such that the contribution of the jump component to the overall variation reflects the results from previous empirical literature (Huang and Tauchen, 2005).

We implement the above discussed tests for jumps in the following way:

- Multipower variation ratio test (BNS) using BV, TPV and QPV to estimate the integrated variance and TPQ or QPQ to estimate the integrated quarticity;
- Threshold bipower variation ratio test (CPR) using threshold TPQ to estimate the integrated quarticity; the choice of threshold follows CPR exactly with ϑ set to 3,4 or 5;
- Median realized volatility ratio test (ADS) using either MinRQ or MedRQ to estimate the integrated quarticity;
- Swap variance ratio test (JO) using either realized quadpower or sixthpower sixticity to estimate the integrated sixticity;
- Two-scale power variation test (ASJ): we set $p = 4$, $k = 2$ as suggested by ASJ, using truncated power variation to estimate the asymptotic variance of the $\hat{S}$ statistics with $\alpha = 0.47$ and ϑ set to 3,4 or 5.
- Truncated power variation test (PZ): we consider $p = 2$ and $p = 4$ and set $\tau = 0.05$, $\vartheta = 0.4$ and $c = 2.3, 3, 4$.
- Test based on local volatility (LM) using BV, TPV and QPV to estimate instantaneous volatility.

For expositional clarity, we summarize the main results in a few tables focusing on the most important differences across the tests and the various data generating process. The simulation results not directly reported here do not provide much additional insight but they are available upon request.

4.2 Size

Table 1 summarizes the simulated size for 1% nominal level. We report results for the LL1F model with moderate mean reversion, LL2F and LLIA with activity index of volatility jumps equal to 0.4. Other parameter configurations of the DGP's yield similar results and are omitted to save space.

We find that while the BNS, CPR and ADS tests exhibit only small size distortions in small samples, the swap variance test (JO) and the truncated power variation test (PZ) tend to be oversized and the ASJ test significantly undersized in moderate samples. This observation is true for all stochastic volatility models considered here.

Starting with the LL1F model with moderate mean reversion, we find that the BNS ratio test based on the use of bipower variation exhibits slight positive size distortions at lower frequencies, as shown by Huang and Tauchen (2005) before, but these are rather negligible from an empirical perspective. When realized tripower variation is used in place of realized bipower variation in the BNS test, the size distortions essentially disappear. However, using realized quadpower variation, in particular when coupled with realized tripower quarticity, reduces the size of the BNS test below the nominal level at low sampling frequencies. Simulation evidence not reported here suggests that this is caused by positive skewness of realized multipower variation at low frequencies, with the degree of skewness increasing with p. Thresholding the bipower variation as suggested by CPR tends to increase the size slightly for low levels of the threshold. This is perhaps not surprising given that it is not optimal from a statistical point of view to truncate the large returns in the absence of jumps. This biases the bipower variation downward and the test statistics upward, hence the slightly higher empirical size.

The jumps tests based on the recently developed median realized volatility (ADS) show relatively stable performance across sampling frequencies. They do tend to be slightly oversized at low sampling frequencies but the distortions are smaller than in the case of the test based on bipower variation. The choice of the estimator of integrated quarticity (MinRQ vs. MedRQ) does not seem to have a practical impact on the size properties.

Turning to the ASJ test, we first note the difference in size depending on the truncation parameter α. The higher α, that is, the larger the threshold employed in the calculation of the truncated power variation, the lower the size. But more importantly, for a given threshold, decreasing the sampling frequency tends to have a significant negative impact on the size of the test. This effect is more pronounced for more conservative significance levels. The simulation evidence indicates that the problem lies with the positive skewness of the ASJ test statistics at lower frequencies. Already at the 2 minute frequency does the ASJ test statistic exhibit significant departures from the standard normal limiting distribution, showing much more probability mass in the right tail than in the left one. This problem becomes more severe at lower sampling frequencies; for example, at 15 minutes, the empirical size of the test is only about one half of the nominal level, irrespective of the speed of mean reversion.

Similar to ASJ, the performance of the PZ test also depends on the choice of threshold. For the relatively small value of the threshold recommended by PZ ($c = 2.3$) we find large positive size distortions for moderate and low sampling frequencies. Most of these distortions are nonetheless alleviated by slightly increasing the threshold ($c = 3$).

We next look at the size properties under the two-factor SV model (LL2F). As shown by Huang and Tauchen (2005), the BNS tests tend to be oversized in this case and this result is confirmed in our simulation for all multipower variation-based tests. Similar results are obtained for the ADS, JO and CPR tests with the latter being much more sensitive to the choice of threshold than in the case of the LL1F model. In the two-factor model, the volatility process experiences sudden erratic movements generating large absolute price increments which can be easily confused with jumps. Setting too small a threshold will eliminate these large but genuine diffusive intraday returns and bias the threshold bipower variation downward, resulting in false rejections of the null hypothesis of no jumps.

The LL2F scenario is much more challenging for the ASJ and PZ test. Not only do these tests become extremely sensitive to the choice of threshold, the size distortions do not seem to disappear in large samples. In fact, increasing the sampling frequency exacerbates the problem, questioning the workings of the limit theory under the two-factor model.

Finally, we investigate whether the pure-jump volatility specification (LLIA) affects the finite-sample properties of the tests for jumps. The results, reported in the web appendix, are very similar to the LL1F case, suggesting that even highly active pure jump volatility process does not adversely affect the inference about jumps beyond the distortions observed for relatively smooth continuous volatility specifications (LL1F).

Summarizing the size simulations, we find that the BNS and ADS tests exhibit most stable performance across the different DGP's and sampling frequencies. The tests that require thresholding (ASJ, PZ and CPR) seem to be very sensitive to the choice of threshold.

4.3 Power against finite-activity jumps

Having examined the size, we now turn to power against finite-activity jumps. We report three different jumps scenarios, ranging from large, infrequent jumps ($\lambda = 0.1$, $\sigma^2 = 2.5$) up to small and frequent ones ($\lambda = 2.0$, $\sigma^2 = 0.5$). It is well-known that when applied on a day-by-day basis, the jump tests are inconsistent (Huang and Tauchen, 2005): for any given finite time-period there is always a positive probability that no jump occurs and hence none of the tests can discriminate between a continuous price process and a price process with jumps of finite activity. For example, with $\lambda = 0.1$, a jump occurs only about every 10 days, and hence the tests will have no chance of detecting jumps on 9 out of 10 days on average. It is therefore more instructive to focus on the ability of the jump test to detect jumps on days when jumps indeed occurred, which can be neatly summarized by the confusion matrix described in what follows.

Tables 2-4 report confusion matrices for the eight tests at different sampling frequencies applied on a day-by-day basis under the LL1F data generating process with moderate mean reversion of -0.100 and significance level of 1%. We only focus on one version of each test, except for the test that require thresholding where we report two test statistics in order to study the dependence on the choice of threshold.

The confusion matrices are constructed as follows. The diagonal elements show the proportion of correctly identified jump days by each test individually. For example, a value of 70 means that the particular test manages to detect 70% of the days on which jumps occurred in the simulation. The off-diagonal elements then report the proportion of jump day jointly flagged by a pair of tests. They provide a measure of agreement between the two tests about the occurrence of jumps.

Starting with the scenario of large, frequent jumps, we find that the PZ test is the most powerful one, closely followed by the LM and JO tests. In case of the former, the power crucially depends on the choice of threshold: a smaller value of the threshold parameter leads to a larger proportion of detected jumps, but recall from the previous section that it also produces important size distortions. This trade-off is particularly pronounced at lower sampling frequencies. For example, for the 15 minute sampling frequency, the PZ test with the threshold value of 2.3, which was recommended by PZ, correctly identifies 79.8% of jump days while with the threshold set equal to 4 the proportion decreases by more than 20 percentage points to 59.1%. The CPR test behaves in a similar way but the dependence on the threshold parameter is less strong.

The BNS and ADS tests deliver good performance, with the latter test being slightly more powerful than the former, as expected. Replacing the bi-power variation by the tri-power or quad-power variations in the BNS test (not reported here) does not lead to improvements in power, however, despite the fact that the TPV and QPV are more robust to jumps than BV. This is probably due to the higher variance of the two measures of integrated variance. Finally, the ASJ test is the least powerful out of all tests at the sampling frequencies reported here. It seems to work well only up to the one-minute frequency and looses power quickly hereafter.

In terms of pairwise agreement among the tests as to whether or not a jump occurred on a given day, which we report in the off-diagonal part of the confusion matrix, we find that the proportion of commonly detected jump days tends to be driven by the test with lower power. In particular, the proportion of commonly detected jumps tends to be slightly smaller that the proportion of jumps detected individually by the test with lower power. This implies that if the test with lower power detects a jump so does the test with higher power. Consider, for example, the swap variance test, JO, which detects 91.20% of jumps days at the 1 minute frequency, together with the ADS test that identifies 88.94% of them. The implied minimum proportion of commonly detected jump days equals 81.14%, while the maximum, given by the ADS performance, is 88.94%. The actual proportion obtained by the simulation equals 88.43%, which is very close to the upper bound. Similar observations are made for the other pair of tests.

The other jump scenarios, reported in Tables 3-4, produce qualitatively similar results in term of the relative performance of the tests. The power decreases across the board as the variance of the jumps decreases and the jump intensity increases. The degree of disagreement among the tests increases slightly as well.

4.4 Power against infinite-activity jumps

We next examine the power against infinite activity jumps. It is important to note that not all of the tests studied here are designed for this kind of departure from the null hypothesis of no jumps. This is because some of the measures of integrated variance that these test employ are not robust to the presence of infinite activity jumps and thus cannot be used to disentangle the continuous and discontinuous components of volatility. In particular, all test based on the bi-power variation and minimum or median realized volatility suffer from this problem. We study them in the context of infinite activity jumps nonetheless for they may still possess non-trivial power against this alternative.

The results are summarized in Table 10. We find the relative performance to be very similar to the case of finite activity jumps. The PZ test with low threshold delivers highest power, followed by the LM and JO tests. The latter test works particularly well in that it retain power even at the low sampling frequency of 15 minutes. The BNS and ADS tests fare relatively well despite being based on non-robust measures of integrated variance. The ASJ test only works at very high frequencies.

5 Microstructure noise

It is now widely recognized that the estimation of the realized variance at very high frequencies is heavily biased by the presence of market microstructure noise (Hansen and Lunde (2006)). This contamination of the "efficient price" arises from a wide range of market frictions including bid-ask spread, infrequent trading, inventory control problems and asymmetric information, among others[2]. The noise dominates the estimation results predominantly at finely sampled data, and thus creates a trade-off between the efficiency and the bias due to contamination. A vast literature has been developed with techniques attempting to reduce or eliminate such frictions, see Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008a) and the references therein.

