Financial Econometrics and Big Data: A Survey of Volatility Estimators and Tests for the Presence of Jumps and Co-Jumps*

Arpita Mukherjee, Weijia Peng, Norman R. Swanson and Xiye Yang

Rutgers University†

October 2018

Abstract

In recent years, the field of financial econometrics has seen tremendous gains in the amount of data available for use in modeling and prediction. Much of this data is very high frequency, and even 'tick-based', and hence falls into the category of what might be termed "big data". The availability of such data, particularly that available at high frequency on an intra-day basis, has spurred numerous theoretical advances in the areas of volatility/risk estimation and modeling. In this paper, we discuss key such advances, beginning with a survey of numerous nonparametric estimators of integrated volatility. Thereafter, we discuss testing for jumps using said estimators. Finally, we discuss recent advances in testing for co-jumps. Such co-jumps are important for a number of reasons. For example, the presence of co-jumps, in contexts where data has been partitioned into continuous and discontinuous (jump) components, is indicative of (near) instantaneous transmission of financial shocks across different sectors and companies in the markets; and hence represents a type of systemic risk. Additionally, the presence of co-jumps across sectors, say, suggests that if jumps can be predicted in one sector, then such predictions may have useful information for modeling variables such as returns and volatility in another sector. As an illustration of the methods discussed in this paper, we carry out an empirical analysis of DOW and NASDAQ stock price returns.

Keywords: Financial econometrics, Integrated volatility, Nonparametric estimator, Continuous time model, Jumps, Co-jumps, Big data, High-frequency data.

JEL Classification: C22, C52, C53, C58.

---

*This paper has been prepared for inclusion in the Handbook of Statistics. The authors are grateful to the editors, and in particular to H.D. Vinod, who has provided invaluable assistance during the preparation of this chapter. Thanks are also owed to Yacine Aït-Sahalia, Valentina Corradi, Frank Diebold, Rob Engle, Jianqing Fan, Eric Ghysels, Yuan Liao, Bruce Mizrach, and Markus Pelger for discussions that have aided in much of the authors' own research that is discussed herein.

†Department of Economics, Rutgers University, 75 Hamilton Street, New Brunswick, NJ - 08901

1 Introduction

The importance of integrated volatility, jumps and co-jumps in the financial econometrics literature and in terms of successful risk management by investors is quite obvious now, given the amount of research that has gone into this field. Measures of integrated volatility are crucial given the advent of numerous volatility based derivative products traded in financial markets while tests for jumps are essential in modeling and predicting volatility and returns. Tests of co-jumps on the other hand are meaningful indicators of transmission of financial shocks across different sectors, companies and markets. The rationale behind this chapter is to discuss some of recent advances in jump and co-jump testing methodology and measurement of integrated volatility, and the properties thereof, in a way which would help both researchers and practitioners in application of such econometric methods in finance. We begin by surveying the most widely used integrated volatility measures, jump and co-jump tests, followed by an empirical analysis using high frequency intra-day stock prices of DOW 30 companies and ETFs.

Daily integrated volatility is unobservable. Econometricians have developed numerous measures which estimate price fluctuations in a variety of ways. One of the earliest measures is the *Realized Volatility* in Andersen et al. (2001). However this measure does not separate jump variation from variation due to continuous components. Barndorff-Nielsen and Shephard (2004) use the product of adjacent intra-day returns to develop jump robust measures *Bipower* and *Tripower Variations*. One of the more recent techniques of separating out the jump component is the truncation methodology which essentially eliminates returns which are above a given threshold as in Corsi et al. (2010) & Aït-Sahalia et al. (2009). One important caveat of high-frequency data is the existence of market microstructure noise which creates a bias in the estimation procedure. Zhang et al. (2005), Zhang et al. (2006) and Kalnina and Linton (2008) solved this problem with noise robust volatility estimators.

In Duong and Swanson (2011), the authors find that 22.8% of the days during the 1993-2000 period had jumps while 9.4% of the days during the 2001-2008 period had jumps. The existence of jumps in financial markets is obvious, which has led many researches to develop techniques which can test for jumps. Jump diffusion is pivotal in analyzing asset movement in financial econometrics and developing jump tests to identify jumps has been the focus for many theoretical econometricians in past few years. Using the ratio of *Bipower Variation* and estimated quadratic variation, Barndorff-Nielsen and Shephard (2006) construct a non parametric test for the existence of jumps. Lee and Mykland (2007) on the other hand propose tests to detect the exact timing of jumps at the intra-day level while Jiang and Oomen (2008) provide a "swap variance" approach to detect the presence of jumps. Instead of the more widely use "fixed time span" tests, Corradi et al. (2014) and Corradi et al. (2018) develop "long time span" jump test, building on earlier work by Aït-Sahalia (2002).

