Journal of the Royal Statistical Society Series C: Applied Statistics, 2025, **00**, 1–19
https://doi.org/10.1093/jrsssc/qlaf025

Original Article

A new method for jump detection: analysis of jumps in the S&P 500 financial index

Khaldoun Khashanah[1], Jing Chen[2], Mike Buckle[3] and Alan Hawkes[3]

[1]School of Business, Stevens Institute of Technology, Hoboken, NJ, USA
[2]School of Mathematics, Cardiff University, Cardiff, UK
[3]School of Management, Swansea University, Swansea, UK

Address for correspondence: Jing Chen, School of Mathematics, Cardiff University, Cardiff, UK. Email: chenj60@cardiff.ac.uk

Abstract

Financial jumps have occurred more frequently with the advent of high-frequency trading enabled by technological advancement. Most existing jump detection methods that treat a jump as a singular, random, and isolated shock event were not designed to capture the clustering of jumps related to contagious behaviour, in which the occurrence of jumps increases the probability of further jumps soon after. This paper presents a new *Med*9 method that addresses the challenges of capturing both singular and consecutive jumps. This approach evaluates the size of individual returns with a measure of local volatility based on the median of consecutive absolute returns. We use this method to detect jumps in both S&P 500 and simulated time series, and compare its performance with several classic jump detection methods. Throughout, our *Med*9 consistently outperforms other approaches applied to both real and simulated financial return series. In addition, we demonstrate that the *Med*9 detection results are not biased or compromised by the intraday volatility pattern.

Keywords: contagion, financial series, jumps, S&P 500 index, volatility

JEL codes: G15, G17, C4, C5

1 Introduction

The increased complexity in financial market trading led by the advancement in technology has seen asset prices exhibit much more volatility. The flash crash on 6 May 2010 saw the S&P 500 index fall around 8.6% and the Dow Jones drop 998.5 points (see Figure 1). On 21 April 2015, the U.S. Department of Justice laid '22 criminal counts, including fraud and market manipulation' relating to the '2010 May Crash' against Navinder Singh Sarao, a trader from a modest background in the UK, for significantly 'spoofing the market'. Nanex, a Chicago-based data provider, has reported 180 flash crashes in Europe, on a scale similar to 2010, between January 2012 and January 2014 across all types of financial assets. The academic literature has also documented an increasing number of crashes including (Brogaard et al., 2018; Prodromou & Westerholm, 2022; Sornette & von der Becke, 2011), etc. They seem to share common features of sudden occurrence, rapid self-recovery, and contagion effects that is a tendency for jump clustering.

Mathematically, a financial jump can be measured using a particular detection method. But empirically, jumps may be defined differently. In ultra-high-frequency trading (e.g. milliseconds or nanoseconds), all price changes can be interpreted as jumps. But in other settings and more commonly, jumps are viewed as extreme but rare events that exhibit discontinuity in time series.[1].

[1] Literature also debates distinctions between jumps, structural breaks, bubbles, etc. and whether flash crashes should be deemed as jumps. However, there is no consensus on what constitutes a jump.

Received: July 5, 2021. **Revised:** March 13, 2025. **Accepted:** March 17, 2025

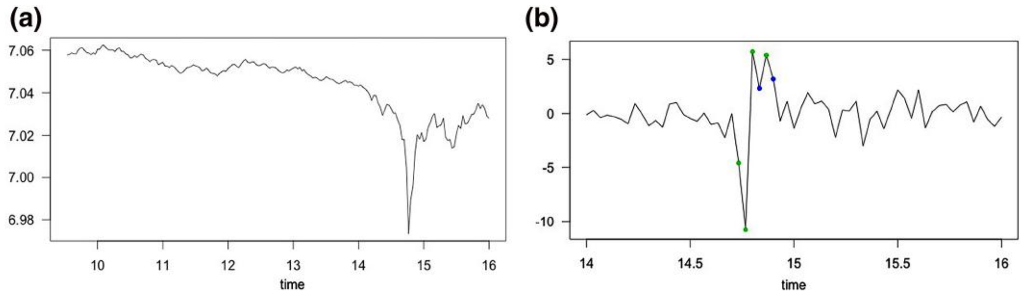

Figure 1. 2-Min log-prices and log-returns of S&P 500 on May 06, 2010. Panel (a) shows the log-prices and Panel (b) shows the standardized log-returns in the afternoon.

We consider jumps to be different from the underlying volatility process as Merton (1976) suggested that diffusion could not describe some rapid large changes in financial price series and introduced 'jumps' in the form of a Poisson process, in conjunction with Geometric Brownian motion. This provides a better description of financial time series and improves option pricing that is based on Brownian motion (see Bachelier, 2006; Black & Scholes, 1973; Samuelson, 1965). Since then, many studies have discussed the properties of Poisson or Lévy jumps, incorporated into a variety of diffusion processes (see Cont & Tankov, 2016 that documents an extensive treatment of their applications to financial problems). However, these models cannot accommodate contagion or clustering of jumps, which could compromise the model's fitness and ability for forecasting and pricing. Hence, in this paper, we introduce a new jump detection method that addresses these issues based on studying a long time series of the S&P500 index at intraday 2-minute frequency. We demonstrate our new detection method outperforms several commonly used existing methods.

A large body of literature examine 'macro-jumps' that occur infrequently within a day and detect jumps based on a statistical significance test with a null hypothesis of no jumps and at least one jump as the alternative at an aggregated level of each day (see Andersen et al., 2000, 2001, 2003, 2010; Barndorff-Nielsen & Shephard, 2004, 2006; Corsi et al., 2010; Huang & Tauchen, 2005; Podolkskij & Ziggel, 2010 and Christoffersen et al., 2008). As a result, some of these detection methods are not designed to capture 'micro-jumps' that occur at higher frequencies and average out over some intraday intervals, and thus do not point to a statistically significant 'macro-jump'. Some studies have suggested a recursive method (see Andersen et al., 2010) to detect additional jumps when one macro-jump at the day frequency is found. Bajgrowicz et al. (2016) use multiple testing procedures to filter out spurious jumps and relate 'true' jumps to news. Bolleslev et al. (2008) and Gilder et al. (2014) suggest that intraday jumps and contemporaneous jumps (co-jumps) rarely occur closely together.

However, as established facts of the increased complexity of trading have shown, micro-jumps, especially when they cluster within a trading day, become highly relevant for everyday risk management. Andersen et al. (2007b), Lee and Mykland (2008) and Andersen et al. (2010) have suggested various methods to identify intraday jumps, which are not based on examining the variance structure over one day cumulatively. However, these methods often failed to detect obvious jumps when multiple jumps occur close to each other, i.e. there are strong jump clustering effects.

We are therefore motivated to introduce a new jump detection method that can effectively capture jumps when contagion is present. We will demonstrate how our method performs better than three commonly used approaches that are designed to detect intraday jumps but have problems dealing with contagious jumps, namely *RV-BV*, *LM* and *ABD* tests, first introduced, respectively, by Barndorff-Nielsen and Shephard (2004, 2006), Lee and Mykland (2008) and Andersen et al. (2007a).

Section 2 introduces existing jump detection methods and highlights the distinction between global and local detection methods and the problems created by contagion of jumps. Section 3 introduces our method of jump detection based on a local approach using a median of 9 consecutive returns. Section 4 introduces the data whilst Section 5 shows the results of applying this new detection method to the data. Section 6 compares the new method with alternative detection methods. Section 7 concludes.

2 Existing jump detection methods

A generic jump diffusion model, assuming the jump process is separate from a Brownian process, can be written as

$$dp(t) = \mu(t) + \sigma(t)dW(t) + k(t)dq(t),\ t \geq 0, \tag{1}$$

where $p(t)$ is the continuous-time log-price process; $\mu(t)$ is the drift; $\sigma(t)$ is the instantaneous volatility process; $W(t)$ is a standard Brownian motion without drift. The counting process of jumps $q(t)$ is normalised, so that $dq(t) = 1$ represents a jump at time t and $dq(t) = 0$ otherwise; $k(t)$ is the jump size if a jump occurs at time t.

Each trading day t is divided into M equal time intervals and the log-return for the jth interval is defined as $r_{t,j} = p_{t,j} - p_{t,(j-1)}$ for $j = 1, \ldots, M$ and $t = 1, \ldots, T$. Looking for price jumps is now equivalent to looking for large absolute log-returns.