The literature typically assumes that the efficient price $X_{t+i/M}$ is contaminated by an additive noise component,

$$X^*_{t+i/M} = X_{t+i/M} + \epsilon_{t+i/M},$$

where $\mathbb{E}[\epsilon_{t+i/M}] = 0$ and $\mathrm{Var}[\epsilon_{t+i/M}] = \omega^2 < \infty$. Various assumptions are made regarding the dependence between $X_{t+i/M}$ and $\epsilon_{t+i/M}$ and the time-series properties of the latter. Here we restrict attention to noise that is independent from the efficient price.

In the context of testing for jumps, Huang and Tauchen (2005) and Andersen, Bollerslev, and Dobrev (2007), use staggered returns in the calculation of the bipower and tri-power variation to eliminate the correlation of two consecutive returns stemming from an *iid* noise process and therefore alleviate the effect of microstructure noise. The realized staggered multipower measures are defined as

$$sMPV_{t,M,k}(p) = \mu_{2/p}^{-p} \frac{M}{M-p(1+k)+1} \sum_{i=0}^{M-p(1+k)} \prod_{j=0}^{p-1} |r_{t_{i+j*(1+k)}}|^{2/p},$$

$$sMPQ_{t,M,k}(p) = \mu_{4/p}^{-p} \frac{M^2}{M-p(1+k)+1} \sum_{i=0}^{M-p(1+k)} \prod_{j=0}^{p-1} |r_{t_{i+j*(1+k)}}|^{4/p},$$

where $\mu_{2/p}^{-p} = \pi^{-1/2} 2^{p/2} \Gamma(\frac{p+1}{2})$.

[2] See O'Hara (1995) and Hasbrouck (2007) for details.

Similarly to the staggered multipower variation measures, one can define staggered median realized volatility as follows:

$$sMedRV_{t,M,k} = c_v \left(\frac{M}{M-2}\right) \sum_{i=0}^{M-4} \text{med}(|r_{t_i}|, |r_{t_{i+(1+k)}}|, |r_{t_{i+2(1+k)}}|)^2,$$

$$sMedRQ_{t,M,k} = c_q \left(\frac{M^2}{M-2}\right) \sum_{i=0}^{M-4} \text{med}(|r_{t_i}|, |r_{t_{i+(1+k)}}|, |r_{t_{i+2(1+k)}}|)^4,$$

where $c_v = \frac{\pi}{6-4\sqrt{3}+\pi}$ and $c_q = \frac{3\pi}{9\pi+72-52\sqrt{3}}$. These staggered measures can be readily plugged into the J tests for jumps in place of the non-staggered measures and applied to testing for jumps in the presence of *iid* or moving-average type of noise.

Alternative approaches are provided by Jiang and Oomen (2008) and Podolskij and Ziggel (2008) who modify their test statistics to account for the presence of *iid* noise. We do not provide the formulae here to save space and refer the reader to the original papers.

To simulate the behavior of the tests for jumps in the presence of microstructure noise, we let the efficient price be governed by the LL1F stochastic volatility model with moderate mean reversion. Following Andersen, Dobrev, and Schaumburg (2009), we model the microstructure noise as an AR(1) process with parameter $\rho \in \{0, 0.95\}$. These scenarios thus include the case of an *iid* noise ($\rho = 0$) typically found in transaction prices as well as a persistent noise process ($\rho = 0.95$) suitable for modeling quotes (Hasbrouck (1999)). We consider two noise-to-signal ratios: large, with $\omega^2/IV = 0.01$ and moderate with $\omega^2/IV = 0.001$ and implement the following tests:

- Staggered multipower variation ratio test (BNS) using sBV, sTPV and sQPV to estimate the integrated variance and sTPQ or sQPQ to estimate the integrated quarticity;
- Staggerd median realized volatility ration test (ADS) using sMinRQ and sMedRQ to the estimated integrated quarticity;
- Swap variance ratio test robust to *iid* microstructure noise (JO);
- Truncated power variation test robust to *iid* microstructure noise (PZ), where, following the recommendation of PZ, we set $\tau = 0.05$, $\vartheta = 0.17$ and $c = 2.3, 3, 4$.

and consider sampling frequencies of 5s, 15s, 30s, 1min and 5 min. The main simulation results are summarized in Tables 5-8.

5.1 Size

Starting with the case of *iid* microstructure noise with moderate noise-to-signal ratio reported in the left panel of Table 5, we see that all four classes of tests possess very good size properties. Similar to the case of no microstructure noise the PZ test exhibits sensitivity to the choice of threshold.

We next introduce dependence into the microstructure noise by allowing it to follow a first-order autoregression with parameter 0.95 and set the noise-to-signal ratio back to

0.001 (moderate noise). It comes as no surprise that the tests based on staggered multipower measures are no longer immune to this type of noise. In fact, staggering can only help if the noise is of the moving-average type. The simulated size of the BNS and ADS tests decreases with the sampling frequency although the distortions are not as dramatic as the highly dependent noise process may suggest. Similar, but a more pronounced, effect is found for the swap variance test (JO). Increasing the variance of the noise implies large size distortions in the same directions.

Overall the best performing test in terms of size is the PZ test. Except for the relatively low frequency of 5 minutes, the PZ test exhibits empirical size very close to the nominal level across the scenarios considered here, as long as moderate or large threshold is used. This is quite remarkable especially in the case of large, highly dependent noise.

5.2 Power

The jump detection ability of the noise-robust tests is presented in Tables 6-8. We focus on the same jump scenarios as in Section 4.3. The most powerful test is again the PZ test but only for a small value of the threshold parameter. Increasing the threshold results into a sharp decrease in power, especially at lower sampling frequencies. The noise-robust JO test shows a very stable performance across the different specifications of the noise process and delivers better power than either BNS or ADS. The latter two work quite well when the noise has moderate variance ($\omega^2 = 0.001$).

Increasing the variance of the noise to 0.01 leads to a sharp drop in power of all tests. The effect is most pronounced when the jumps are small and frequent (Table 8).

6 Zero returns

It is well-known that prices do not change at equidistant points in time (see, for example, Engle and Russell (1998)). There generally tends to be more activity taking place in the market shortly after opening and towards the end of the trading session than around lunchtime. As a result, when sampling in calendar time some intraday returns may be equal to zero, which may in turn distort the inference about jumps.

To see this, consider the BNS test based on bipower variation. Since the latter is calculated as a sum of products of two consecutive returns, one zero intraday return will set two summands equal to zero as opposed to the realized volatility, where only one summand will be knocked out of the sum of squared returns. As a result, the difference between RV and BV will be upward biased and consequently the test based on this difference oversized. It is clear that this effect will be more pronounced for tests based on multipower variations of higher order. This observation has motivated ADS to propose the median realized volatility as a more robust measure of integrated variance and quarticity in the presence of infrequent trading.

To shed more light on the impact of zeros returns on the alternative tests for jumps, we consider the following simple model of sparse sampling. The efficient price follows the LL1F model as before but it is only observed at random points in time, whereby the durations between consecutive observations are assumed to be independently exponentially distributed with mean $\phi(t)$. To calibrate the mean duration as a function of the time of day, $\phi(t)$, we follow **?** and fit a cubic spline to the price durations of the S&P 500 futures contract between 2003-2007 (the data is described in greater detail below).

The average duration between consecutive price changes is found to be about 15 seconds and we observe a large difference between average morning durations (10 seconds) and lunchtime durations (20 seconds). We use this diurnal pattern function throughout the simulations but re-scale it such that the mean price duration over the course of the trading day is equal to either 5 seconds, 15 seconds or 30 seconds. This allows us to study the impact of different levels of nontrading on the size of the alternative tests for jumps. We study the same test statistics as in the size simulations.

The main simulation results are summarized in Table 9. Consistent with intuition, the most affected by the presence of zero intraday returns are the tests based on multipower variations (BNS and CPR). In case of the CPR test, the problem is further exacerbated by the presence of the bipower variation in construction of the threshold. The negative bias in the bipower variation due to zero returns tends to reduce the threshold value and as we have seen in the simulations before this translates into more frequent false rejections of the null hypothesis. The tests based on the median realized volatility (ADS) are slightly more robust to the presence of zeros although the gains are not very large, at least not for the type of infrequent trading considered here.

The impact of zeros on the swap variance test (JO) tends to be much smaller. It operates primarily through the realized sixticity appearing in the denominator of the JO test statistics. The downward bias of the realized sixticity implies more frequent rejections that consistent with the nominal significance level. Similarly effected is the ASJ test, which requires the use of multipower variation to estimate the quarticity appearing in the denominator of the test statistics.

Overall the best performance in terms of empirical size in the presence of zero returns is afforded by the PZ test, as long as one chooses a sufficiently large constant c when calculating the threshold. Even when the mean duration of nontrading is large (30 seconds), the PZ test provides reasonable inference at frequencies as high as 2 minutes, at which all other test already suffer from substantial size distortions.

7 Diurnal volatility

The last challenge that the test for jumps will be subjected to in this paper is the deterministic diurnal volatility component. It is well-known that unconditional intraday volatility tends to exhibit an asymmetric U-shaped pattern. It is typically highest in the morning, drops significantly around lunchtime and then picks up again towards market close. We follow Hasbrouck (1999) and model the unconditional intraday volatility as

$$E(\sigma_t) = A(e^{-at} + e^{-b(1-t)}). \tag{2}$$

We set $A = 0.0795$, $a = -2.5$ and $b = -3.2$. With these parameter values the mean volatility approximately equals 2, 0.5 and 1 in the morning, mid-day and evening, respectively.

We first study the size. The results, reported in Table 15, resemble those obtained under the LL2F model. All tests are substantially oversized, those that require thresholding are very sensitive to the choice of the threshold, and the performance of PZ and LM test actually deteriorates as the sampling frequency increases. The BNS and ADS tests seem to be the best choice similarly to the case of LL2F with no diurnal volatility.

In terms of the ability of the test to detect jumps, the ranking is again similar to the case of LL1F. To save space, we only report the scenario of moderate jump intensity and size noting that the results for the other jump scenarios are qualitatively similar. The PZ test with a low value of threshold delivers highest power, closely followed by the LM and JO tests. Recall, however, that the two former tests are severely oversized.