Co-jump tests which are instrumental in identifying systemic risk across multiple sectors and markets are relatively new in the literature. Co-jumps reflect market correlation and have important implication for portfolio management and risk hedging. There are tests which utilize univariate jump tests to identify co-jumps among multivariate processes (Gilder et al. (2014)), while co-jump tests can also be directly applied to multiple price processes (see, e.g., Jacod and Todorov (2009), Bandi and Reno (2016), Bibinger and Winkelmann (2015) and Caporin et al. (2017)). Gnabo et al. (2014) propose a co-jump test based on bootstrapping methods, Bandi and Reno (2016) develop a nonparametric infinitesimal moments method to detect co-jumps between asset returns and volatilities and Caporin et al. (2017) build a co-jump test based on the comparison between smoothed realized variance and smoothed random realized variation.

As an illustration of the aforementioned testing methodologies and estimation procedures, an empirical analysis is carried out using high frequency intra-day stock prices of six DOW 30 companies and ETFs which include The Boeing Company (BA), Exxon Mobile Corporation (XOM), Johnson & Johnson (JNJ), JPMorgan Chase & Co. (JPM), Microsoft Corporation (MSFT) and Walmart Inc. (WMT) and two SPDR sector ETFs XLE & XLK. We use three jump tests; ASJ test (Aït-Sahalia et al. (2009)), BNS test Barndorff-Nielsen and Shephard (2006) and LM test (Lee and Mykland (2007)). In terms of co-jump tests we use, JT test (Jacod and Todorov (2009)), BLT test (Bollerslev et al. (2008)) and GST coexceedance rule (Gilder et al. (2014)). For estimation of integrated volatility we make use of *Realized Volatility* (Andersen et al. (2001)), *Bipower Variation* and *Tripower Variation* (Barndorff-Nielsen and Shephard (2004)), *Truncated Realized Volatility* (Aït-Sahalia et al. (2009)), *MedRV* and *MinRV* (Andersen et al. (2012)). In our findings, we report the volatility movement of the different stocks and ETFs, percentage of days identified as having jumps and co-jumps, kernel density plots of the different jump and co-jump test statistics as well the proportion of jump variation to the total variation in the asset prices.

The important empirical findings can be summarized as follows. Over the entire sample period JPMorgan has the highest and Johnson & Johnson has the lowest mean estimated integrated volatility. Amongst all the volatility measures, *Bipower Variation* reports the lowest mean volatility estimate while *Realized Volatility* reports the highest mean volatility estimate for any given stock or ETF. This can be explained by the fact that in the presence of frequent jumps, *Realized Volatility* overestimates integrated volatility. All individual stocks achieve their highest volatility in the fourth quarter of 2008 during the financial crisis. XLK sector ETF has the largest percentage of jump days (38%) and ratio of jump to total variation (45%) among all other ETFs and individual stocks. BNS jump test detected more jumps and reported a lager percentage of jump days when compared with the other two jump tests. When the sampling frequency is reduced from 1-minute to 5-minute, the ASJ jump test reports lesser number of jumps as well as smaller proportion of jump to total variation in the sample data. We detect co-jumps between Exxon & JPMorgan, Exxon & Microsoft, Exxon & XLE, JPMorgan & Microsoft, Microsoft & XLK and XLE & XLK through JT co-jump test and the GST co-exceedance rule. The results show that the percentage of co-jump days range from 0.4%-2.5% for JT co-jump test and from 2.8%-9.5% for the GST co-exceedance rule. The higher percentage of co-jump days in case of the co-exceedance rule, which uses the results at the intersection of BNS and LM jump tests, could be because the test has a large false rejection rate. We use BLT co-jump test to detect co-jumps among six stocks including Boeing, Exxon, Johnson&Johnson, JPMorgan, Microsoft and Walmart. The percentage of co-jumps days is 0.2% during financial crisis period and 0.1% after financial crisis period.