Andersen and Bolleslev (1998), Andersen et al. (2001) and Barndorff-Nielsen and Shephard (2002) define realized variance for day t as:

$$RV_t = \sum_{j=1}^{M} r_{t,j}^2,\ t = 1, \ldots, T. \tag{2}$$

RV_t estimates the daily increment to the quadratic variation for the underlying log-price process represented by equation (1). When $M \to \infty$,

$$RV_t \to_p \int_{t-1}^{t} \sigma^2(s)ds + \sum_{s \in (t-1,t]} k^2(s),\ t = 1, \ldots, T. \tag{3}$$

In the absence of jumps, the second term vanishes and realized variance (the first term) consistently estimates integrated variance. Barndorff-Nielsen and Shephard (2004, 2006) show that this first term of equation (3) can be separately identified using the realized bi-power variation defined as

$$BV_t = \mu_1^{-2} \sum_{j=2}^{M} |r_{t,j-1}|\,|r_{t,j}|,\ t = 1, \ldots, T, \tag{4}$$

where $\mu_1 = \sqrt{\frac{2}{\pi}}$. When $M \to \infty$, equation (4) can be written as.[2]

$$BV_t \to_p \int_{t-1}^{t} \sigma^2(s)ds,\ t = 1, \ldots, T. \tag{5}$$

Different jump detection approaches are discussed within Subsection 2.1 based on the difference of RV_t and BV_t over each day (the global approach); while Subsection 2.2 examines methods that compare the magnitude of individual log-returns with some volatility measure such as BV_t (the local approach).

2.1 The global detection approach

The contribution from jumps to total return variation on day t is estimated by the difference between equations (3) and (5). That is, as $M \to \infty$

$$RV_t - BV_t \to_p \sum_{s \in (t-1,t]} k^2(s),\ t = 1, \ldots, T. \tag{6}$$

[2] The rationale for this is that, for large M, there is at most one jump in any two adjacent intervals of length $1/M$. Since the contribution of each absolute return associated with the diffusion in the limit is very small, any product involving a jump return will be vanishingly small asymptotically.

In absence of jumps, $RV_t - BV_t$ could be negative or take on small positive values.

Barndorff-Nielsen and Shephard (2004, 2006) test the null hypothesis of no jumps against the alternative of at least one jump on day t using the test statistic

$$Z_t \equiv \sqrt{M}\frac{ln(RV_t) - ln(BV_t)}{((\mu_1^{-4} + 2\mu_1^{-2} - 5)TQ_t BV_t^{-2})^{1/2}} \rightarrow_d N(0, 1), \tag{7}$$

where TQ_t represents realized tri-power quarticity and is defined as

$$TQ_t \equiv M\mu_{4/3}^{-3}\sum_{j=3}^{M}|r_{t,j-2}|^{4/3}|r_{t,j-1}|^{4/3}|r_{t,j}|^{4/3},\ t = 1,\ \ldots,\ T, \tag{8}$$

where $\mu_{4/3}^{-3} = 2^{2/3}\Gamma(7/6)/\Gamma(1/2)$ with $\Gamma(\cdot)$ denoting the Gamma function.

The test, referred as the *BNS* test, is based on asymptotic theory, for interval sizes tending to zero and judged as significant at level α if the statistic Z_t exceeds the $1-\alpha$ fractile of the standard normal distribution. On days where there are no significant jumps, the continuous volatility component is measured by realized variance; otherwise, volatility is measured by realized bi-power variation.[3]

Huang and Tauchen (2005) carried out extensive simulations to compare some statistics that are asymptotically equivalent to that given by equations (7) and (8). Andersen et al. (2010) introduce a recursive procedure that removes the largest jump and recalculates the statistic to test further jumps until there are no more significant jumps.

Corsi et al. (2010) found that bi-power variation can sometimes be seriously inflated if there are large jumps in adjacent intervals and proposed *threshold bi-power variation* (*TBV*) that replaces large returns above a set threshold by zero to counter this. This is similar to bi-power variation, in which they replace large returns above a set threshold by zero. They further suggest the Corrected Threshold measure, *C-TBV*, to implement the replacement with a an expected absolute value from a Normal distribution instead of zero. They modify and improve the *BNS* test for no jumps on day t using a *C-TZ* statistic rather than the *BNSZ* statistic. However, the work does not look for individual jumps. Ferriani and Zoi (2020) introduced $s-CPR$ using the corrected measure *C-TZ* of Corsi et al. (2010) to carry out a sequential test similarly to Andersen et al. (2010).

2.2 The local detection approach

Andersen et al. (2007a) proposed a test (*ABD*) to examine individual returns divided by the square root of bi-power (continuous) variation calculated over the whole day. They used a statistic $|r_{t,j}|/\sqrt{BV_t/M}$ to capture and report jumps only at the end of each day.

In contrast, Lee and Mykland (2008) proposed a local jump detection method, the *LM* test, that calculates the ratio of the local return r_j to a volatility measure $\widehat{\sigma}(t_i)$ over an interval (t_{j-1}, t_j) over a period preceding that point. The statistic that tests for a jump in the interval (t_{j-1}, t_j) is defined as

$$L(i) = \frac{r_i}{\widehat{\sigma}(t_i)}, \tag{9}$$

where $\widehat{\sigma}^2(t_i) = \frac{1}{K-1}\sum_{j=i-K+1}^{i-1}|r_{j-1}|\,|r_j|$. K is the window size. If $L(i)$ exceeds a certain threshold, that observation is a jump. The local volatility measure is similar to the square root of bi-power variation, with a different scaling constant: it will therefore be liable to the same problems of inflation if large returns occur in adjacent intervals.

Tsai and Shackleton (2016) modified both *ABD* and *LM* tests, using so-called *interpolated Bipower Variation*. This uses a heterogeneous autoregression (HAR) model, combining estimates of bipower variation averaged over previous daily, weekly, monthly, and quarterly periods, and a current local estimate of bi-power variation. They also use a generalized extreme value (*GEV*) distribution to deal with extreme behaviour in the tail and allow for the effect of multiple comparisons on the test size.

[3] Bi-power variation can be generalised to multipowers, see (Kolokolov & Renò, 2017)

Recall LM uses a volatility divisor which is essentially the square root of a BV measure averaged over K previous returns. Ferriani and Zoi (2020) use an exponentially weighted moving average to replace observations beyond a threshold by zero. Then, they take a simple average of forward and backward versions of the weighted moving average to obtain volatility information from before and after the return being tested. They claim their modified method ($m - LM$) is more powerful.

2.3 The problem of contagious jumps

Global methods like Andersen et al. (2010) are useful for identifying jumps at the daily level and depend on the bi-power variation that is the product of absolute returns in successive intervals (see Equation 4). If jumps occur in two or more consecutive intervals, BV_t could be inflated and become larger than RV_t, signaling no jumps. Although ocal methods like (Lee & Mykland, 2008) consider the magnitude of price changes, and hence jumps, they need to be determined by conditioning on local volatility.

In reality, jump clustering is often seen in financial time series, especially at an intraday level. Recall the *flash crash day* on 6 May 2010 in Figure 1. Panels (a) and (b) show that morning prices are fairly quiet but in the afternoon a sharp decline and rebound (8.6%) occurred in S&P 500 returns. At least two jumps, one downward and one upward, happened within a short period between 14:42 and 15:54. It is associated with a run of six consecutive large absolute returns (at 2-min intervals) with standardized magnitudes, −4.59, −10.80, 5.72, 2.33, 5.39 and 3.20, that inflates the BV causing it to exceed the RV. We calculated $RV_t = 0.00238$, $BV_t = 0.00259$, so that $RV_t < BV_t$: a strong indication of no jumps on the day according to Andersen et al. (2010). Ferriani and Zoi (2020) also concluded that large neighbouring returns could mask the detection of jumps, hence implementing the $s - CPR$ and $m - LM$ tests before fitting Hawkes process models to 5-minute returns of S&P 500 and Euro StOXX 50 where jump clustering was present.

3 Jump detection based on medians

Andersen et al. (2012) introduced an estimator of realized volatiity, ($Med3RV_{t,M}$), based on a rolling sum of the median of three consecutive absolute returns as follows:

$$Med3RV_{t,M} = \frac{\pi}{6 - 4\sqrt{3} + \pi}\left(\frac{M}{M-2}\right)\sum_{j=2}^{M-1} med(|r_{t,j-1}|\,|r_{t,j}|\,|r_{t,j+1}|)^2. \tag{10}$$

This estimator is claimed to be robust to the presence of single jumps in the absolute returns and against occasional zero returns when the market lacks of activity. However, it has some limitations. As Ferriani and Zoi (2020) suggested, jump clustering tends to occur at frequencies higher than 5-min, and it takes time to fully unfold: e.g. the downward jump on the *flash crash day* took place over approximately 12 min (six 2-min intervals). Hence, $Med3RV_{t,M}$ would not fully cover the entire course of the occurrence of the contagious jump. Therefore, we extend Andersen et al. (2012) and generalise this estimator and use it, not only to measure realized volatility, but also to find jumps that occur in the presence of contagion.