8 Empirical Application

In this section, the we apply aforementioned jump tests to empirical data. The analysis is carried out using high frequency data from three markets: the foreign exchange inter-dealer market, the equity futures market, and the stock market. Specifically, the currency pairs of EUR/USD and USD/JPY are analyzed together with the S&P 500 Futures Index and equity data from five corporations, listed on the New York Stock Exchange (NYSE) namely, Citigroup, IBM, McDonald's, Disney and General Electric. For the cleaning of the data, we follow the procedure outlined by Barndorff-Nielsen, Hansen, Lunde, and Shephard (2008b). Below, we give a brief description of the various datasets employed.

8.1 Data Description and Preliminaries

8.1.1 Foreign exchange

We study the EUR/USD spot exchange rate during the period between January 4, 2000 and May 31, 2007. The mid-quotes are extracted from the Electronic Broking Services (EBS) Market Data database, which is currently the larger of the two electronic venues that make up the inter-dealer spot FX market, after Reuters. In addition, EBS has become the major trading platform for the two most traded currency pairs, the USD/JPY and the EUR/USD. This data has been only recently made available to academic researchers. As is customary in the literature, observations recorded between 21:00 GMT on Friday and 21:00 GMT on Sunday as well as holidays are discarded. In addition, days with low trading activity due to public and bank holidays are also excluded from the data. This leaves us with 1820 days in the sample.

8.1.2 Individual stocks

We collect equity data for five corporations listed on the New York Stock Exchange (NYSE), namely McDonald's and IBM over the period between July 2, 2001 and December 29, 2005, which yields a total of 1126 days. Only the mid-quotes recorded between 9:30 EST and 16:00 EST are considered. The data is extracted from the Trades And Quotes (TAQ) database of NYSE.

8.1.3 S&P 500 Futures

We focus on the most liquid (front) S&P 500 futures contract over the period from June 2, 2003 to December 28, 2007. Only observations between the hours of 8:30 EST to 15:00 EST are considered and holidays are omitted, leaving us with a total of 1174 days in the sample. This data was obtained from TickData Inc.

8.1.4 Preliminaries

All four series of high-frequency prices have been filtered using the approach proposed by **?**. In order to gain some intuition about the level of the jump component, microstructure noise and flat trading in the various datasets, we provide in Figure 1 in the signature plots of the average daily realized volatility, medium realized volatility, bi-power and tri-power variation. The level of microstructure noise appears to be higher for the equity and futures data than for the FX data, for which a 30-second sampling frequency seems adequate to avoid the impact of microstructure noise on the estimation of volatility. In case of the individual stocks and S&P 500 futures, frequencies between 2 and 5 minutes deliver stable results.

While the difference between the realized volatility and and the jump-robust measures (MedRV, BV, TPV) provides information about the magnitude of the jump component, this information is only reliable at moderate and small sampling frequencies due to the presence of zero returns. The signatures plots reveal that the equity data has a larger proportion of zero returns than the FX and futures data, and the jump-robust measures become severely downward biased for frequencies higher than a minute. For the foreign exchange and futures data, on the other hand, MedRV and BV seem to stabilize already at the 30-second frequency, whereas, not surprisingly, TPV requires slightly lower frequency to avoid the effect of zero returns.

Tables 18-20 provide some descriptive statistics for the various datasets. In Table 18, the average number of quotations per day is reported for each calendar year in the sample together with the percentage of unique quotes. We observe that the foreign exchange data has the largest average number of quotes per day relative to the operating hours in the market, while the futures data has the smallest number of quotes. In addition, for all datasets, with the exception of the futures data, the number of quotes increases significantly over the sample period, the increase being higher for the individual stocks. The futures data possesses the highest percentage of unique quotes amongst all the data examined. Only less than 3% of consecutive quotes in the future data are identical. The number is much larger for the foreign exchange data and equity data. For these datasets, the number of consecutive duplicate quotes increases significantly over time.

Table 19 reports the average durations between successive quotes. When accounting for duplicate quotes, the quote arrival rates are quite similar across the datasets, except for S&P 500 futures which shows average duration about twice as high as the other assets considered. Finally, Table 20 reports the yearly percentage of zero returns for five different sampling frequencies, extending from 30 seconds to 15 minutes. Consistent with the volatility signature plots, the percentage of zero returns decreases significantly with the sampling frequency. Nonetheless, the proportion of zeros remains nontrivial even at the relatively low and commonly employed 5-minute frequency implying possible positive bias in the estimated contribution of jumps to the overall variation in the assets' prices.

8.2 Results of Jump Tests

We now apply the set of jumps test to the data. For each asset and day of the sample we test for the presence of jumps at three different sampling frequencies: 1 minute, 5 minute and 15 minutes. To balance the trade-off between size and power we use intermediate values for the threshold parameters in the CPR, ASJ and PZ tests. We adopt the 1% significance level throughout.

We summarize the results in a set of confusion matrices: we calculate the proportion of days that each test identified jumps individually and also for each pair of the tests the proportion of jump days detected by the corresponding tests jointly. The latter is again interpreted as a measure of the degree of agreement among the different tests. The confusion matrices are reported in Tables 21 and 22.

The empirical results are consistent with those obtained in the simulations. There is a clear association between the number of zero returns and the proportion of days identified as jump days. As expected, the tests that are affected the most are the BNS and LM ones. The proportion of detected jump days also radically increases with the sampling frequency, except for the JO test which exhibits relatively stable performance.

The largest proportion of jump days is detected for the EUR/USD exchange rate. At the 5 minute frequency, the tests indicate between 23 and 53% of jump days in the sample, if we ignore the ASJ test which is known to have low power and the LM test which is substantially biased due to zero returns. There are much less jumps in the S&P 500 future index, for about 10 to 15 % of the days in the sample do the test signal the presence of jumps. The IBM stock price jumps relatively infrequently (6 - 11%), while McDonalds exhibits similar behavior to S&P 500 futures.

9 Conclusion

This paper aims to evaluate the performance of seven different approaches developed for testing for the presence of jumps in asset price processes. Extensive simulation results examining the size and power of the tests under different data generating scenarios reveal that there is no clear “winner”. The performance of each test depends on a particular scenario: while some tests perform well in the absence of frictions, they face considerable difficulties when confronted with noisy data. The tests employing thresholds suffer further from a trade-off between size and power; small threshold improves the ability to detect true jumps but at the same time increases the probability of spurious jump detection in periods when no jumps occurred. Further research is therefore called for to address the choice of the threshold in order to balance this trade-off.

An important feature of the jump tests that we document in this paper is their sensitivity to the presence of zero intraday returns. A natural remedy to this problem is to resort to tick-time sampling (see e.g. **?**, 2006 for a discussion of the benefits of tick-time sampling for estimation of volatility). Nevertheless, the limit theories underlying the tests studied here are derived under the assumption of equidistant sampling and hence it remains to be shown whether they remain valid when sampling time becomes random. Further research will almost surely tackle this issue.

The conclusions from the empirical application based on individual stock, foreign exchange and equity futures data conform to those of the Monte Carlo simulation. They show that the test statistics are very sensitive to the presence of zero returns and microstructure noise. It is therefore very important in empirical work to be aware of this problem and interpret the results accordingly.

References

Aït-Sahalia, Y. (2002): "Telling from Discrete Data Whether the Underlying Continuous-Time Model is a Diffusion," *Journal of Finance*, 57(5), 2075–2112.

Aït-Sahalia, Y., and J. Jacod (2009): "Testing for Jumps in a Discretely Observed Process," *Annals of Statistics*, 37(1), 184–222.

Andersen, T. G., and T. Bollerslev (1998): "Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts," *International Economic Review*, 39(4), 885–905.

Andersen, T. G., T. Bollerslev, and F. X. Diebold (2004): "Some Like it Smooth, and Some Like it Rough: Untangling Continuous and Jump Components in Measuring, Modeling and Forecasting Asset Return Volatility," Working Paper, University of Pennsylvania.

Andersen, T. G., T. Bollerslev, and D. Dobrev (2007): "No-Arbitrage Semi-Martingale Restrictions for Continuous-Time Volatility Models subject to Leverage Effects, Jumps and i.i.d. Noise: Theory and Testable Distributional Implications," *Journal of Econometrics*, 138(1), 125–180.

Andersen, T. G., D. Dobrev, and E. Schaumburg (2009): "Jump Robust Volatility Estimation," Northwestern University.

Bakshi, G., C. Cao, and Z. Chen (1997): "Empirical Performance of Alternative Option Pricing Models," *Journal of Finance*, 52(5), 20032049.

Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde, and N. Shephard (2008a): "Designing Realized Kernels to Measure the Ex-post Variation of Equity Prices in the Presence of Noise," *Econometrica*, 76(6), 14811536.

——— (2008b): "Realized Kernels in Practice: Trades and Quotes," *Journal of Econometrics*, 4, 1–32.

Barndorff-Nielsen, O. E., and N. Shephard (2004): "Power and Bipower Variation with Stochastic Volatility and Jumps," *Journal of Financial Econometrics*, 2(1), 1–48.

——— (2006): "Econometrics of Testing for Jumps in Financial Economics Using Bipower Variation," *Journal of Financial Econometrics*, 4(1), 1–30.

Bates, D. (2000): "Post-87 Crash Fears in the S&P 500 Futures Option Market," *Journal of Econometrics*, 94, 181–238.

Carr, P., and L. Wu (2003): "What Type of Process Underlies Options? A Simple Robust Test," *Journal of Finance*, 58(6), 2581–2610.

Chernov, M., A. R. Gallant, E. Ghysels, and G. Tauchen (2003): "Alternative Models for Stock Price Dynamics," *Journal of Econometrics*, 116, 225–257.

Corsi, F., D. Pirino, and R. Renò (2008): "Volatility Forecasting: The Jumps Do Matter," Università di Siena.

Das, S. R. (2002): "The Surprise Element: Jumps in Interest Rates," *Journal of Econometrics*, 106(1), 2765.

Eberlein, E., and S. Raible (1999): "Term Structure Models Driven by General Lévy Processes," *Mathematical Finance*, 9(1), 31–53.

Engle, R. F., and J. Russell (1998): "Autoregressive Conditional Duration: A New Model for Irregularly-spaced Transaction Data," *Econometrica*, 66, 1127–1162.

Hansen, P. H., and A. Lunde (2006): "Realized Variance and Market Microstructure Noise," *Journal of Business and Economic Statistics*, 24(2), 127–161.

Hasbrouck, J. (1999): "The Dynamics of Discrete Bid and Ask Quotes," *Journal of Finance*, 54(6), 21092142.