The rest of the paper is organized as follows. Section 2 gives the theoretical background and setup. Sections 3, 4 and 5 give detailed descriptions of the different integrated volatility measures, jump tests and co-jump tests respectively. Section 6 discusses the empirical methodology and reports the findings. Finally Section 7 concludes. Accompanying R code for the measures and tests is provided in the appendix.

2 Setup

We represent the log-price of a financial asset at continuous time t, as Y_t. It is assumed that the log-price is a Brownian semimartingale process with jumps and it can be denoted as[1]:

$$Y_t = Y_0 + \int_0^t \mu_s ds + \int_0^t \sigma_s dW_s + J_t \tag{1}$$

In (1) μ_s the drift term is a predictable process, σ_s the diffusion term is a cádlág process, W_s is a standard Brownian motion and J_t is a pure jump process. J_t can be defined as the sum of all discontinuous log price movements up to time t,

$$J_t = \sum_{s \leq t} \Delta Y_s \tag{2}$$

When this jump component is a finite activity jump process, i.e. a compound poisson process (CPP), then

$$J_t = \sum_{j=1}^{N_t} \xi_j \tag{3}$$

where N_t is a poisson process with intensity λ, the jumps occur at the corresponding times given as $(\tau_j)_{j=1,..,N_t}$ and ξ_j refers to $i.i.d$ random variables measuring the size of jumps at time τ_j. The finite activity jump assumption has been widely used in financial econometrics literature. Log-price Y_t can be decomposed into a continuous Brownian component Y_t^c and a discontinuous component Y_t^d (due to jumps). The "true variance" of process Y_t can be given as,

$$QV_t = [Y, Y]_t = [Y, Y]_t^c + [Y, Y]_t^d \tag{4}$$

where QV stands for quadratic variation. The variation due to the continuous component is

$$[Y, Y]_t^c = \int_0^t \sigma_s^2 ds, \tag{5}$$

and the variation due to the discontinuous jump component is

$$[Y, Y]_t^d = \sum_{j=1}^{N_t} \xi_j^2 \tag{6}$$

[1]We follow the setup and notation as in Corradi et al. (2011) and Mukherjee and Swanson (2018)

Integrated volatility which is the continuous part of QV is denoted as

$$IV_t = \int_{t-1}^{t} \sigma_s^2 ds, \quad t = 1, ..., T \tag{7}$$

where IV is the (daily) integrated volatility at day t. Since IV is unobservable, different realized measures of integrated volatility are used as its substitute. The presence of market frictions in high frequency financial data has been documented in recent literature. To take care of this, the observed log price process X can then be given as

$$X = Y + \epsilon \tag{8}$$

where Y is the latent log price and ϵ captures market microstructure noise. We consider M equi-spaced intradaily observations for each of T days for process X which leads to a total of MT observations, i.e.

$$X_{t+j/M} = Y_{t+j/M} + \epsilon_{t+j/M}, \quad t = 0, .., T \;\; \& \;\; j = 1, .., M \tag{9}$$

where ϵ follows a zero mean independent process. The intradaily return or increment of process X follows,

$$\Delta_j X = X_{t+(j+1)/M} - X_{t+j/M} \tag{10}$$

The noise containing realized measure, RM of the integrated volatility is computed using process X given in (9) and can be expressed as the sum of IV and measurement error N, i.e.

$$RM_{t,M} = IV_t + N_{t,M} \tag{11}$$

RM can be used to estimate IV if k^{th} moment of the measurement error decays to zero at a fast enough rate or there exists a sequence b_M with $b_M \to \infty$ such that $E(|N_{t,M}|^k) = O(b_M^{-K/2})$, for some k $\geq$2.