Essentially, with M equally spaced log-returns $\{r_{t,j}\}$ on day t, we suggest a family of jump-robust estimators of daily realized volatility $MedkRV_{t,M}$, based on running medians of k absolute returns and defined by

$$MedkRV_{t,M} = \frac{M}{M+1-k}\sum_{j=(k+1)/2}^{M-(k-1)/2} LVk_{t,j} \tag{11}$$

$$LVk_{t,j} = c_k med(|r_{t,j-(k-1)/2}| \cdots |r_{t,j}| \cdots |r_{t,j+(k-1)/2}|)^2, \tag{12}$$

where $c_k = 1.41936$, 1.62360, 1.74332, and 1.82184 for $k = 3$, 5, 7, and 9, respectively. The factors c_k are chosen so that $MedkRV_{t,M}$ is an unbiased estimator of daily realized volatility in the event that returns were to be distributed as $i.i.d.N(0, \sigma^2)$.[4] $LVk_{t,j}$ estimates the local volatility in the neighbourhood of the jth interval of day t. $MedkRV_{t,M}$ averages these local volatility over day t. Furthermore, in any block of k consecutive returns it is capable of eliminating up to 1, 2, 3, or 4 large returns for $k = 3, 5, 7$, or 9, respectively. Hence, it is useful when jumps cluster. We can also derive the *standardized returns* for which each return is divided by a measure of volatility based on all returns of day t. For example, the individual standardized returns for $k = 9$ are written as

$$r^*_{t,j} = \frac{r_{t,j}}{\sqrt{Med9RV_{t,M}/M}}, \tag{13}$$

Empirically, analysts can choose a suitable critical threshold upon which any standardized return whose absolute value exceeds that threshold would be identified as a jump: we discuss possible choices in Section 5. Our *Med*9 also allows for identification of multiple jumps in a day, which relaxes the assumption that jumps are rare events.

4 S&P 500 data

We use the 1-min S&P 500 prices from Thomson Reuters ™ Tick History from 3 January 2006 to 13 March 2015. Data outside the normal trading hours (09:30–16:00) were deleted. Other data exclusion include 15 half-trading days just before or after major public holidays; 2 when Hurricane Sandy caused exchange closure; and a further 8 days that contain blocks of 5 or more consecutive minutes with missing data. A small number of other minutes with missing data, mostly isolated minutes but with at most four consecutive missing minutes, had data values inserted by backfilling. This cleaning process provided us with complete data for 2,285 trading days.

We then converted the time series into 2-min intervals to avoid problems such as microstructure noise, consistent with Andersen et al. (2012). We omit the first 2-min interval each trading day, as data burning at the start of trading are established facts. Subsequently, we obtain 195 log-prices and 194 log-returns for each trading day (between 9:32 and 16:00, see Figure 2). We can clearly identify the 2008 financial crisis and several large absolute values (jumps) in panel (b).

5 Jump detection for S&P 500

We report jump detection results based on the *Medk* approach, for various values of k along with RV and BV, using the S&P 500 2-min sample in Subsection 5.1. We conclude that the measures with $k = 9$ are the most appropriate for these data. Subsequently, we use $Med9RV_{tM}$ to define standardized returns using equation (13) and consider jumps to be returns whose standardized absolute values exceed a certain threshold. We study the time of day at which such jumps are likely to occur and also consider the number of jumps per day as a time series. These are discussed in Subsection 5.2.

5.1 Local volatility and jump-robust estimators of daily volatility

Table 1 reports the statistics of volatility estimators $MedkRV_{tM}(k = 3, 5, 7, 9)$, BV and RV. They are calculated as a square root multiplied by a scale factor of 1,000 for each day (e.g. the statistics in column 1 are calculated on $1{,}000\sqrt{RV_t}$, $t = 1, \ldots 2{,}285$).The distributions of all $MedkRV_{tM}$ measures are similar but have lower means/medians than BV. The BV statistics are generally lower than RV except for the maximum value, indicating the problem of inflated BV in real data. *Med*7 and *Med*9 seem to cope with BV inflation better, compared to *Med*3 and *Med*5, given their lower maximums.

We also analyze a set of matched pair differences recorded in Table 2. Each test is based on 2,285 differences of two variables (one difference per day). Note that the mean differences are consistent with the differences of the means in the penultimate row of Table 1. With 2,285 values, the distributions of the mean are close to Normal with known variances. The Z statistic, the mean

[4] We thank Assad Jalali for his initial calculation of these factors (also see Jalali, 2014).

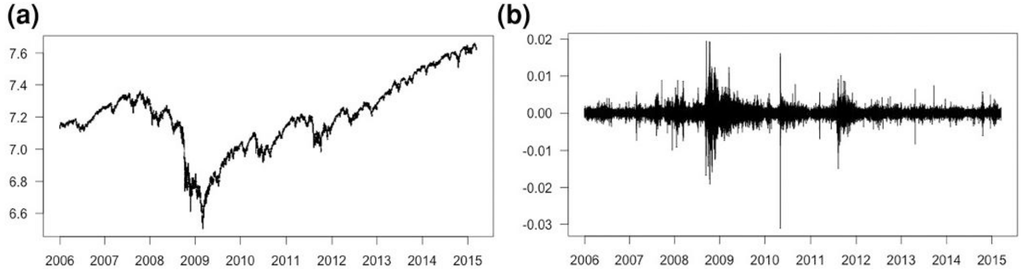

Figure 2. 2-min log-prices (a) and log-returns (b) of S&P 500 between 3 January 2006 and 13 March 2015.

difference divided by its standard error, provides a test of the hypothesis that the mean difference is zero. The mean differences are all clearly positive, including RV-BV, but smaller for $MedkRV_{tM}$ and the means of the *Medk* volatility decrease with increasing k.

The choice of an appropriate *MedkRV* varies according to the dataset. For example, Yang et al. (2018) use *Med7RV* for their 15-min level data while *Med9RV* is the best for our study. We illustrate this in Figure 3, with $BV = 3.93$, $RV = 3.62$, $Med3 = 3.31$, $Med5 = 3.10$, $Med7 = 2.59$, $Med9 = 2.34$. There is little difference among the local estimators except for a short section just after 10:00, where there are three large absolute returns within a run of five successive intervals: the effect of those returns is partially eliminated in $LV7$, and more so in $LV9$.

Panel (a) of Figure 4 shows the square root of the ratio of the realized volatility and a jump-robust estimate, $\sqrt{RV/Med9RV}$. The two largest values occur on 23 April 2013 and 18 September 2013. Panel (b) plots daily $\sqrt{Med9RV}$: there is considerable volatility in late 2008 and 2011, and *flash crash day*. Figure 5 illustrates the intraday volatility profile for our sample. For each 2-min interval, we take the median value of $\sqrt{LV9}$ of that interval (multiplied by 1,000) of each day and compute the average value over the 2,285 days. This reveals the typical U-shaped intraday volatility pattern (IVP) discussed in the literature like Andersen et al. (2012), Boudt et al. (2011) and others.[5]

5.2 Standardized returns and jumps

Table 3 summarizes the daily counts of three different-sized jumps ($n = 3, 4, 5$). The daily averages for J_3, J_4, and J_5 are 3.72 (too many), 1.10 and 0.41 (too few), respectively. We further find that well over half of the sample period have 0 or 1 J_4 jumps but as many as 5 jumps in a day is not uncommon and occasionally there are 6, 7 or even 8 jumps. Hence, we will work with J_4 jumps for the various tests.[6].

Figure 6 shows the counts of J_4 jumps within each 2-min interval of the day summed over the sample. Many jumps occurred in the first half hour of trading, then the rate drops. Between 11:00 and 14:00, the market appears quiet, followed by more activity toward the end of the day when traders start to balance their order books to protect their overnight positions. What is striking is the very large number of jumps that occur just after 10am: remember the little blip about 10:00 in the intraday volatility profile in Figure 5. We believe this relates to news announcements routinely issued at 10:00.[7] We also see some concentration of jumps between 14:00 and14:20, similar to results found by Boudt and Petitjean (2014) for the Dow Jones Industrial Average (DJIA) index, coincident with some market report announcements.

[5] The IVP refers to the seasonality pattern where the market is more volatile at the beginning and end of a trading day (especially at the beginning). We take $MedkRV_{t,M}$ to be constant during day t instead of adjusting for IVP because this popular treatment could introduce arbitrary information that is not carried by real data (see some examples in Figure 8a). Also note a slight blip just after 10:00 (discussed later).

[6] Note that we use a 'Data Science' approach to choose a critical test value (n) based on data rather than a traditional significance level approach that needs assumptions about a probability distribution that we do not know. After all, it is the number of events that is important here. We also implemented simulations to test the robustness of the J_4 jump detection method. These are not included in the paper to meet the publication requirements of the article length, but can be requested from the authors.

[7] e.g. University of Michigan Consumer Sentiment Survey, Institute for Supply Management purchasing indices, housing reports, US Bureau of Labor Statistics and US Census Bureau reports.