——— (2007): *Empirical Market Microstructure*. Oxford University Press.

Huang, X., and G. Tauchen (2005): "The Relative Contribution of Jumps to Total Price Variation," *Journal of Financial Econometrics*, 3(4), 456–499.

Jiang, G. J., and R. C. A. Oomen (2008): "Testing for Jumps When Asset Prices Are Observed With Noise A Swap Variance Approach," *Journal of Econometrics*, 144(2), 352–370.

Johannes, M. (2004): "The Statistical and Economic Role of Jumps in Continuous-Time Interest Rate Models," *Journal of Financ*, 59(1), 227–260.

Lee, S. S., and P. A. Mykland (2008): "Jumps in Finacial Markets: A New Nonparametric Test and Jump Dynamics," *Review of Financial Studies*, 21, 2535–2563.

Mancini, C. (2001): "Disentangling the jumps of the diffusion in a geometric jumping Brownian motion," *Giornale dell'Instituto degli Attuari*, (LXIV), 19–47.

Merton, R. K. (1976): "Option Pricing When Underlying Stock Returns are Discontinuous," *Journal of Financial Economics*, 3, 125–144.

Neuberger, A. (1994): "The Log Contract: A New Instrument to Hedge Volatility," *Journal of Portfolio Management*, 20, 74–80.

O'Hara, M. (1995): *Market Microstructure Theory*. Blackwell Publishing Limited.

Pan, J. (2002): "The Jump-Risk Premia Implicit in Options: Evidence from an Integrated Time Series Study," *Journal of Financial Economics*, 63.

Piazzesi, M. (2005): "Bond Yields and the Federal Reserve," *Journal of Political Economy*, 113(2), 311344.

Podolskij, M., and D. Ziggel (2008): "New Tests for Jumps: A Threshold-Based Approach," CREATES Research Papers, 34, School of Economics and Management, University of Aarhus.

Rosiński, J. (2001): "Series Representations of Lévy Processes from the Perspective of Point Processes," in *Lévy Processes-Theory and Applications*, ed. by O. E. Barndorff-Nielsen, T. Mikosch, and S. Resnick. Boston: Birkhäser.

Todorov, V. (2007): "Estimation of Continuous-Time Stochastic Volatility Models with Jumps using High-Frequency Data," Duke University.

Todorov, V., and G. Tauchen (2008): "Volatility Jumps," Northwestern University.

Veraart, A. E. D. (2008): "Inference for the Jump Part of Quadratic Variation of It Semimartingales," CREATES Research Paper 2008-17.

A Tables

μ	0.030
β_0	0.000
β_1	0.125
α_v	$\{-0.137e-2, -0.100, -1.386\}$

Table 1: Parameters of the LL1F model used in the simulations.

μ	0.030
β_0	-1.200
β_1	0.040
β_2	1.500
α_{v1}	-0.137e-2
α_{v2}	-1.386
β_{v2}	0.250
ρ_1	-0.300
ρ_2	-0.300

Table 2: Parameters of the LL2F model used in the simulations.

μ	0.030
β_0	0.000
β_1	0.125
α_v	$\{-0.137e-2, -0.100, -1.386\}$
λ	2.500
(c, α)	{(2.424,0.400),(1.635,0.800),(0.894,1.200),(0.325,1.600)}

Table 3: Parameters of the LLIA model used in the simulations.

λ	{0.1,0.5,1.0,1.5,2.0}
σ^2	{0.5,1.0,1.5,2.0,2.5}

Table 4: Jump intensity and variance of jump size for LL1F-FAJ model.

α	{0.1,0.5}
k	{0.0119,0.0161}
c	{0.125,0.4}
λ	{0.015,0.015}

Table 5: Parameters of the jump process in the LL1F-IAJ model.

	LL1F					LL2F					LLIA				
	30 s	1 min	2 min	5 min	15 min	30 s	1 min	2 min	5 min	15 min	30 s	1 min	2 min	5 min	15 min
A. Multipower variation ratio tests (BNS)															
$J_{bv,tpq}$	1.17	1.19	1.28	1.23	1.46	2.01	2.30	2.78	3.70	4.55	1.10	1.12	1.25	1.37	1.50
$J_{bv,qpq}$	1.19	1.20	1.30	1.26	1.53	2.13	2.56	3.08	4.17	5.10	1.11	1.13	1.27	1.37	1.55
$J_{tpv,tpq}$	0.96	0.94	0.96	0.89	0.80	1.62	1.79	2.08	2.59	3.06	0.93	0.94	0.91	0.90	0.87
$J_{tpv,qpq}$	1.01	0.99	1.05	1.00	0.96	1.86	2.19	2.53	3.45	4.05	0.99	1.01	0.98	0.98	1.00
$J_{qpv,tpq}$	0.78	0.70	0.63	0.46	0.25	1.29	1.34	1.32	1.43	1.29	0.83	0.70	0.67	0.45	0.24
$J_{qpv,qpq}$	0.89	0.84	0.86	0.80	0.63	1.64	1.85	2.01	2.57	2.85	0.95	0.86	0.89	0.76	0.72
B. Threshold bipower variation ratio tests (CPR)															
$J^{(3)}_{tbv,ttpq}$	1.37	1.37	1.41	1.37	1.54	2.60	3.21	4.35	7.11	8.97	1.28	1.29	1.39	1.53	1.56
$J^{(4)}_{tbv,ttpq}$	1.17	1.20	1.28	1.24	1.47	2.05	2.39	2.99	4.32	5.84	1.10	1.14	1.26	1.38	1.51
$J^{(5)}_{tbv,ttpq}$	1.17	1.19	1.28	1.23	1.46	2.01	2.32	2.85	3.87	5.05	1.10	1.12	1.26	1.37	1.50
C. MedRV ratio tests (ADS)															
$J_{medrv,minrq}$	1.04	1.07	1.12	1.03	1.20	2.10	2.66	3.30	4.22	4.76	1.01	1.05	1.10	1.14	1.23
$J_{medrv,medrq}$	1.00	1.04	1.08	1.01	1.19	1.73	2.16	2.71	3.63	4.77	1.01	1.03	1.08	1.13	1.24
D. Swap variance ratio tests (JO)															
SwV_{qps}	1.25	1.37	1.65	2.55	5.42	2.81	3.60	5.14	8.31	15.01	1.16	1.44	1.77	2.58	5.43
SwV_{sps}	1.27	1.45	1.89	3.20	7.26	3.26	4.50	6.56	11.43	21.81	1.21	1.54	1.94	3.05	7.32
E. Power variation ratio test (ASJ)															
$S_3(4,2)$	0.69	0.56	0.38	0.20	0.04	5.56	3.83	1.90	0.36	0.02	0.75	0.57	0.43	0.21	0.05
$S_4(4,2)$	0.41	0.32	0.25	0.15	0.04	1.79	1.05	0.62	0.14	0.01	0.47	0.34	0.30	0.17	0.04
$S_5(4,2)$	0.41	0.32	0.25	0.15	0.04	0.64	0.43	0.29	0.06	0.01	0.46	0.32	0.29	0.17	0.04
F. Threshold power variation difference test (PZ)															
$S_{2.3}(2)$	1.84	2.50	3.54	4.61	5.82	66.03	60.75	53.51	41.36	26.67	1.75	2.39	3.32	4.70	5.71
$S_3(2)$	1.04	0.95	0.92	0.87	0.86	21.50	19.14	16.04	11.64	7.52	0.93	0.93	0.99	0.91	0.85
$S_4(2)$	1.08	0.95	0.91	0.87	0.52	3.79	3.41	2.93	2.21	1.62	1.03	0.92	1.00	0.78	0.45
G. Local volatility test (LM)															
LM_{bv}	1.21	1.17	1.29	1.41	1.36	67.29	59.86	51.24	39.80	26.88	1.21	1.09	1.09	1.03	0.96

Table 6: **Size**. Simulated size of the tests for jumps for 1% nominal level.

	J_{bv}	$J_{tbv}^{(3)}$	$J_{tbv}^{(5)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_5(4,2)$	$S_{2.3}(2)$	$S_4(2)$	LM_{bv}
Sampling frequency: 1 minute										
J_{bv}	**88.68**									
$J_{tbv}^{(3)}$	88.68	**89.88**								
$J_{tbv}^{(5)}$	88.68	89.44	**89.44**							
J_{medrv}	87.93	88.56	88.37	**88.94**						
SwV_{qps}	88.18	89.27	88.91	88.43	**91.20**					
$S_3(4,2)$	74.36	74.93	74.81	74.53	75.50	**75.86**				
$S_5(4,2)$	71.34	71.91	71.78	71.53	72.43	72.52	**72.52**			
$S_{2.3}(2)$	88.56	89.73	89.31	88.83	91.12	75.79	72.50	**93.97**		
$S_4(2)$	87.74	88.54	88.35	87.99	88.75	74.74	71.78	88.81	**88.85**	
LM_{bv}	88.52	89.69	89.27	88.79	91.12	75.75	72.50	93.60	88.81	**93.68**
Sampling frequency: 5 minutes										
J_{bv}	**78.27**									
$J_{tbv}^{(3)}$	78.27	**82.01**								
$J_{tbv}^{(5)}$	78.25	80.18	**80.18**							
J_{medrv}	77.49	79.87	79.07	**80.47**						
SwV_{qps}	77.77	81.23	79.68	79.70	**84.19**					
$S_3(4,2)$	1.64	2.08	1.93	1.93	2.20	**2.33**				
$S_5(4,2)$	1.62	1.99	1.91	1.91	1.99	1.99	**1.99**			
$S_{2.3}(2)$	78.21	81.92	80.12	80.33	83.75	2.33	1.99	**87.57**		
$S_4(2)$	75.92	77.18	77.09	76.84	77.09	1.76	1.76	77.18	**77.26**	
LM_{bv}	78.14	81.84	80.05	80.18	83.41	2.29	1.99	85.85	77.18	**86.00**
Sampling frequency: 15 minutes										
J_{bv}	**64.01**									
$J_{tbv}^{(3)}$	63.97	**70.75**								
$J_{tbv}^{(5)}$	63.95	66.89	**66.93**							
J_{medrv}	62.17	66.64	64.37	**67.98**						
SwV_{qps}	63.22	69.64	66.07	66.83	**75.77**					
$S_3(4,2)$	0.00	0.00	0.00	0.00	0.02	**0.04**				
$S_5(4,2)$	0.00	0.00	0.00	0.00	0.00	0.02	**0.02**			
$S_{2.3}(2)$	63.91	70.67	66.83	67.81	74.68	0.02	0.00	**79.76**		
$S_4(2)$	58.22	58.98	58.93	58.41	58.85	0.00	0.00	59.00	**59.08**	
LM_{bv}	63.05	69.56	65.97	66.79	72.96	0.02	0.00	75.46	58.93	**76.59**

Table 7: **Confusion matrix**. LL1F-FAJ with medium mean reversion. $\lambda = 0.1$, $\sigma^2 = 2.5$. Significance level: 1%.