3 Realized Measures of Integrated Volatility

Volatility measures variation in the asset prices and thus can be regarded as an indicator of risk. Accurate volatility estimation is very important in both asset allocation and risk management. Since volatility is inherently unobservable, the first two types of parametric models developed to estimate the latent volatility were continuous time (e.g. stochastic volatility) and discrete time models (e.g. ARCH-GARCH models). However, these parametric models have been proven to be misspecified in capturing volatilities implied by option pricing and other financial return variables. With the availability of high frequency data, a series of nonparametric models have been proposed to examine integrated volatility at intra-day level. Andersen et al. (2001) first introduce a nonparametric volatility measure, termed *Realized Volatility* by summing over intra-day squared returns. The authors showed that *Realized Volatility* is an error free estimator of integrated volatility in the absence of noise and jumps. When the sampling frequency of the data is relatively high, microstructure noise creates a bias in the volatility estimation procedure. Zhang et al. (2005), Zhang et al. (2006) and Kalnina and Linton (2008) solve this problem with microstructure noise robust estimators based on sub-sampling with multiple time scales. Barndorff-Nielsen et al. (2008) and Barndorff-Nielsen et al. (2011) on the other hand, use kernel based estimators to account for the microstructure noise in finely sampled data. When estimating integrated volatility in the presence of jumps within the underlying price process, jump components should be separated from the quadratic variation. Barndorff-Nielsen et al. (2003), Barndorff-Nielsen and Shephard (2004) provide asymptotically unbiased integrated volatility estimators, the bipower and tripower variations, which are robust to the presence of jumps. Aït-Sahalia et al. (2009) propose a threshold method to identify and truncate jumps and further develop a consistent non-parametric jump robust estimator of the integrated volatility. Corsi et al. (2010) introduce threshold bipower variation by combining the concepts from Barndorff-Nielsen et al. (2003) and Mancini (2009). Jacod et al. (2014) estimate local volatility by using the empirical characteristic function of the return and then remove bias due to jump variation. When combining both jumps and microstructure noise in the price process, Fan and Wang (2007) propose a wavelet-based multi-scale approach to estimate integrated volatility. Podolskij et al. (2009) design modulated bipower variation, an estimator that filters the impact of microstructure noise then use bipower variation for volatility estimation. Andersen et al. (2012) use the concept of "nearest neighbor truncation" to establish jump and noise robust volatility estimators. On the other hand Brownlees et al. (2016) create truncated two scaled realized volatility by adopting a jump signaling indicator as in Mancini (2009) and noise robust sub-sampling as in Zhang et al. (2005). In addition to the above mentioned work, discussion regarding nonparametric estimation of integrated volatility and functionals of volatility can also be found in Barndorff-Nielsen et al. (2006), Mykland and Zhang (2009), Todorov and Tauchen (2012), Hautsch and Podolskij (2013), Jacod et al. (2013), Jing et al. (2014) and Jacod et al. (2017). What follows in the next section, is a detailed review of *12* of the most commonly used integrated volatility measures.[2]

3.1 Realized Volatility (RV)

Realized Volatility or RV as developed in Andersen et al. (2001) is one of the first empirical measures that used high-frequency intra-day returns to compute daily return variablility without having to explicitly model the intra-day data. The authors show that under suitable conditions RV is an unbiased and highly efficient estimator of QV as in (4). By extension it can be shown that in the absence of jumps or when jumps populate the data infrequently, RV converges in probability to IV as $M \longrightarrow \infty$. It should also be noted that RV has been used widely as part of the HAR-RV forecasting models. Here

$$RV_{t,M} = \sum_{j=1}^{M-1} (X_{t+(j+1)/M} - X_{t+j/M})^2 \tag{12}$$

3.2 Realized Bipower Variation (BPV)

In Barndorff-Nielsen and Shephard (2004), the authors demonstrate that they could untangle the continuous component of quadratic variation from its discontinuous component (jumps). This led them to develop *Realized Bipower Variation* (BPV), one of the first asymptotically unbiased estimators of IV which was robust to the presence of price jumps. It takes the following form

$$BPV_{t,M} = (\mu_1)^{-2} \sum_{j=2}^{M-1} |\Delta_j X||\Delta_{j-1} X| \tag{13}$$

where $\Delta_j X$ is the same as in (10) and $\mu_1 = 2^{\frac{1}{2}} \frac{\Gamma(1)}{\Gamma(\frac{1}{2})}$.

3.3 Tripower Variation (TPV)

The *Realized Bipower Variation* does not allow the consistency of the IV estimate to be impacted by finite activity jumps. However it is subject to finite sample jump distortions

[2]We follow the notation and description as in Mukherjee and Swanson (2018)

or upward bias. To counter this problem, BPV is generalized to *Tripower Variation* in Barndorff-Nielsen and Shephard (2004), by utilizing products of the (lower order) power of three adjacent intra-day returns. Theoretically speaking, although *Tripower Variation* (TPV) is more efficient, it is also more vulnerable to microstructure noise of the high frequency return data compared to BPV. TPV can be given as

$$TPV_{t,M} = (\mu_{\frac{2}{3}})^{-3} \sum_{j=3}^{M-1} |\Delta_j X|^{2/3} |\Delta_{j-1} X|^{2/3} |\Delta_{j-2} X|^{2/3} \tag{14}$$

where $\Delta_j X$ is the same as in (10) and $\mu_{\frac{2}{3}} = 2^{\frac{1}{3}} \frac{\Gamma(\frac{5}{6})}{\Gamma(\frac{1}{2})}$.