Table 1. Basic statistics for $MedkRV_{tM}(k=3, 5, 7, 9)$, BV and RV on S&P 500

Statistic	RV	BV	Med3	Med5	Med7	Med9
min	1.74	1.73	1.67	1.51	1.49	1.44
LQ	4.42	4.19	4.13	3.99	3.93	3.89
median	6.08	5.81	5.65	5.57	5.50	5.43
UQ	9.11	8.71	8.57	8.43	8.26	8.18
max	81.04	87.15	82.42	82.02	78.98	77.49
mean	7.85	7.55	7.42	7.28	7.19	7.13
sd	6.04	5.98	5.91	5.89	5.84	5.81

Table 2. Matched pair analysis for volatility measures

Statistic	RV - BV	BV - Med3	Med3 - Med5	Med5 - Med7	Med7 - Med9
mean	0.3033	0.1269	0.1426	0.0869	0.0599
SE of mean.	0.0095	0.0074	0.0073	0.0052	0.0041
Z statistic	31.9	17.1	19.6	16.8	14.7

Figure 7 shows autocorrelations and partial autocorrelations of the daily numbers of J_4 jumps for S&P 500. There are several significant (at 5%) positive correlations within the first ten lags (shown as falling outside horizontal dotted lines on the graphs), suggesting that the contagion effect of jumps also spills over several days.

6 Comparison of detection methods

We compare our *Med*9 with other common jump detection methods: *Med*9 vs. RV-BV method by Barndorff-Nielsen and Shephard (2004, 2006) in 6.1 and *Med*9 vs. *LM* test by Lee and Mykland (2008) and *ABD* test of Andersen et al. (2007a) in 6.2.

6.1 *Med9* vs. *RV-BV* methods

The original RV-BV method and other extensions (e.g. Andersen et al., 2010 recursive method) test the null hypothesis of no jumps on a day against the alternative of at least one jump on the day. However, these methods sometimes indicate no jumps when actual jumps took place (e.g. *flash crash day*). We compared the detection results on 12 jump days reported by Nanex in our sample (see Table 4). There are nine days when the RV-BV tests are not significant (NS), but we find at least one J_4 jump. Sometimes the BV is larger than RV, resulting a negative $1 - BV/RV$, hence no jumps.

Table 5 shows how the numbers of days with positive RV-BV tests is related to the number of J_4 jumps detected in a day by the *Med*9 method. For example, of the 897 days on which there are no J_4 jumps, there are 155 days (17% of 897) where the RV-BV test suggests at least one jump. As the number of J_4 jumps in a day, n, increases, RV-BV find fewer days with at least one jump. Even when $n >= 3$, there are still about 1/3 such days where the RV-BV test indicates no jumps on that day. Overall, the RV-BV method seems to suffer from inflation in BV when large jumps occur consecutively, but not so much with isolated medium-sized jumps. The median method (e.g. *Med*9) performs much better by mitigating the effect of large individual jumps.

6.2 *Med9* vs. *LM* & *ABD* tests

The LM volatility estimator $\hat{\sigma}^2(t_i) = \frac{1}{K-1}\sum_{j=i-K+1}^{i-1} |r_{j-1}|\,|r_j|$ is proportional to bi-power variation, hence subject to a similar inflation problem if there are adjacent large absolute returns.

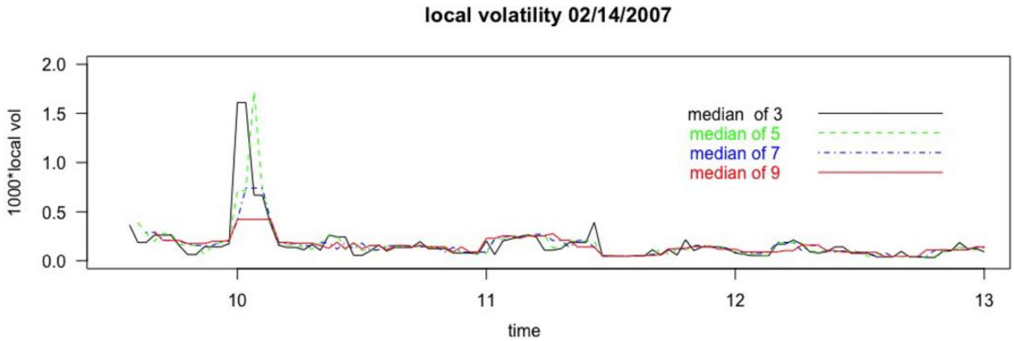

Figure 3. Local volatility ($\sqrt{LVk_{t,j}}$) of S&P 500 on the morning of 14 February 2007.

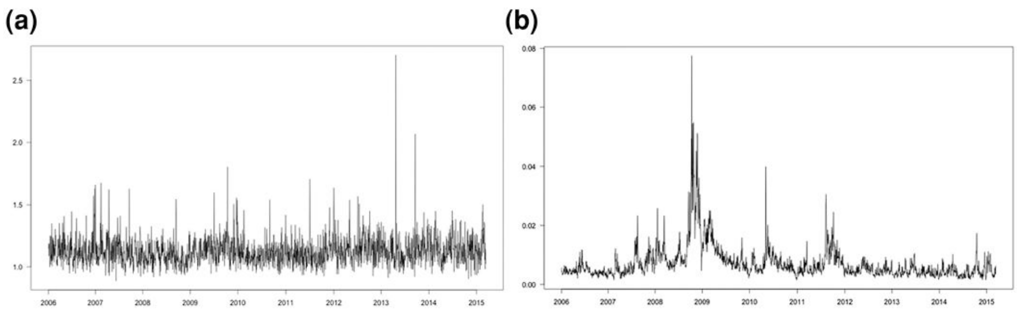

Figure 4. Daily realized volatility of S&P 500 over the sample period. Panel (a) shows $\sqrt{RV/Med9RV}$ and Panel (b) shows $\sqrt{Med9RV}$.

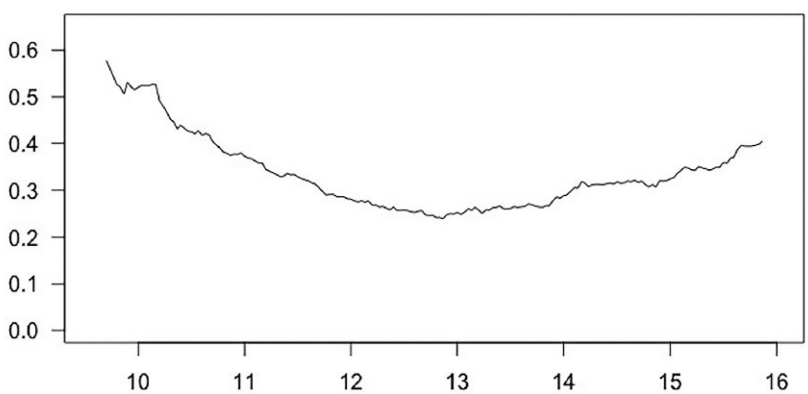

Figure 5. Intraday local volatility of S&P 500.

Table 3. The count of J_3, J_4, and J_5 jump days in S&P 500

Jumps/day	0	1	2	3	4	5	6	7	8	>8	Total	Avg.
J_3	115	271	390	386	366	285	200	121	75	76	8,494	3.72
J_4	897	733	358	172	87	29	7	1	1	0	2,515	1.10
J_5	1,599	493	146	37	8	2	0	0	0	0	938	0.41

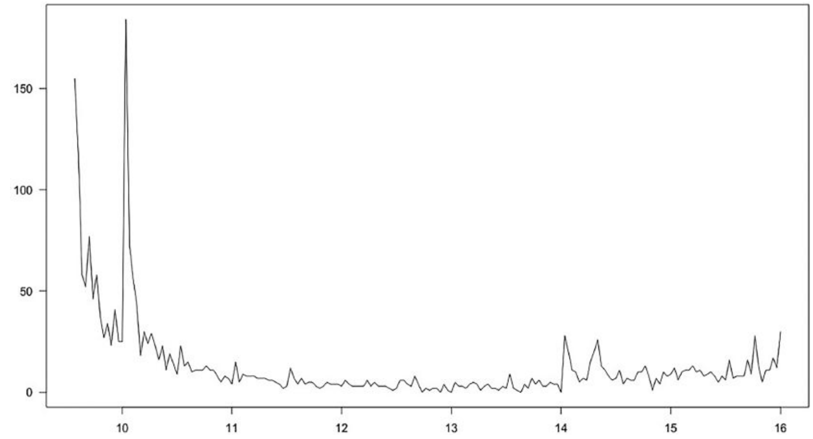

Figure 6. J_4 jumps in the S&P 500 index returns over the sample.