	J_{bv}	$J_{tbv}^{(3)}$	$J_{tbv}^{(5)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_5(4,2)$	$S_{2.3}(2)$	$S_4(2)$	LM_{bv}
Sampling frequency: 1 minute										
J_{bv}	**86.15**									
$J_{tbv}^{(3)}$	86.15	**88.02**								
$J_{tbv}^{(5)}$	86.15	87.11	**87.11**							
J_{medrv}	84.97	85.90	85.57	**86.39**						
SwV_{qps}	85.14	86.71	86.02	85.30	**88.85**					
$S_3(4,2)$	68.84	69.59	69.30	68.84	69.74	**70.57**				
$S_5(4,2)$	61.60	62.25	62.01	61.57	62.37	62.89	**62.89**			
$S_{2.3}(2)$	86.03	87.88	86.99	86.28	88.71	70.55	62.87	**92.78**		
$S_4(2)$	85.00	86.08	85.79	85.04	85.88	69.30	62.13	86.42	**86.51**	
LM_{bv}	86.00	87.84	86.96	86.24	88.65	70.52	62.86	92.13	86.41	**92.23**
Sampling frequency: 5 minutes										
J_{bv}	**72.87**									
$J_{tbv}^{(3)}$	72.86	**77.74**								
$J_{tbv}^{(5)}$	72.87	75.15	**75.16**							
J_{medrv}	70.47	73.32	72.02	**74.21**						
SwV_{qps}	70.66	74.87	72.83	71.72	**78.41**					
$S_3(4,2)$	1.81	2.31	2.10	1.99	2.40	**2.62**				
$S_5(4,2)$	1.51	1.79	1.72	1.70	1.77	1.82	**1.82**			
$S_{2.3}(2)$	72.72	77.59	75.01	74.08	78.04	2.60	1.80	**85.27**		
$S_4(2)$	68.93	70.45	70.32	68.69	69.20	1.98	1.67	70.45	**70.54**	
LM_{bv}	72.11	76.70	74.37	73.31	76.77	2.54	1.80	81.17	70.39	**81.31**
Sampling frequency: 15 minutes										
J_{bv}	**52.87**									
$J_{tbv}^{(3)}$	52.84	**60.51**								
$J_{tbv}^{(5)}$	52.85	55.89	**55.92**							
J_{medrv}	48.86	52.93	50.52	**55.54**						
SwV_{qps}	50.02	56.58	52.83	52.34	**65.71**					
$S_3(4,2)$	0.00	0.01	0.00	0.01	0.02	**0.03**				
$S_5(4,2)$	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**			
$S_{2.3}(2)$	52.79	60.37	55.84	55.38	64.28	0.02	0.00	**73.98**		
$S_4(2)$	42.80	43.56	43.45	41.60	42.63	0.00	0.00	43.60	**43.69**	
LM_{bv}	50.24	57.22	53.23	52.56	58.88	0.02	0.00	64.14	43.11	**65.01**

Table 8: **Confusion matrix**. LL1F-FAJ with medium mean reversion. $\lambda = 1.0$, $\sigma^2 = 1.5$. Significance level: 1%.

	J_{bv}	$J_{tbv}^{(3)}$	$J_{tbv}^{(5)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_5(4,2)$	$S_{2.3}(2)$	$S_4(2)$	LM_{bv}
Sampling frequency: 1 minute										
J_{bv}	**67.93**									
$J_{tbv}^{(3)}$	67.92	**72.87**								
$J_{tbv}^{(5)}$	67.92	70.41	**70.41**							
J_{medrv}	65.80	68.46	67.42	**69.62**						
SwV_{qps}	64.00	67.79	66.19	65.17	**72.57**					
$S_3(4,2)$	44.39	46.31	45.51	44.99	45.52	**48.46**				
$S_5(4,2)$	30.09	31.35	30.91	30.44	31.17	32.48	**32.48**			
$S_{2.3}(2)$	67.75	72.66	70.23	69.39	72.36	48.35	32.43	**84.78**		
$S_4(2)$	63.84	66.19	65.55	64.47	64.56	44.80	30.89	66.91	**67.06**	
LM_{bv}	67.67	72.54	70.15	69.31	72.25	48.32	32.43	83.23	66.90	**83.50**
Sampling frequency: 5 minutes										
J_{bv}	**36.89**									
$J_{tbv}^{(3)}$	36.87	**45.81**								
$J_{tbv}^{(5)}$	36.88	39.51	**39.52**							
J_{medrv}	33.23	38.23	35.11	**40.37**						
SwV_{qps}	31.70	38.34	34.11	34.11	**45.92**					
$S_3(4,2)$	0.87	1.42	1.12	1.16	1.57	**1.98**				
$S_5(4,2)$	0.55	0.83	0.73	0.76	0.82	0.86	**0.86**			
$S_{2.3}(2)$	36.56	45.46	39.20	39.91	44.90	1.93	0.83	**63.45**		
$S_4(2)$	24.68	25.89	25.77	24.91	24.76	0.89	0.66	25.93	**26.25**	
LM_{bv}	35.31	43.57	37.94	38.38	42.61	1.83	0.83	54.65	25.89	**55.39**
Sampling frequency: 15 minutes										
J_{bv}	**13.92**									
$J_{tbv}^{(3)}$	13.89	**18.84**								
$J_{tbv}^{(5)}$	13.90	14.62	**14.65**							
J_{medrv}	10.55	13.28	11.08	**16.22**						
SwV_{qps}	11.30	15.18	12.00	13.08	**27.62**					
$S_3(4,2)$	0.00	0.00	0.00	0.00	0.01	**0.04**				
$S_5(4,2)$	0.00	0.00	0.00	0.00	0.00	0.03	**0.03**			
$S_{2.3}(2)$	13.65	18.49	14.38	15.67	23.81	0.01	0.00	**36.62**		
$S_4(2)$	4.81	4.93	4.88	4.60	4.83	0.00	0.00	4.93	**5.17**	
LM_{bv}	10.78	15.13	11.50	12.72	17.87	0.01	0.00	23.04	4.79	**25.21**

Table 9: **Confusion matrix**. LL1F-FAJ with medium mean reversion. $\lambda = 2.0$, $\sigma^2 = 0.5$. Significance level: 1%.

	J_{bv}	$J_{tbv}^{(3)}$	$J_{tbv}^{(5)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_5(4,2)$	$S_{2.3}(2)$	$S_4(2)$	LM_{bv}
Sampling frequency: 1 minute										
J_{bv}	**10.73**									
$J_{tbv}^{(3)}$	10.71	**12.28**								
$J_{tbv}^{(5)}$	10.73	11.38	**11.40**							
J_{medrv}	9.32	10.01	9.72	**11.06**						
SwV_{qps}	9.26	10.34	9.87	9.51	**13.40**					
$S_3(4,2)$	6.15	6.66	6.45	6.29	6.97	**7.97**				
$S_5(4,2)$	5.24	5.66	5.50	5.36	5.94	6.36	**6.36**			
$S_{2.3}(2)$	9.86	11.30	10.53	10.19	12.30	7.53	6.07	**19.48**		
$S_4(2)$	8.65	9.28	9.09	8.83	9.41	6.27	5.45	9.51	**10.30**	
LM_{bv}	9.73	11.12	10.40	10.09	12.20	7.48	6.07	17.07	9.50	**17.58**
Sampling frequency: 5 minutes										
J_{bv}	**5.71**									
$J_{tbv}^{(3)}$	5.69	**6.94**								
$J_{tbv}^{(5)}$	5.71	6.06	**6.08**							
J_{medrv}	4.50	5.19	4.79	**6.19**						
SwV_{qps}	4.46	5.42	4.83	4.91	**8.95**					
$S_3(4,2)$	0.13	0.25	0.19	0.22	0.30	**0.52**				
$S_5(4,2)$	0.12	0.20	0.17	0.19	0.20	0.32	**0.32**			
$S_{2.3}(2)$	5.00	6.22	5.37	5.48	7.13	0.38	0.21	**13.03**		
$S_4(2)$	3.48	3.66	3.65	3.61	3.67	0.16	0.16	3.68	**4.34**	
LM_{bv}	4.55	5.61	4.91	5.06	6.40	0.36	0.20	8.52	3.68	**9.09**
Sampling frequency: 15 minutes										
J_{bv}	**3.40**									
$J_{tbv}^{(3)}$	3.38	**4.21**								
$J_{tbv}^{(5)}$	3.40	3.59	**3.61**							
J_{medrv}	2.17	2.65	2.34	**3.66**						
SwV_{qps}	2.42	3.15	2.63	2.78	**8.89**					
$S_3(4,2)$	0.00	0.00	0.00	0.00	0.00	**0.06**				
$S_5(4,2)$	0.00	0.00	0.00	0.00	0.00	0.05	**0.05**			
$S_{2.3}(2)$	3.06	3.87	3.27	3.26	5.53	0.00	0.00	**10.31**		
$S_4(2)$	1.34	1.37	1.37	1.33	1.37	0.00	0.00	1.37	**1.85**	
LM_{bv}	1.98	2.72	2.19	2.36	3.52	0.00	0.00	4.28	1.36	**5.08**

Table 10: **Confusion matrix**. LL1F-IAJ with medium mean reversion and jump activity index equal to 0.5. Significance level: 1%.