3.4 Two Scale Realized Volatility (TSRV)

It is found that when the sampling interval of the asset prices is small, microstructure noise issues become more prominent and *Realized Volatility* ceases to function as a robust volatility estimator. Due to the bias introduced by the market microstructure noise in the finely sampled data, initially longer time horizons are preferred by econometricians. It is found that ignoring microstructure noise works well for intervals more 10 minutes. However sampling over lower frequencies does not quantify and correct the noise effect on volatility estimation. As a solution, *Two Scale Realized Volatility* ($TSRV$) is introduced in Zhang et al. (2005) by combining estimators obtained over two time scales, avg and M. It forms an unbiased and consistent, microstructure noise robust estimator of IV in the absence of jumps. It takes the following form

$$TSRV_{t,M} = [X,X]^{avg} - \frac{1}{K}[X,X]^M \tag{15}$$

where

$$[X,X]^{m_i} = \sum_{j=1}^{m_i-1} (X_{t+((j+1)K+i)/M} - X_{t+(jK+i)/M})^2, \;\; i = 1,..,K \;\;\&\;\; m_i = \frac{M}{K} \tag{16}$$

$$[X,X]^{avg} = \frac{1}{K} \Sigma_{i=1}^{K} [X,X]^{m_i} \tag{17}$$

$$[X,X]^M = \sum_{j=1}^{M-1} (X_{t+(j+1)/M} - X_{t+j/M})^2 \tag{18}$$

$K = cM^{2/3}$ is the number of subsamples, $\frac{M}{K}$ is subsmaple size, $c > 0$ is a constant and M is the number of equispaced intra daily observations.

3.5 Multi Scale Realized Volatility (MSRV)

The $TSRV$ estimator though has many desirable properties, is not efficient. The rate of convergence for $TSRV$ is not satisfactory, it converges to the true volatility (IV in the absence of jumps) only at the rate of $M^{-1/6}$. The *Multi Scale Realized Volatility* ($MSRV$) is proposed in Zhang et al. (2006). This is a microstructure noise robust measure which converged to IV (in the absence of jumps) at the rate of $M^{-1/4}$. While $TSRV$ uses two time scales, $MSRV$ on the other hand uses N different time scales. $MSRV$ takes the following form

$$MSRV_{t,M} = \sum_{n=1}^{N} a_n [X,X]^{(M,K_n)}, \;\; n = 1,..,N \tag{19}$$

where

$$a_n = 12\frac{n}{N^2}\frac{n/N - 1/2 - 1/(2N)}{1 - 1/N^2}, \;\; \sum_{n=1}^{N} a_n = 1 \;\; \& \;\; \sum_{n=1}^{N} a_n/n = 0 \tag{20}$$

$$[X,X]^{(M,K_n)} = \frac{1}{K_n}\sum_{l=1}^{K_n}\sum_{j=1}^{m_{n,l}-1}(X_{t+((j+1)K_n+l)/M} - X_{t+(jK_n+l)/M})^2 \tag{21}$$

Here $l = 1,..,K_n$ & $m_{n,l} = \frac{M}{K_n}$. We take $N = 3, K_1 = 1, K_2 = 2, K_3 = 3$.

3.6 Realized Kernel (RK)

Barndorff-Nielsen et al. (2008) introduce Realized Kernel (RK) which as the name suggests is a realized kernel type consistent measure of IV in the absence of jumps. It is robust to endogenous microstructure noise and for particular choices of weight functions it can be asymptotically equivalent to $TSRV$ and $MSRV$ estimators, or even more efficient. RK can be given as

$$RK_{t,M} = \gamma_0(X) + \sum_{h=1}^{H}\kappa(\frac{h-1}{H})\{\gamma_h(X) + \gamma_{-h}(X)\} \tag{22}$$

where

$$\gamma_h(X) = \sum_{j=1}^{M-1}(X_{t+(j+1)/M} - X_{t+j/M})(X_{t+(j+1-h)/M} - X_{t+(j-h)/M}) \tag{23}$$

Here c is a constant. For our analysis we take a Turkey-Hanning$_2$ kernel which gives $\kappa(x) = sin^2\{\pi/2(1-x)^2\}$ and $H = cM^{1/2}$.