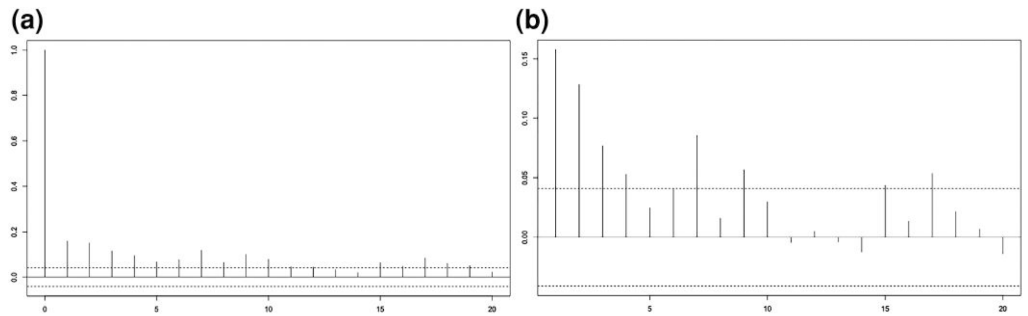

Figure 7. Panels (a) and (b) show, respectively, autocorrelations and partial autocorrelations of the daily numbers of J_4 jumps for S&P 500.

Table 4. Jump detection comparison for selected days

Date	J_4	$RV - BV$	$1 - BV/RV$
05/10/2006	6	NS	0.048
05/09/2007	2	NS	−0.245
05/15/2008	0	S	0.098
08/05/2008	2	NS	−0.147
11/24/2009	1	NS	−0.094
05/06/2010	4	NS	−0.088
03/31/2011	2	NS	0.083
06/14/2012	4	NS	0.034
10/31/2012	0	S	0.140
04/23/2013	3	NS	0.082
09/18/2013	3	S	0.532
04/21/2014	2	NS	0.005

Note. S or NS refers to a significant or not significant null hypothesis of no jumps under the RV-BV test at the 5% significance level.

Table 5. Two-way table of *RV-BV* and J_4 counts

n	0	1	2	3	4	≥ 5	*Total*
# days with n J_4 jumps	897	733	358	172	87	38	2,285
# days RV-BV test positive	155	279	188	110	56	25	813
% days RV-BV test positive	17	38	53	64	64	66	36

Note. It reports the number of days with n J_4 jumps and the number of such days when the RV-BV indicates at least one jump.

It treats the returns as one long series, not divided into different days, while the *Med9* uses a median of variation based on all returns on the same day. $\hat{\sigma}^2(t_i)$ is measured on returns prior to the return being tested but *Med9* is for end-of-the-day analysis. We propose two comparisons:1) *Med9* vs. *Pseudo LM* that uses the bi-power variance estimate over the whole day in which the return we wish to test is located, like the *ABD* test by Andersen et al. (2007a); and 2)The *LM* that calculates the local volatility prior to the return being tested, and compare it with *Med9* which is shifted to estimate the volatility over the same prior period (so this includes some returns on both previous days and the current day): call this modified *Med9* ($Med9_{mod}$).

6.2.1 *ABD (Pseudo LM)* vs. *Med9*

We begin by generating a series of $2{,}285 \times 194$ i.i.d $N(0, 1)$ variables to compare the variances estimated based on *Med9* and BV that is proportional to the variance quoted by *ABD* (see Table 6). Taking $M = K = 194$ as the number of intervals per day, we calculate for each day the variance per interval as $\hat{\sigma}^2_{Med9}(t) = Med9_{t,M}/M$ and

$$\hat{\sigma}^2_{ABD}(t) = \hat{\sigma}^2_{BV}(t) = \frac{\pi}{2(M-1)} \sum_{j=2}^{M} |r_{t,j-1}|\,|r_{t,j}|, \; t = 1, \; \dots, 2{,}285. \tag{14}$$

Both means are close to 1, consistent with $N(0, 1)$ data (white noise with no jumps). The estimator based on BV/ABD has a slightly more compact distribution.

For each test, we consider each of the $2{,}285 \times 194$ standardized returns, $r_{t,i}/\hat{\sigma}(t)$ to be divided by the standard deviation for that test on that day, and obtain the upper 0.01% critical values: 3.914 for the *ABD* test and 4.101 for *Med9*.[8] We apply both tests to the S&P 500 returns, using these critical 0.01% values. The number of days that n jumps occur is reported in Table 7.[9] Indeed *ABD* misses about 20% of the jumps that the *Med9* method finds (compare bottom two rows of Table 7).

When considering the statistic *ABD*, we use the median estimator $\hat{v}_{t,j} = \hat{\sigma}_{Med9}(t)$ for each day (i.e. for all $j = 1, \dots, M$) and the cut-off value used for *Med9*. Thus we replace $r_{t,j}$ in equation (14) by zero if $|r_{t,j}|/\hat{\sigma}_{Med9}(t) > 4.101$. This treatment could lead to a small volatility measure $\hat{\sigma}^2_{BV}(t)$ and yield too many '*ABD*' jumps. So, we propose to replace with a fraction $\alpha r_{t,j}$. Call α, $0 \leq \alpha \leq 1$, the shrink factor:[10] $\alpha = 0$ gives the original Threshold BV method and $\alpha = 1$ gives the unadjusted BV method. The count of jump days with various α are in Table 7. As expected, the ($\alpha = 0$) case results in too many jumps while ($\alpha = 1$) leads to less than one jump per day. When $\alpha = 0.3$, the results for the *ABD* and *Med9* methods are comparable.

8 We also have the critical value for the upper 0.1% level (3.328 for *ABD* and 3.386 for *Med9*. But 400+ false positives are found (about 1 in every 5 days on average, with some days having 2 or 3) - far too many for white noise. At the 0.01% level, there is roughly one false positive every 50 days.

9 The *Med9* test with critical value 4.101 reported 2,247 jumps ($J_{4.101}$)in total, slightly fewer than the total 2,515 J_4 jumps in Table 3. This is because bi-power variation causes the standardized return statistic using *ABD* to be reduced on some days: therefore possible jumps may be missed.

10 Corsi et al. (2010) proposes to replace large absolute returns by an expected value under a Normality assumption. This assumption seems to be far from how the real data behave.

Table 6. Distributions of variances from white noise simulation

	Min	Mean	Max	st.dev
$\hat{\sigma}^2_{ABD}$	0.658	0.998	1.472	0.116
$\hat{\sigma}^2_{Med9}$	0.551	0.999	1.619	0.145

6.2.2 *LM* vs. $\mathrm{Med9}_{\mathrm{mod}}$

We study the *LM* test using the 194 returns prior to the return that is to be tested to estimate the volatility. We compare these results with the $Med9_{mod}$ calculated over the same prior times.[11] We use the same critical values (i.e. 3.914 and 4.101 for the *LM* test and $Med9_{mod}$ test, respectively), and shrink factors for *LM*. The results of n-jump day counts for S&P 500 are given in Table 8.

Comparing the results with those of Table 7 suggests that looking at volatility on previous days tends to give rather wild results, presumably because the volatility changes from day to day: more jumps will be detected in a volatile period that follows a quiet period. As before the modified *Med9* method finds more jumps than the raw *LM* method; the threshold *LM* method is fairly stable with respect to the value of α but a shrinkage value of about 0.3 for the *LM* method is probably again about right.[12]

6.3 Piecewise constant intraday volatility patterns

Many studies have noted intraday volatility patterns (IVPs) as evidence of time-varying volatility by assuming the U-shape volatility profile and testing its validity and robustness. Standard approaches include (Andersen & Bolleslev, 1997) decomposition of intraday volatility, (Parkinson, 1980) high-low approach, open-high-low-close profiling by Garman and Klass (1980) (also see Chan & Lien, 2003), Andersen et al. (2001) and Andersen et al. (2007a)'s filtered J-statistics (a periodicity test), and Boudt et al. (2011)'s extension to robust filtered J-statistics etc.[13]

Boudt et al. (2011) tested the hypothesis that IVP introduces biases in jump detection (also see Gilder et al. (2014)) but concluded that the null is not significant. Bolleslev et al. (2008) suggest the sampling frequency (e.g. Dimitru & Urga, 2012) and stale prices (e.g. Corsi et al., 2010 and Schulz & Mosler, 2011) are the real causes for such biases. Andersen et al. (2019) and Andersen et al. (2023) further suggest that the real intraday volatility profiling would be far from IVP and believe the variation in the IVP is partly driven by the current level of volatility (also see Torben et al., 2019 and Vatter et al., 2015).

We propose a new simulation approach, *Piecewise Constant IVP* (PC-IVP), to demonstrate that IVP does not affect jump detection performance, hence extending the literature. The PC-IVP method can also be customized to fit different frequencies of intraday volatility, thus is more precise in capturing intraday volatility patterns in real data.

To demonstrate the method, we test daily volatility patterns at both 1- and 2-min S&P log-returns over the 2,285-day sample. Instead of considering a constant volatility averaged over a day, we divide each day into n sub-windows. Though the partition does not need to be evenly spaced, we use the equal sized windows for simplicity. The volatility within each sub-interval is assumed to be constant. We calculate the piecewise approximation of daily volatility over n sub-intervals (P_n-volatility). We have several significant findings: first, we demonstrate the U-shape exists; second, we find the IVP is the dominant pattern for the intraday time-varying volatility; and finally, other volatility patterns, often, noisy, are also identified within a day.