	iid noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.01$				
	5 s	15 s	30 s	1 min	5 min	5 s	15 s	30 s	1 min	5 min	5 s	15 s	30 s	1 min	5 min
A. Staggered multipower variation ratio tests (BNS)															
$J_{sbv,stpq}$	0.96	1.04	1.10	1.17	1.44	0.67	0.64	0.89	1.10	1.35	0.47	0.40	0.80	1.07	1.27
$J_{sbv,sqpq}$	0.97	1.05	1.11	1.16	1.46	0.67	0.66	0.90	1.12	1.36	0.47	0.41	0.80	1.09	1.30
$J_{stpv,stpq}$	0.93	0.96	0.90	0.93	1.05	0.71	0.67	0.83	0.88	1.05	0.50	0.45	0.76	0.86	0.95
$J_{stpv,sqpq}$	0.96	1.04	0.93	0.99	1.18	0.71	0.69	0.88	0.95	1.16	0.52	0.48	0.80	0.92	1.04
$J_{sqpv,stpq}$	0.87	0.88	0.75	0.70	0.55	0.68	0.64	0.69	0.71	0.58	0.54	0.46	0.70	0.71	0.50
$J_{sqpv,sqpq}$	0.93	0.95	0.88	0.88	0.93	0.73	0.72	0.79	0.87	0.88	0.56	0.54	0.81	0.85	0.90
B. Staggered median realized volatility ratio tests (ADS)															
$J_{smedrv,sminrv}$	1.04	1.04	1.08	1.10	1.29	0.70	0.74	0.91	1.01	1.32	0.45	0.50	0.79	0.99	1.16
$J_{smedrv,smedrv}$	1.04	1.05	1.05	1.10	1.31	0.70	0.75	0.90	0.99	1.28	0.44	0.51	0.79	1.00	1.20
C. Swap variance ratio tests (JO)															
SwV^*	1.03	1.11	1.09	1.14	1.28	0.21	0.40	0.73	1.09	1.25	0.03	0.12	0.37	0.96	1.24
D. Threshold power variation difference test (PZ)															
$S_{2.3}^{noise}(2)$	1.68	2.32	2.87	2.59	0.80	1.96	2.75	3.05	2.71	0.88	1.65	1.79	1.49	1.22	0.87
$S_3^{noise}(2)$	1.01	0.91	1.04	1.00	0.72	1.02	0.95	1.00	1.03	0.80	0.97	0.96	0.96	0.98	0.84
$S_4^{noise}(2)$	1.00	0.90	1.03	0.99	0.72	1.02	0.94	0.98	1.02	0.80	0.96	0.96	0.96	0.98	0.84

Table 11: **Size**. Simulated size of the test for jumps robust to microstructure noise for 1% nominal. LL1F model with medium mean reversion and microstructure noise. The autoregressive parameter of the noise process equals 0.95.

	iid noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.01$				
	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$
Sampling frequency: 15 sec															
J_{sbv}	**84.26**					**86.49**					**66.21**				
J_{smed}	83.20	**84.32**				85.66	**86.59**				64.22	**66.43**			
SwV^*	83.98	84.04	**89.23**			86.33	86.33	**90.51**			65.94	65.96	**77.22**		
$S^*_{2.3}(2)$	83.96	84.08	88.76	**91.46**		86.23	86.31	89.88	**91.52**		66.15	66.23	77.06	**86.45**	
$S^*_4(2)$	82.78	82.84	85.05	85.31	**85.31**	84.12	84.08	85.03	85.11	**85.11**	65.62	65.68	74.81	76.39	**76.39**
Sampling frequency: 30 sec															
J_{sbv}	**83.94**					**85.09**					**65.38**				
J_{smed}	83.10	**84.50**				84.00	**85.35**				63.04	**65.74**			
SwV^*	83.69	84.16	**88.84**			84.71	85.03	**89.29**			64.73	65.15	**76.31**		
$S^*_{2.3}(2)$	83.65	84.20	88.07	**90.49**		84.73	84.93	88.44	**90.43**		65.05	65.56	76.06	**85.31**	
$S^*_4(2)$	81.62	81.89	82.96	83.16	**83.16**	82.03	82.01	82.90	83.06	**83.06**	64.38	64.91	73.00	74.36	**74.36**
Sampling frequency: 1 min															
J_{sbv}	**83.77**					**83.47**					**66.41**				
J_{smed}	82.92	**84.34**				82.39	**83.81**				64.38	**67.46**			
SwV^*	83.29	83.89	**87.77**			83.12	83.35	**87.85**			65.60	66.51	**76.21**		
$S^*_{2.3}(2)$	83.06	83.65	86.25	**87.97**		82.92	83.25	86.37	**87.93**		66.13	67.10	75.64	**82.47**	
$S^*_4(2)$	77.53	77.67	77.83	77.95	**77.95**	77.65	77.57	77.97	78.17	**78.17**	63.45	63.79	67.30	68.07	**68.07**

Table 12: **Confusion matrices**. LL1F-FAJ with medium mean reversion and microstructure noise. $\lambda = 0.1$, $\sigma^2 = 2.5$. Significance level: 1%.

	iid noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.01$				
	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$
Sampling frequency: 15 sec															
J_{sbv}	**82.22**					**85.22**					**59.46**				
J_{smed}	80.79	**82.05**				84.20	**85.31**				56.97	**59.90**			
SwV^*	81.53	81.30	**87.93**			84.70	84.80	**89.54**			58.21	58.48	**72.27**		
$S^*_{2.3}(2)$	82.03	81.84	87.49	**91.38**		84.97	85.07	88.93	**91.29**		59.34	59.78	72.18	**85.29**	
$S^*_4(2)$	80.45	80.31	83.02	83.78	**83.78**	82.15	82.20	83.09	83.50	**83.50**	58.70	58.97	68.84	71.88	**71.88**
Sampling frequency: 30 sec															
J_{sbv}	**82.53**					**83.64**					**58.38**				
J_{smed}	81.27	**82.79**				82.42	**83.94**				55.75	**59.37**			
SwV^*	81.70	81.89	**87.53**			82.83	83.10	**87.97**			56.77	57.68	**70.88**		
$S^*_{2.3}(2)$	82.22	82.43	86.69	**89.88**		83.27	83.55	87.16	**89.97**		58.16	59.17	70.63	**84.10**	
$S^*_4(2)$	79.11	79.12	80.26	80.87	**80.87**	79.52	79.54	80.23	80.74	**80.74**	56.97	57.75	66.29	69.23	**69.23**
Sampling frequency: 1 min															
J_{sbv}	**81.45**					**81.49**					**58.49**				
J_{smed}	80.00	**81.88**				80.02	**81.94**				56.05	**60.27**			
SwV^*	80.29	80.70	**85.89**			80.27	80.66	**85.92**			56.39	57.96	**70.54**		
$S^*_{2.3}(2)$	80.70	81.13	84.38	**87.12**		80.78	81.19	84.32	**87.15**		58.08	59.81	69.60	**80.12**	
$S^*_4(2)$	72.79	72.70	72.92	73.60	**73.60**	72.77	72.68	72.74	73.50	**73.50**	53.42	54.39	57.83	59.41	**59.41**

Table 13: **Confusion matrices**. LL1F-FAJ with medium mean reversion and microstructure noise. $\lambda = 1.0$, $\sigma^2 = 1.5$. Significance level: 1%.

	iid noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.001$					AR(1) noise, $\omega^2 = 0.01$				
	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$	J_{sbv}	J_{smed}	SwV^*	$S^*_{2.3}(2)$	$S^*_4(2)$
Sampling frequency: 15 sec															
J_{sbv}	**57.80**					**66.24**					**13.01**				
J_{smed}	54.46	**58.58**				63.80	**67.04**				10.18	**14.29**			
SwV^*	54.95	55.46	**70.84**			63.65	64.21	**75.22**			10.66	11.53	**29.29**		
$S^*_{2.3}(2)$	57.20	57.97	69.66	**80.83**		65.47	66.11	73.34	**80.43**		12.78	14.02	29.03	**63.29**	
$S^*_4(2)$	50.70	51.13	55.92	58.13	**58.13**	54.65	54.83	55.83	57.07	**57.07**	10.47	11.20	20.88	25.78	**25.78**
Sampling frequency: 30 sec															
J_{sbv}	**58.19**					**61.32**					**12.43**				
J_{smed}	55.30	**59.80**				58.80	**62.95**				9.40	**14.14**			
SwV^*	54.76	55.97	**68.85**			57.90	59.20	**70.27**			9.78	10.83	**26.84**		
$S^*_{2.3}(2)$	57.28	58.69	66.81	**76.98**		60.26	61.75	68.07	**76.88**		11.93	13.56	26.30	**59.52**	
$S^*_4(2)$	45.59	46.15	47.62	49.27	**49.27**	46.66	47.02	47.64	49.01	**49.01**	8.77	9.57	16.90	20.97	**20.97**
Sampling frequency: 1 min															
J_{sbv}	**55.27**					**55.70**					**13.38**				
J_{smed}	52.34	**57.46**				52.92	**58.11**				10.28	**15.10**			
SwV^*	50.93	52.72	**64.41**			51.48	53.28	**64.66**			10.05	11.31	**27.52**		
$S^*_{2.3}(2)$	52.97	54.82	59.81	**68.37**		53.40	55.29	60.15	**68.36**		12.20	13.82	25.23	**47.65**	
$S^*_4(2)$	29.89	30.12	30.15	31.13	**31.13**	30.03	30.22	30.33	31.19	**31.19**	5.17	5.64	7.84	9.20	**9.20**

Table 14: **Confusion matrices**. LL1F-FAJ with medium mean reversion and microstructure noise. $\lambda = 2.0$, $\sigma^2 = 0.5$. Significance level: 1%.