3.7 Truncated Realized Volatility (TRV)

Truncated Realized Volatility (TRV) is one of the first volatility measures that tried to estimate IV by identifying when price jumps greater than an adequately defined threshold occurred as in Aït-Sahalia et al. (2009). The truncation level for the jumps are chosen in a data-driven manner; the cutoff level α (given below) is set equal to a particular number times estimated standard deviations of the continuous part of the semimartingale. The price jump robust measure can be given as

$$TRV_{t,M} = \sum_{j=1}^{M-1} |\Delta_j X|^2 1_{\{|\Delta_j X| \leq \alpha \Delta_M^{\varpi}\}} \tag{24}$$

where

$$\alpha = 5\sqrt{\sum_{j=1}^{M-1} |\Delta_j X|^2 1_{\{|\Delta_j X| \leq \Delta_M^{1/2}\}}} \tag{25}$$

Here $\varpi = 0.47$. $\Delta_M = 1/M$

3.8 Modulated Bipower Variation (MBV)

Modulated Bipower Variation (MBV) as in Podolskij et al. (2009) consistently estimates IV and is robust to both market microstructure noise and finite activity jumps. It takes the following form

$$MBV_{t,M} = \frac{(c_1 c_2/\mu_1^2) mbv_{t,M} - \vartheta_2 \hat{\omega}^2}{\vartheta_1} \tag{26}$$

where

$$\vartheta_1 = \frac{c_1(3c_2 - 4 + max((2-c_2)^3, 0))}{3(c_2-1)^2}, \quad \vartheta_2 = \frac{2min((c_2-1), 1)}{c_1(c_2-1)^2} \tag{27}$$

$$mbv_{t,M} = \Sigma_{b=1}^{B} |\bar{X}_b^{(R)}||\bar{X}_{b+1}^{(R)}| \tag{28}$$

$$\bar{X}_b^{(R)} = \frac{1}{M/B - R + 1} \Sigma_{j=(b-1)M/B}^{bM/B-R} (X_{t+(j+R)/M} - X_{t+j/M}) \tag{29}$$

Here $c_1 = 2$, $c_2 = 2.3$, $R \approx c_1 M^{0.5}$, $B = 6$, $\mu_1 = 0.7979$, $\hat{\omega}^2 = \frac{1}{2M} RV_{t,M}$, $RV_{t,M}$ is given by (12).

3.9 Threshold Bipower Variation (TBPV)

Corsi et al. (2010) introduce a jump robust measure, *Threshold Bipwer Variation* ($TBPV$) which is constructed by combining the concepts of *Realized Bipower Variation* and *Threshold Realized Variance* (Mancini (2009). The authors show that $TBPV$ is robust to the choice of threshold function (υ as given below).

$$TBPV_{t,M} = \mu_1^{-2} \sum_{j=2}^{M-1} |\Delta_{j-1}X||\Delta_j X| I_{\{|\Delta_{j-1}X|^2 \leq \upsilon_{j-1}\}} I_{\{|\Delta_j X|^2 \leq \upsilon_j\}} \tag{30}$$

where

$$\upsilon_j = c_\upsilon^2 \hat{V}_j \tag{31}$$

$$\hat{V}_j^z = \frac{\Sigma_{i=-L}^{L} \kappa(\frac{i}{L})(\Delta_{j+i}X)^2 I_{\{(\Delta_{j+i}X)^2 \leq c_\upsilon^2 \hat{V}_{j+i}^{z-1}\}}}{\Sigma_{i=-L}^{L} \kappa(\frac{i}{L}) I_{\{(\Delta_{j+i}X)^2 \leq c_\upsilon^2 \hat{V}_{j+i}^{z-1}\}}}. \tag{32}$$

and $\Delta_j X$ is given by (10). Here we take $L = 25, c_\upsilon = 3, \hat{V}^0 = +\infty$. υ_j is the threshold for removal of large returns at each j. $\hat{V}_j^z$ gives estimated local variance in the presence of jumps at each iteration z for any j. Large returns are removed at each iteration according to $\{(\Delta_j X)^2 \leq c_\upsilon^2 \hat{V}_j^{z-1}\}$ and the estimated variance at that iteration is multiplied by c_υ^2 to get the threshold for the next iteration. When large returns cannot be removed any more, the iterations stop. Typically z is taken to be 2.