[11] i.e. The *Med9* variance estimator is modified to be averaged over 186 medians in the same window as the corresponding *LM* variance is calculated. The first return to be tested is the first return of day 2 with the variance estimated on the previous 194 returns. Thereafter the variance estimator rolls forward one interval at a time and is applied to the corresponding test return that moves forward in parallel.

[12] We also compare the *LM* with *Med9*. We find that the total jumps are 2,500, 2,484, 2,476, and 2,456 for $\alpha = 0$, 0.3, 0.4, 1, respectively

[13] Early theoretical works on Ergodicity of IVP include Birkhoff (1942) and Walters (1982).

Table 7. The count of n-jump days with shrink α in S&P 500 using *ABD* and *Med*9

	Shrink α	0	1	2	3	4	5	6	7	>7	Total Jump	Daily Avg.
	0	978	682	334	151	85	30	17	2	6	2,460	1.08
ABD	0.3	990	716	332	151	65	21	7	2	1	2,262	0.99
	0.4	994	737	328	144	54	24	3	0	1	2,187	0.96
	1	1,038	806	345	76	19	1				1805	0.79
*Med*9	–	995	717	325	149	69	25	3	2		2,247	0.98

Table 8. The count of n-jumps in S&P 500 using *LM* and $Med9_{mod}$

	Shrink α	0	1	2	3	4	5	6	7	>7	Total Jump	Daily Avg.
	0	1,017	638	310	169	69	46	20	8	67	2,505	1.10
LM	0.3	1,023	635	309	168	68	47	19	8	7	2,491	1.09
	0.4	1,028	631	309	169	66	48	18	8	7	2,481	1.09
	1	1,034	632	307	164	69	47	16	10	5	2,456	1.08
$Med9_{mod}$	–	1,013	584	288	171	94	62	33	15	24	2,881	1.26

We report the results of P_3 and P_6-volatility. The P_3-volatility patterns are shown in Figure 8 and it is clear that our $Med9P_3$-volatility is comparable to the ABD method in capturing the IVP. This implies the IVP does not affect jump detection whether the *Med9* or another detection method is applied. We further refine each P_3 partition into two equal sub-intervals, forming the P_6-volatility illustrated in Figure 9. It is also clear that the $Med9P_6$-volatility remains a U-shape pattern, suggesting that IVP does not affect jump detection.[14]

With the P_3-volatility, there could be up to eight pattern, for instance (high, low, high) denoted as $(+, -, +)$ etc.[15]. We compute the percentage of the occurrence of each pattern to demonstrate the P_3-volatility profile under both *ABD* and *Med9* methods at both frequencies. We identify four patterns (see Table 9) that are distinctively visible but the other four are not. We not only find the U-shape volatility represented by Pattern $(+, -, +)$, but also show the IVP is dominant. We further identify that pattern($(+, -, -)$) also has significant counts though less dominant than the volatility smile. Two other patterns also have effective counts. Moreover, our *Med9* method shows a slight advantage over *ABD* in capturing the IVP (71.33% vs. 69.50 & 72.82% vs. 70.97%). We also summarize the P_6-volatility pattern counts (5 distinctive patterns) in Table 10. It is more prominent that IVP is not the only pattern for intraday volatility but still somewhat dominant.

We further report the count of n-jumps using P_3 and P_6- volatility (*Med9* & *ABD*) at 2-min frequency in Table 11.[16] This shows that the daily average detection of jumps for $ABD(P_3)$ drops about 30% and a further 16% for $ABD(P_6)$. In contrast, the $Med9(P_3)$ detection drops less than 10% and the $Med9(P_6)$ performance is better. It is noticeable that $ABD(P_6)$ does not perform well for jumps higher than 5. This result demonstrates that *Med9* is more robust than ABD with respect to IVP partition.

[14] Similar patterns are found for other scenarios: 2-min *Med*9, 1- & 2-min *ABD*.

[15] The eight patterns include: (high, higher, higher)–$(+, +, +)$,(low, lower, lower)–$(-, -, -)$,(high, higher, low)–$(+, +, -)$, (low, lower, high)–$(-, -, +)$; (high, low, lower)–$(+, -, -)$,(high, low, high)–$(+, -, +)$, (low, high, low)–$(-, +, -)$ and (low, high, higher)–$(-, +, +)$.

[16] 1-min results are similar. $ABD(P_3)$ detections drop 41% and $Med9(P_3)$ decrease 27%. $ABD(P_6)$ drops another 26% while $Med9(P_6)$ reduces a further 5%.

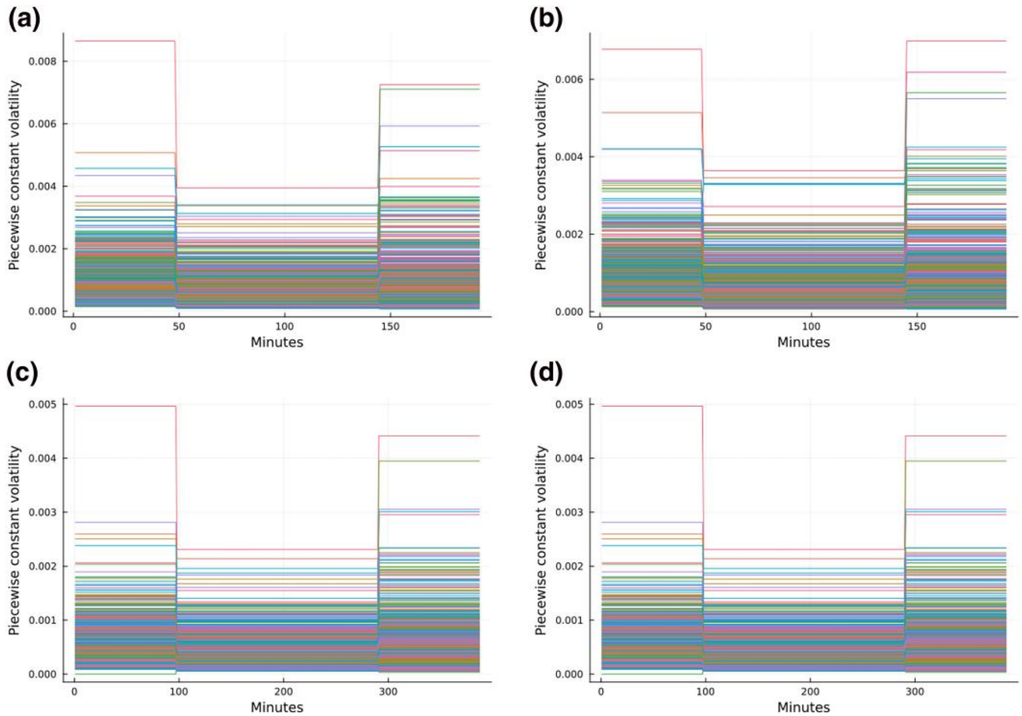

Figure 8. P_3-volatility patterns for *Med*9 and *ABD* methods. Panels (a) (b)are the 2-min P_3-volatility volatilities while Panels (c) (d) are the 1-min patterns. (a) 2-min *ABD*. (b) 2-min *Med*9. (c) 1-min *ABD*. (d) 1-min *Med*9.

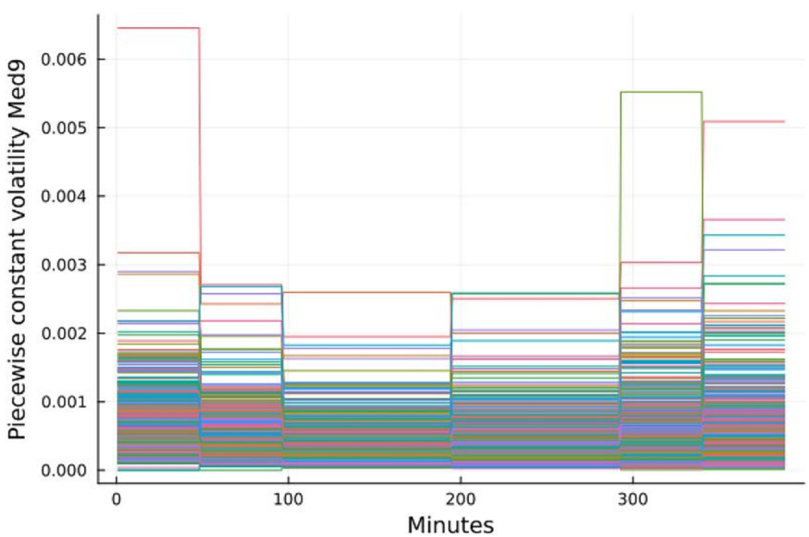

Figure 9. 1-min P_6-volatility pattern using *Med*9.