	Nontrading duration: 5 sec					Nontrading duration: 15 sec					Nontrading duration: 30 sec				
	30 s	1 min	2 min	5 min	15 min	30 s	1 min	2 min	5 min	15 min	30 s	1 min	2 min	5 min	15 min
A. Multipower variation ratio tests (BNS)															
$J_{bv,tpq}$	8.80	1.75	1.41	1.34	1.44	94.76	5.73	1.53	1.30	1.48	100.0	99.71	14.63	1.74	1.49
$J_{bv,qpq}$	8.88	1.76	1.42	1.36	1.51	95.11	5.81	1.54	1.33	1.55	100.0	99.80	14.99	1.75	1.53
$J_{tpv,tpq}$	7.81	1.33	1.03	0.96	0.79	98.13	4.78	1.11	0.86	0.80	100.0	99.98	14.72	1.10	0.85
$J_{tpv,qpq}$	8.11	1.39	1.11	1.08	0.90	98.34	5.05	1.23	0.98	0.90	100.0	99.98	15.97	1.23	0.93
$J_{qpv,tpq}$	7.23	1.00	0.70	0.52	0.25	99.36	4.17	0.82	0.45	0.21	100.0	100.0	14.35	0.57	0.26
$J_{qpv,qpq}$	7.80	1.18	0.95	0.83	0.61	99.38	4.74	1.14	0.77	0.58	100.0	100.0	16.82	0.96	0.68
B. Threshold bipower variation ratio tests (CPR)															
$J^{(3)}_{tbv,ttpq}$	10.19	2.04	1.61	1.47	1.54	95.57	6.62	1.78	1.47	1.56	100.0	99.77	16.11	1.98	1.57
$J^{(4)}_{tbv,ttpq}$	8.87	1.77	1.41	1.34	1.44	94.82	5.81	1.55	1.31	1.48	100.0	99.73	14.77	1.76	1.50
$J^{(5)}_{tbv,ttpq}$	8.81	1.74	1.41	1.34	1.44	94.77	5.73	1.53	1.30	1.48	100.0	99.71	14.65	1.74	1.49
C. MedRV ratio tests (ADS)															
$J_{medrv,minrq}$	6.70	1.53	1.18	1.21	1.13	60.08	4.28	1.48	1.20	1.24	100.0	80.67	8.05	1.46	1.24
$J_{medrv,medrq}$	6.50	1.48	1.15	1.21	1.13	57.80	4.13	1.44	1.16	1.27	100.0	77.90	7.60	1.43	1.25
B. Swap variance ratio tests (JO)															
SwV_{qps}	2.03	1.70	1.90	2.54	5.33	6.03	2.46	2.01	2.73	5.49	36.49	13.23	4.55	2.99	5.52
SwV_{sps}	2.27	1.79	2.17	3.15	7.25	7.95	2.74	2.21	3.31	7.26	58.24	20.44	5.58	3.66	7.51
E. Power variation ratio test (ASJ)															
$S_3(4,2)$	1.25	0.61	0.36	0.16	0.05	6.37	0.81	0.43	0.19	0.04	84.98	8.50	0.64	0.21	0.04
$S_4(4,2)$	0.61	0.34	0.21	0.11	0.05	1.85	0.42	0.29	0.14	0.04	70.47	1.70	0.28	0.16	0.04
$S_5(4,2)$	0.55	0.33	0.21	0.11	0.05	1.20	0.39	0.27	0.14	0.04	42.95	0.89	0.25	0.16	0.04
F. Threshold power variation difference test (PZ)															
$S_{2.3}(2)$	5.20	3.22	3.66	4.55	5.82	26.96	6.08	4.05	4.82	5.94	99.99	65.17	13.57	5.88	5.87
$S_3(2)$	1.02	0.94	0.99	0.88	0.93	2.56	1.06	0.93	0.89	0.90	95.71	13.06	1.50	0.92	0.88
$S_4(2)$	0.93	0.98	0.95	0.81	0.48	1.08	0.98	0.92	0.78	0.49	41.71	1.50	0.84	0.90	0.58
G. Local volatility test (LM)															
LM_{bv}	4.79	1.80	1.41	1.45	1.36	72.70	9.72	2.56	1.52	1.33	99.99	55.96	76.16	19.14	13.12

Table 15: **Size**. Simulated size of the tests for jumps for 1% nominal. LL1F model with medium mean reversion and infrequent trading.

	30 s	1 min	2 min	5 min	15 min
A. Multipower variation ratio tests (BNS)					
$J_{bv,tpq}$	2.01	2.52	3.07	4.20	5.27
$J_{bv,qpq}$	2.15	2.71	3.38	4.55	5.67
$J_{tpv,tpq}$	1.70	1.95	2.43	3.26	3.94
$J_{tpv,qpq}$	1.95	2.30	3.03	4.01	4.66
$J_{qpv,tpq}$	1.36	1.45	1.67	1.95	1.68
$J_{qpv,qpq}$	1.73	2.03	2.55	3.33	3.61
B. Threshold bipower variation ratio tests (CPR)					
$J^{(3)}_{tbv,ttpq}$	2.45	3.16	4.21	6.91	7.61
$J^{(4)}_{tbv,ttpq}$	2.06	2.60	3.20	4.74	5.91
$J^{(5)}_{tbv,ttpq}$	2.02	2.54	3.10	4.32	5.43
C. MedRV ratio tests (ADS)					
$J_{medrv,minrq}$	2.11	2.68	3.26	4.12	4.66
$J_{medrv,medrq}$	1.93	2.33	2.90	3.97	5.04
B. Swap variance ratio tests (JO)					
SwV_{qps}	2.33	3.39	5.01	8.48	15.66
SwV_{sps}	2.71	4.20	6.44	11.79	22.81
E. Power variation ratio test (ASJ)					
$S_3(4,2)$	5.00	3.56	1.91	0.64	0.03
$S_4(4,2)$	1.36	0.91	0.52	0.23	0.03
$S_5(4,2)$	0.53	0.35	0.20	0.07	0.03
F. Threshold power variation difference test (PZ)					
$S_{2.3}(2)$	64.79	57.44	48.49	36.04	24.00
$S_3(2)$	10.53	10.39	9.39	7.94	6.31
$S_4(2)$	1.24	1.20	1.19	1.17	1.18
G. Local volatility test (LM)					
LM_{bv}	92.96	79.21	29.91	28.35	11.48

Table 16: **Size**. Simulated size of the tests for jumps for 1% nominal. LL1F model with medium mean reversion and deterministic diurnal volatility.

	J_{bv}	$J_{tbv}^{(3)}$	$J_{tbv}^{(5)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_5(4,2)$	$S_{2.3}(2)$	$S_4(2)$	LM_{bv}
Sampling frequency: 1 minute										
J_{bv}	**84.96**									
$J_{tbv}^{(3)}$	84.96	**86.90**								
$J_{tbv}^{(5)}$	84.96	86.06	**86.06**							
J_{medrv}	83.36	84.40	84.04	**85.02**						
SwV_{qps}	83.30	84.71	84.20	83.22	**86.66**					
$S_3(4,2)$	28.34	28.79	28.61	28.42	28.71	**29.60**				
$S_5(4,2)$	9.64	9.85	9.78	9.64	9.85	10.01	**10.01**			
$S_{2.3}(2)$	84.95	86.89	86.05	84.99	86.65	29.56	10.00	**96.83**		
$S_4(2)$	83.99	85.38	84.90	83.88	85.07	28.81	9.90	86.58	**86.60**	
LM_{bv}	84.93	86.87	86.03	84.99	86.65	29.55	10.00	96.18	86.58	**97.59**
Sampling frequency: 5 minutes										
J_{bv}	**72.92**									
$J_{tbv}^{(3)}$	72.92	**78.85**								
$J_{tbv}^{(5)}$	72.91	75.35	**75.35**							
J_{medrv}	69.43	72.26	70.90	**73.28**						
SwV_{qps}	69.84	74.29	72.04	69.86	**77.99**					
$S_3(4,2)$	0.23	0.35	0.30	0.28	0.36	**0.46**				
$S_5(4,2)$	0.18	0.23	0.22	0.22	0.23	0.24	**0.24**			
$S_{2.3}(2)$	72.86	78.75	75.29	73.21	77.60	0.45	0.24	**89.97**		
$S_4(2)$	68.48	70.72	70.19	67.91	69.31	0.27	0.22	70.84	**70.90**	
LM_{bv}	72.00	77.42	74.42	72.30	76.22	0.43	0.24	84.17	70.74	**84.87**
Sampling frequency: 15 minutes										
J_{bv}	**56.18**									
$J_{tbv}^{(3)}$	56.14	**63.81**								
$J_{tbv}^{(5)}$	56.16	58.91	**58.95**							
J_{medrv}	50.60	53.96	51.87	**57.09**						
SwV_{qps}	53.17	59.88	55.86	53.86	**72.53**					
$S_3(4,2)$	0.00	0.00	0.00	0.00	0.00	**0.01**				
$S_5(4,2)$	0.00	0.00	0.00	0.00	0.00	0.00	**0.00**			
$S_{2.3}(2)$	56.10	63.64	58.87	56.90	69.37	0.01	0.00	**78.80**		
$S_4(2)$	44.58	45.62	45.38	42.81	44.96	0.00	0.00	45.74	**45.80**	
LM_{bv}	52.34	59.19	55.07	52.55	61.74	0.00	0.00	65.89	45.04	**67.07**

Table 17: **Confusion matrix**. LL1F-FAJ with medium mean reversion and deterministic diurnal volatility. $\lambda = 1.0$, $\sigma^2 = 1.5$. Significance level: 1%.

	Quotes/day							
Asset	2000	2001	2002	2003	2004	2005	2006	2007
EUR/USD	32582	30928	30412	36080	37065	38000	37766	36401
	(63.20)	*(61.54)*	*(54.32)*	*(61.32)*	*(63.51)*	*(54.11)*	*(47.32)*	*(40.25)*
S& P 500	–	–	–	2042	1890	1749	1665	1819
	–	–	–	*(99.11)*	*(98.38)*	*(98.28)*	*(97.85)*	*(98.12)*
McDonald's	–	2512	3007	5450	6782	8785	–	–
	–	*(42.41)*	*(37.51)*	*(32.14)*	*(29.38)*	*(28.45)*	–	–
IBM	–	5034	6213	7223	8338	8777	–	–
	–	*(53.36)*	*(38.52)*	*(49.93)*	*(50.50)*	*(44.28)*	–	–

Table 18: The average number of quotes per day for each for the years in the sample and for each dataset. The percentage of observations for which the quoted price was different from the previous one is given in parentheses.

Asset		2000	2001	2002	2003	2004	2005	2006	2007
EUR/USD	Quotes	6.33	5.91	6.43	4.95	4.64	4.84	5.28	4.58
	Quotes ($\Delta_p \neq 0$)	9.09	9.03	10.87	7.57	7.01	8.18	9.83	10.81
S & P 500	Quotes	–	–	–	11.69	13.56	15.36	16.15	14.59
	Quotes ($\Delta_p \neq 0$)	–	–	–	11.79	14.56	17.46	18.67	16.19
MacDonald's	Quotes	–	9.39	7.92	4.36	3.51	2.69	–	–
	Quotes ($\Delta_p \neq 0$)	–	22.39	21.29	13.87	12.22	9.73	–	–
IBM	Quotes	–	4.64	3.89	3.31	2.83	2.67	–	–
	Quotes ($\Delta_p \neq 0$)	–	8.73	10.08	6.78	5.67	6.10	–	–

Table 19: Average durations between successive quotes.

Asset	Sampl. Freq.	2000	2001	2002	2003	2004	2005	2006	2007
EUR/USD	1 min	27.21	28.98	33.48	25.53	23.52	26.75	31.37	37.15
	5 min	11.06	12.08	14.69	10.03	9.27	11.02	13.34	16.52
	15 min	5.77	6.57	8.04	5.46	4.94	5.90	7.41	9.80
S & P 500	1 min	–	–	–	15.93	16.46	17.13	17.63	15.50
	5 min	–	–	–	7.63	8.08	8.28	8.41	7.30
	15 min	–	–	–	4.75	4.89	5.36	5.87	5.35
McDonald's	1 min	–	23.34	21.81	27.48	28.41	25.24	–	–
	5 min	–	7.99	6.56	10.23	10.48	8.73	–	–
	15 min	–	4.57	4.26	5.49	5.44	5.06	–	–
IBM	1 min	–	5.64	5.90	6.83	7.57	9.44	–	–
	5 min	–	2.68	2.25	3.01	2.90	3.66	–	–
	15 min	–	2.02	1.50	2.02	1.87	2.11	–	–

Table 20: Yearly percentages of zero returns.

	EUR/USD							S&P 500						
	J_{bv}	$J_{tbv}^{(4)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_3(2)$	LM_{bv}	J_{bv}	$J_{tbv}^{(4)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_3(2)$	LM_{bv}
Sampling frequency: 1 minute														
J_{bv}	**77.20**							**36.07**						
$J_{tbv}^{(4)}$	77.20	**86.45**						32.48	**22.15**					
J_{medrv}	56.84	59.64	**60.68**					16.61	17.86	**16.61**				
SwV_{qps}	26.20	28.43	22.97	**29.38**				8.09	9.83	6.47	**11.99**			
$S_3(4,2)$	22.14	24.97	19.63	11.40	**26.81**			4.00	4.96	3.24	2.30	**9.49**		
$S_3(2)$	67.46	76.03	55.28	28.31	25.86	**87.48**		10.73	12.91	8.26	8.09	4.43	**20.95**	
LM_{bv}	77.20	86.45	60.68	29.50	26.76	87.48	**100.0**	20.93	23.88	16.64	13.04	6.60	18.70	**57.29**
Sampling frequency: 2 minutes														
J_{bv}	**54.58**							**17.18**						
$J_{tbv}^{(4)}$	54.52	**65.06**						17.09	**19.65**					
J_{medrv}	37.65	41.97	**45.33**					9.20	10.60	**17.04**				
SwV_{qps}	20.36	24.86	18.85	**28.70**				6.22	7.69	5.79	**16.36**			
$S_3(4,2)$	13.68	16.28	12.40	8.90	**20.41**			1.11	1.71	1.28	1.28	**4.43**		
$S_3(2)$	44.61	54.01	39.04	26.81	19.19	**75.48**		6.73	8.80	6.98	7.07	2.56	**16.53**	
LM_{bv}	54.47	64.93	45.33	28.70	20.41	75.48	**98.94**	8.92	11.43	11.32	10.29	2.92	13.72	**38.25**
Sampling frequency: 5 minutes														
J_{bv}	**25.98**							**10.17**						
$J_{tbv}^{(4)}$	25.98	**35.34**						10.17	**11.45**					
J_{medrv}	16.35	19.73	**26.36**					4.68	5.13	**9.45**				
SwV_{qps}	12.57	17.11	12.97	**23.36**				4.60	5.56	4.00	**15.16**			
$S_3(4,2)$	3.28	4.91	3.62	3.00	**8.73**			0.00	0.00	0.00	0.09	**0.68**		
$S_3(2)$	20.08	29.26	20.97	20.36	7.23	**52.63**		3.83	4.70	3.32	6.13	0.17	**9.28**	
LM_{bv}	24.58	33.95	25.36	22.79	8.38	52.39	**83.13**	4.72	5.76	4.80	8.15	0.17	6.52	**24.19**

Table 21: **Confusion matrices**. Jump detection for EUR/USD and S&P500. Significance level: 1%.

	IBM							McDonald's						
	J_{bv}	$J_{tbv}^{(4)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_3(2)$	LM_{bv}	J_{bv}	$J_{tbv}^{(4)}$	J_{med}	SwV_{qps}	$S_3(4,2)$	$S_3(2)$	LM_{bv}
Sampling frequency: 1 minute														
J_{bv}	**15.33**							**52.32**						
$J_{tbv}^{(4)}$	15.33	**18.09**						52.14	**57.49**					
J_{medrv}	8.61	9.98	**15.80**					26.47	27.45	**29.58**				
SwV_{qps}	5.15	6.60	5.24	**11.99**				10.12	11.14	7.46	**12.34**			
$S_3(4,2)$	1.78	2.23	1.78	1.51	**7.65**			6.57	7.58	3.20	2.31	**10.39**		
$S_3(2)$	8.17	10.25	9.15	8.17	3.02	**24.42**		32.42	36.90	21.14	10.75	8.26	**53.56**	
LM_{bv}	13.60	16.31	13.69	11.18	5.28	23.08	**65.95**	48.30	53.32	28.53	12.25	10.29	52.68	**91.41**
Sampling frequency: 2 minutes														
J_{bv}	**7.40**							**19.16**						
$J_{tbv}^{(4)}$	7.40	**9.09**						19.07	**22.55**					
J_{medrv}	4.17	4.99	**8.43**					9.33	10.34	**13.77**				
SwV_{qps}	2.84	4.01	3.55	**11.81**				5.51	6.68	4.35	**10.04**			
$S_3(4,2)$	0.44	0.89	0.80	1.42	**4.52**			0.89	1.52	0.89	0.80	**4.62**		
$S_3(2)$	4.00	5.53	4.71	6.04	2.13	**15.46**		10.39	13.55	8.53	7.82	2.93	**30.11**	
LM_{bv}	5.37	6.90	6.35	9.12	2.68	13.24	**39.27**	14.58	17.74	11.09	8.68	3.58	26.30	**58.86**
Sampling frequency: 5 minutes														
J_{bv}	**5.88**							**9.27**						
$J_{tbv}^{(4)}$	5.88	**7.48**						9.27	**10.87**					
J_{medrv}	3.20	3.74	**6.13**					5.15	5.61	**8.26**				
SwV_{qps}	2.84	3.92	3.82	**11.9**				4.97	6.15	4.53	**12.87**			
$S_3(4,2)$	0.00	0.00	0.00	0.36	**0.80**			0.00	0.18	0.00	0.27	**0.89**		
$S_3(2)$	3.55	4.55	4.00	5.95	0.53	**10.48**		6.22	7.84	5.15	7.37	0.36	**15.01**	
LM_{bv}	3.85	5.29	4.20	7.78	0.54	8.41	**27.55**	5.81	7.26	5.28	8.94	0.63	12.08	**36.31**

Table 22: **Confusion matrices**. Jump detection for IBM and McDonald's mid-quotes. Significance level: 1%.

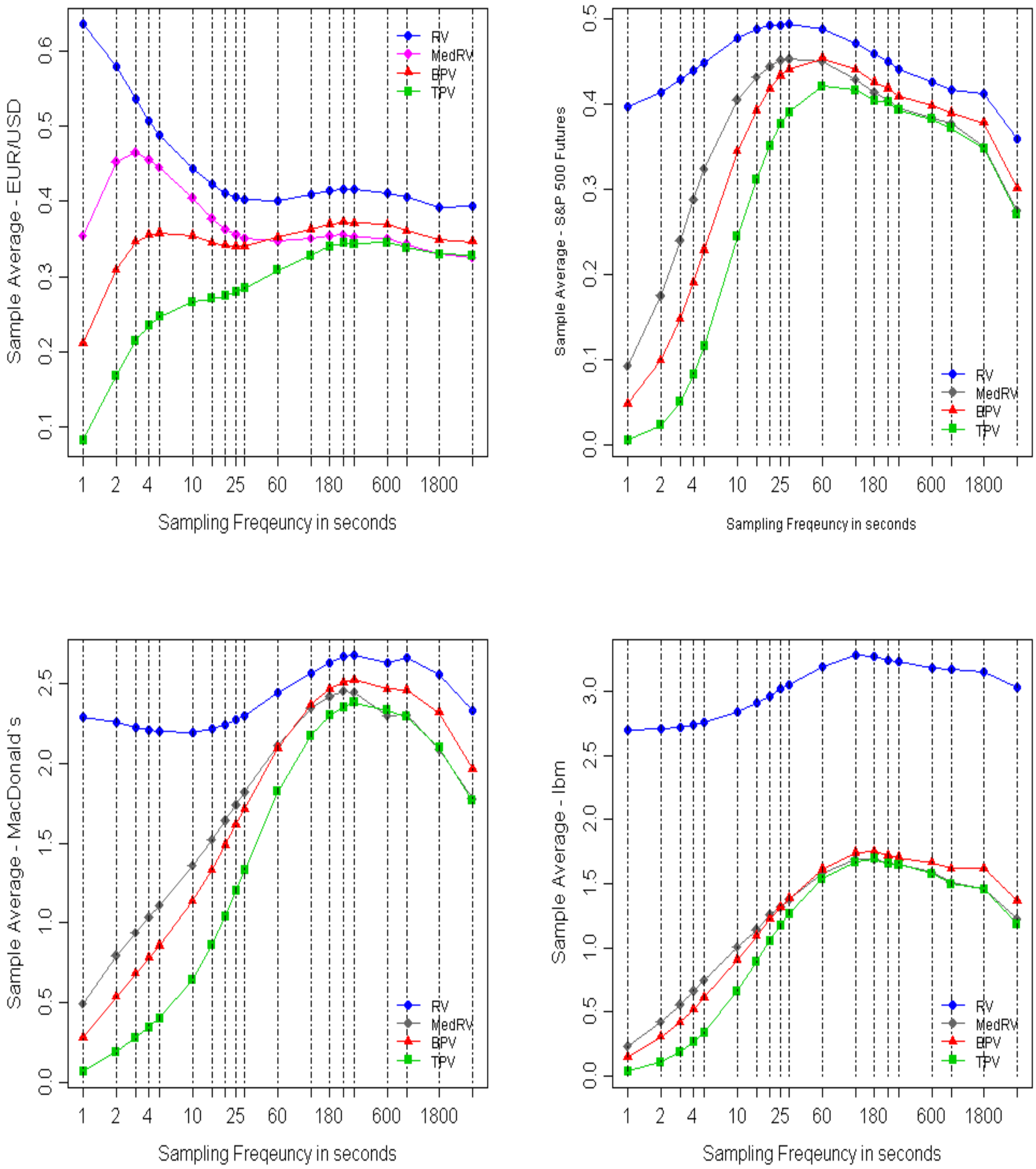

Figure 1: Volatility signature plots for EUR/USD, IBM and McDonalds based on mid-quotes, and S&P 500 futures contract based on transaction prices.