6.4 Simulations

We compare *Med9* and *ABD* with a shrink factor α using simulations. We have performed various simulations but specifically report the cases representing jump clustering (contagion) effects, which is the core reason for proposing the *Med9* method (see Subsections 6.4.1 and 6.4.2.

Table 9. Intraday volatility pattern counts with P_3-volatility

	2-min		1-min	
Pattern	ABD %	Med9 %	ABD %	Med9 %
(+, −, +)	69.50	71.33	70.97	72.82
(+, −, −)	23.81	22.89	23.23	22.09
(−, +, +)	3.81	2.97	3.51	3.82
(−, +, −)	2.89	2.80	2.28	1.27

Table 10. Intraday volatility pattern counts with P_6-volatility

	2-min		1-min	
Pattern	ABD %	Med9 %	ABD %	Med9 %
(+ − − − ++)	18.29	14.75	20.16	19.54
(− − − − −+)	9.93	8.79	11.11	10.67
(− + − − ++)	3.02	4.29	1.84	2.90
(− − − + −−)	3.59	2.76	3.38	2.5
(− + + + ++)	0.00	0.00	0.00	0.00

Table 11. The count of *n*-jumps with P_3 & P_6-volatility at 2-min

	0	1	2	3	4	5	6	7	>7	Total Jump	Daily Average
ABD(P1)	1,038	806	345	76	19	1	0	0	0	1,805	0.79
ABD(P_3)	1,358	648	224	44	9	2	0	0	0	1,274	0.56
ABD(P_6)	1,460	612	171	37	5	0	0	0	0	1,085	0.47
Med9(P1)	995	717	325	149	694	25	3	2	0	2,247	0.98
Med9(P_3)	1,092	676	302	130	55	18	8	4	0	2,056	0.90
Med9(P_6)	979	670	374	146	66	32	11	4	3	2,399	1.049

To achieve this, we simulate a Hawkes-jump-diffusion process by superposing a one-dimensional Hawkes process (see Yang et al., 2018) on a Gaussian white noise series and calibrate the parameters to the S&P500. The rationale is to take advantage of the Hawkes processes' ability to generate consecutive jumps (contagion) through its self-exciting intensity. We stress test for both methods through raising the intensity of jump occurrences controlled by λ_t. We see the *ABD* method starts to fail to detect jumps but *Med9* remains robust when λ_t increases.

We start by generating a series of $2{,}285 \times 194$ from an i.i.d. N(0, 1) process called Gaussian white noise. The summary statistics are as follows

	Min	Mean	Max	st.dev
$\hat{\sigma}^2_{ABD}$	0.6824	0.9993	1.4136	0.1163
$\hat{\sigma}^2_{Med9}$	0.5645	0.9995	1.5832	0.1449

The Hawkes-jump process, generated using package *HawkesProcesses.jl* in Julia, is then added to the Gaussian white noise, which symbolically reads $N(0, 1) + H$ where H has the intensity of jump occurrence λ_t

$$\lambda_t = \lambda_0 + \sum_{t>t_i} \gamma e^{-\beta(t-t_i)}$$

The jump size can be generated from an independent log-normal distribution. A jump size from a Lévy distribution was also examined. Then, the matrix of $2{,}285 \times 194$ (representing normalized returns) is constructed wherein the time of jump occurrences to reconstruct the intensity λ_t calibrated to the S&P500 sample using Bayesian statistics.

6.4.1 Jump clustering simulation at 2-min level

This simulation, with intensity parameters $(\lambda_0, \gamma, \beta) = (0.15, 2.2, 5)$, shows the day count of n-jumps in a simulated series that has a high percentage of self-excitations indicating consecutive jumps (contagion). As shown in Table 12, the *ABD* method fails to detect sufficient total jumps regardless of the shrink parameters while *Med9* performs consistently in identifying jumps. It is evident that *Med9* is a superior method for detecting jumps when jump clustering occurs because the *ABD* method is not particularly effective when there are ¿2 jumps in a day while *Med9* detects well even when $n > 7$. If we look back at the singular jump scenario, *ABD* suggests that jumps occur over a maximum of 2% of the sample period (47 out of 2,285 days), which is far from the reality. While the *Med9* reports approximately a much more realistic 15% jump days.

6.4.2 Jump clustering simulation at 1-min level

To show the robustness, we increase the simulated sample frequency to 1-minute, which increases the data matrix size to $2,285 \times 388$. With the same intensity parameters $(\lambda_0, \gamma, \beta) = (0.15, 2.2, 5))$, Table 13 reports the new results of the day count of n-jumps. First, we observe that doubling the data frequency generally worsens the effectiveness of the *ABD*

Table 12. The count of jump clustering at 2-min frequency

	Shrink α	0	1	2	3	4	5	6	7	>7	Total Jump	Daily Average
	0	2,224	47	7	3	3	0	1	0	0	88	0.039
ABD	0.3	2,240	38	4	1	2	0	0	0	0	57	0.025
	0.4	2,242	38	3	0	2	0	0	0	0	52	0.023
	1	2,256	28	1	0	0	0	0	0	0	30	0.013
Med9		1,407	351	200	120	70	51	27	20	39	2,306	1.009

Table 13. The count of jump clustering at 1-min frequency

	Shrink α	0	1	2	3	4	5	6	7	>7	Total Jump	Daily Average
	0	2,241	38	5	1	0	0	0	0	0	51	0.022
ABD	0.3	2,246	34	5	0	0	0	0	0	0	44	0.019
	0.4	2,247	33	5	0	0	0	0	0	0	43	0.0188
	1	2,252	32	1	0	0	0	0	0	0	34	0.0148
Med9		1,268	425	219	126	97	47	41	22	40	2,663	1.65

detection but *Med9* continues to show strong detection performance (see total jumps). *ABD* appears to perform even more poorly in recognizing clustered jumps for different shrink factors; but *Med9* can detect up to over seven consecutive jumps. Finally, *ABD* single jump detection declines, indicating a maximum 1.7% of the 2,285 days having a jump while that of the *Med9* improves to 18.6%.

7 Conclusion

Many existing methods of jump detection in financial series are based on integration over a whole day while others concentrate on looking for large individual returns. However, both types often fail to detect jumps if there are large returns in consecutive intervals. We propose a new detection method in which individual returns are assessed against a jump-robust measure of daily volatility, *Med9RV*, using a 9-interval rolling median. Thus a bigger return is required on volatile days than on quiet days to be called a jump. The volatility measure ensures that the test denominator is less likely be inflated by one or several large adjacent returns, which could cause a jump to be missed/masked like several traditional methods of jump detections.

We test the *Med9* method on a 2-min S&P 500 returns series over a 2,285-day sample period. First, the *Med9* method find jumps more effectively while *RV-BV* type methods often fail to identify jumps on days when most people would agree that jumps have occurred. Second, comparisons with the local volatility approaches of Lee and Mykland (2008) and Andersen et al. (2007a) also suggest that our *Med9* approach is consistently better at detecting jumps, especially identifying large jumps, because it does not suffer from inflated bi-power variation or insufficient time to allow for jump contagion to naturally unfold.

We have conducted simulations using a Hawkes-jump-diffusion process, particularly constructing the case of jump clustering, to demonstrate the robustness of the *Med9* vs. *ABD* methods. We further enhance the evidence in supporting the *Med9* method through increasing the frequency of the simulated series from 2 to 1 min. We find the *Med9* method consistently performs better. We find the *ABD* method not only detects an unrealistically low number of single-jump days, but also becomes ineffective in the presence of multiple jumps within a day ($n > 2$). In summary, we conclude that monitoring the local volatility is important for detecting jumps in price and return series. The proposed *Medn* approach overcomes the challenges of measuring intraday volatility when jump contagion and clustering exist and more accurately locates jumps.

Conflicts of interest: To the knowledge of all authors, we do not have any conflict of of interest for all categories identified by the journals' guidelines (referring to the 'Conflict of Interest Disclosure' form by JRSSC).

Data availability

The data are obtained from Thomson Reuters ™ Tick History under the paid data subscription license of Swansea university. Data can be shared on request to the authors within the remit of such subscription license. Alternatively, researchers could directly request a sample similar to this dataset from Thomson Reuters ™ Tick History or other data sources for educational purpose.

References

Andersen T. G., & Bollerslev T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, *4*(2–3), 115–158. https://doi.org/10.1016/S0927-5398(97)00004-2

Andersen T. G., & Bollerslev T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*(4), 885–905. https://doi.org/10.2307/2527343

Andersen T. G., Bollerslev T., & Das A. (2001). Variance-ratio statistics and high-frequency data: Testing for changes in intraday volatility patterns. *Journal of Finance*, *56*, 305–327. https://doi.org/10.1111/0022-1082.00326

Andersen T. G., Bollerslev T., & Debrov D. (2007a). 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. https://doi.org/10.1016/j.jeconom.2006.05.018

Andersen T. G., Bolleslev T., & Diebold F. X. (2007b). Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*, *89*(4), 701–720. https://doi.org/10.1162/rest.89.4.701

Andersen T. G., Bolleslev T., Diebold F. X., & Labys P. (2000). Exchange rate returns standardized by realized volatility are (nearly) Gaussian. *Multinational Finance Journal*, *4*(3/4), 159–179. https://doi.org/10.17578/4

Andersen T. G., Bolleslev T., Diebold F. X., & Labys P. (2001). The distribution of exchange volatility. *Journal of the American Statistical Association*, *96*(453), 42–55. https://doi.org/10.1198/016214501750332965

Andersen T. G., Bolleslev T., Diebold F. X., & Labys P. (2003). Modeling and forecasting realized volatility. *Econometrica: Journal of the Econometric Society*, *71*(2), 579–625. https://doi.org/10.1111/ecta.2003.71.issue-2

Andersen T. G., Bolleslev T., Frederiksen P., & Nielsen M. (2010). Continuous time models, realized volatilities and testable distributional implications for daily stock returns. *Journal of Applied Econometrics*, *25*(2), 233–261. https://doi.org/10.1002/jae.v25:2

Andersen T. G., Dobrev D., & Schaumburg E. (2012). Jump robust volatility estimation using nearest neighbor truncation. *Journal of Econometrics*, *169*(1), 75–93. https://doi.org/10.1016/j.jeconom.2012.01.011

Andersen T. G., Su T., Todorov V., & Zhang Z. (2023). Intraday periodic volatility curves. *Journal of the American Statistical Association*, *119*, 1–11. https://doi.org/10.1080/01621459.2023.2177546

Andersen T. G., Thyrsgaard M., & Todorov V. (2019). Time-varying periodicity in intraday volatility. *Journal of the American Statistical Association*, *114*(528), 1695–1707. https://doi.org/10.1080/01621459.2018.1512864

Bachelier L. (2006). In M. H. Davis, & A. Etheridge (Eds.), *Louis Bachelier's theory of speculation*. Princeton University Press.

Bajgrowicz P., Scaillet O., & Treccani A. (2016). Jumps in high-frequency data: Spurious detections, dynamics and news. *Management Science*, *62*(8), 2198–2217. https://doi.org/10.1287/mnsc.2015.2234

Barndorff-Nielsen O. E., & Shephard N. (2002). Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B*, *64*(2), 253–280. https://doi.org/10.1111/1467-9868.00336

Barndorff-Nielsen O. E., & Shephard N. (2004). Power and bipower variation with stochastic volatility and jumps. *Journal of Financial Econometrics*, *2*(1), 1–37. https://doi.org/10.1093/jjfinec/nbh001

Barndorff-Nielsen O. E., & Shephard N. (2006). Econometrics of testing for jumps in financial economics using bipower variation. *Journal of Financial Econometrics*, *4*(1), 1–30. https://doi.org/10.1093/jjfinec/nbi022

Birkhoff G. D (1942). What is the ergodic theorem? *The American Mathematical Monthly*, *49*(4), 222–226. https://doi.org/10.1080/00029890.1942.11991212

Black F., & Scholes M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*(3), 637–654. https://doi.org/10.1086/260062

Bolleslev T., Law T. H., & Tauchen G. (2008). Risk, jumps and diversification. *Journal of Econometrics*, *144*(1), 234–256. https://doi.org/10.1016/j.jeconom.2008.01.006

Boudt K., Croux C., & Laurent S. (2011). Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Financial Economics*, *18*(2), 353–367. https://doi.org/10.1016/j.jempfin.2010.11.005

Boudt K., & Petitjean M. (2014). intra-day liquidity dynamics and news releases around price jumps: Evidence from the DJIA stocks. *Journal of Financial Markets*, *17*(2), 121–149. https://doi.org/10.1016/j.finmar.2013.05.004

Brogaard J., Carrion A., Moyaert T., Riordan A., Shkilko A., & Sokolov K. (2018). High-frequency trading and extreme price movements. *Journal of Financial Economics*, *128*(2), 253–265. https://doi.org/10.1016/j.jfineco.2018.02.002

Chan L., & Lien D. (2003). Using high, low, open, and closing prices to estimate the effects of cash settlement on futures prices. *International Review of Financial Analysis*, *12*(1), 35–47. https://doi.org/10.1016/S1057-5219(02)00125-4

Christoffersen P., Jacobs P. K., Ornthanalai C., & Wang Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics*, *90*(3), 272–297. https://doi.org/10.1016/j.jfineco.2007.12.003

Cont R., & Tankov P. (2016). *Financial modelling with jump processes*. Chapman & Hall/CRC.

Corsi F., Pirino D., & Reno R. (2010). Threshold bipower variation and the impact of jumps on volatility forecasting. *Journal of Econometrics*, *159*(2), 276–288. https://doi.org/10.1016/j.jeconom.2010.07.008

Dimitru A. M., & Urga G. (2012). Identifying jumps in financial assets: A comparison between nonparametric jump tests. *Journal of Business and Economic Statistics*, *30*(2), 242–255. https://doi.org/10.1080/07350015.2012.663250

Ferriani F., & Zoi P. (2020). The dynamics of price jumps in the stock market: An empirical study on Europe and US. *The European Journal of Finance*, *28*, 718–742. https://doi.org/10.1080/1351847X.2020.1740288.

Garman M., & Klass M. (1980). On the estimation of security price volatilities from historical data. *Journal of Business*, *53*(1), 67–78. https://doi.org/10.1086/jb.1980.53.issue-1

Gilder D., Shackelton M., & Taylor S. (2014). Cojumps in stock prices: Empirical evidence. *Journal of Banking and Finance*, *40*(2), 443–459. https://doi.org/10.1016/j.jbankfin.2013.04.025

Huang X., & Tauchen G. (2005). The relative contribution of jumps to total price variation. *Journal of Financial Econometrics*, *3*(4), 456–499. https://doi.org/10.1093/jjfinec/nbi025

Jalali A. (2014). On expected values of maxima of samples from Gamma distributions. Personal communication.

Kolokolov A., & Renò R. (2017). Efficient multipowers. *Journal of Financial Econometrics*, *16*(4), 629–659. https://doi.org/10.1093/jjfinec/nbx018

Lee S. S., & Mykland P. (2008). Jumps in financial markets: A new nonparametric test and jump-dynamics. *The Review of Financial Studies*, *21*(6), 2535–2563. https://doi.org/10.1093/rfs/hhm056

Merton R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics*, *3*(1-2), 125–144. https://doi.org/10.1016/0304-405X(76)90022-2

Parkinson M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business*, *53*, 61–65. http://dx.doi.org/10.1086/296071

Podolkskij M., & Ziggel D. (2010). New tests for jumps in semimartingale models. *Statistical Inference for Stochastic Processes*, *13*(1), 15–41. https://doi.org/10.1007/s11203-009-9037-8

Prodromou T., & Westerholm P. J. (2022). Are high frequency traders responsible for extreme price movements? *Economic Analysis and Policy*, *73*(2), 94–111. https://doi.org/10.1016/j.eap.2021.11.001

Samuelson P. A. (1965). Rational theory of warrant pricing. *Industrial Management Review*, *6*, 13–22.

Schulz F. C, & Mosler K. (2011). The effect of infrequent trading on detecting price jumps. *AStA Advances in Statistical Analysis*, *95*(1), 27–58. https://doi.org/10.1007/s10182-010-0137-y

Sornette D., & von der Becke S. (2011). Crashes and high frequency trading: An evaluation of risks posed by high-speed algorithmic trading. Foresight Driver Review – DR 7 for The Future of Computer Trading in Financial Markets, Government Office for Science, UK. Date accessed June 1, 2022.

Torben G. A., Martin T., & Viktor T. (2019). Iime-varying periodicity in intraday volatility. *Journal of the American Statistical Association*, *114*(528), 1695–1707. https://doi.org/10.1080/01621459.2018.1512864

Tsai P.-C., & Shackleton M. B. (2016). *Chapter 6, Handbook of high-frequency trading and modeling in finance*. Wiley.

Vatter T., Wu H. T., Chavez-Demoulin V., & Yu B. (2015). Non-parametric estimation of intraday spot volatility: Disentangling instantaneous trend and seasonality. *Econometrics*, *3*(4), 864–887. https://doi.org/10.3390/econometrics3040864

Walters P. (1982). *An introduction to ergodic theory*. Springer-Verlag.

Yang S. Y., Liu A., Chen J., & Hawkes A. G. (2018). Application of a multivariate Hawkes process to joint modelling of sentiment and market return events. *Quantitative Finance*, *18*(2), 295–310. https://doi.org/10.1080/14697688.2017.1403156