3.10 Subsampled Realized Kernel (SRK)

Barndorff-Nielsen et al. (2011) constructed *Subsampled Realized Kernel* (SRK) by combining the concepts of subsmapling (Zhang et al. (2005)) and realized kernels (Barndorff-Nielsen et al. (2008)). The main benefit of subsampling in this context is that it can overpower the inefficiency that stems from the poor selection of kernel weights that might be the case in *Realized Kernel*. SRK takes the following form

$$SRK_{t,M} = \frac{1}{S} \sum_{s=1}^{S} K^s(X) \tag{33}$$

where

$$K^s(X) = \gamma_0^s(X) + \sum_{h=1}^{H} \kappa(\frac{h-1}{H})\{\gamma_h^s(X) + \gamma_{-h}^s(X)\} \tag{34}$$

$$\gamma_h^s(X) = \sum_{j=1}^{\frac{M}{S}} x_j^s x_{j-h}^s \tag{35}$$

$$x_j^s = X_{t+(j+\frac{s-1}{S})/M} - X_{t+(j+\frac{s-1}{S}-1)/M} \tag{36}$$

Here the smooth Turkey-Hanning$_2$ kernel function gives $\kappa(x) = sin^2\{\pi/2(1-x)^2\}$, $S = 13$ and $H = 3$.

3.11 MedRV & MinRV

As alternatives to *Realized Bipower Variation* and *Tripower Variation*, Andersen et al. (2012) provide two alternative measures $MedRV$ and $MinRV$ which are robust to jumps and/or microstructure noise by using "nearest neighbor truncation". The basic concept behind these new measures is that the neighboring returns control the level of truncation of absolute returns. On one hand where $MinRV$ compares and takes the minimum of two adjacent absolute returns, $MedRV$ takes the median of three adjacent absolute returns and carries out two-sided truncation. Unlike the typical truncated realized measures as in Corsi et al. (2010), these new measures do not have to deal with the selection of an ex-ante threshold.

$$MinRV_{t,M} = \frac{\pi}{\pi-2}(\frac{M}{M-1}) \sum_{j=1}^{M-1} min(|\Delta_j X|, |\Delta_{j+1} X|)^2 \tag{37}$$

$$MedRV_{t,M} = \frac{\pi}{6-4\sqrt{3}+\pi}(\frac{M}{M-2}) \sum_{j=2}^{M-1} med(|\Delta_{j-1} X|, |\Delta_j X|, |\Delta_{j+1} X|)^2 \tag{38}$$

where $\Delta_j X$ is given by (10).

4 Jump Testing

Jump diffusion has been increasingly important in characterizing dynamic movement of asset prices. Early studies about jump diffusions can be seen in Andersen et al. (2002), Chernov et al. (2003), Pan (2002), and Eraker et al. (2003). Differentiating jumps from continuous process is particularly useful because it has implications for both researchers and practitioners in financial econometrics. Thus, a strand of literature has addressed the methodologies to identify jumps in the discretely sampled financial data. Aït-Sahalia (2002) rely on the transition density to test the existence of jumps under the option pricing model. Focusing on the risk-neutral dynamics of the underlying option prices, Carr and Wu (2003) propose a method to use the convergence rates of option prices to distinguish jumps from continuous process. Johannes (2004) propose a jump test to identify jump-induced misspecification. However, these tests only use limited low frequency data. With availability of high frequency data, the mechanism behind jump testing methodology has evolved. Barndorff-Nielsen and Shephard (2006) use the ratio of bipower variation and realized quadratic variation to construct a nonparametric test for the existence of jumps. Huang and Tauchen (2005) design extensive Monte Carlo experiments to evaluate the properties of newly proposed jump tests (see Andersen et al. (2003), Barndorff-Nielsen and Shephard (2004), and Barndorff-Nielsen and Shephard (2006)). Lee and Mykland (2007) propose tests to detect the exact timing of jumps at the intra-day level while Jiang and Oomen (2008) provide a "swap variance" approach to detect the presence of jumps. Mancini (2009) and Corsi et al. (2010) devise unique threshold or truncation techniques in their testing methodology. Aït-Sahalia et al. (2009) compare two higher order realized power variations to develop a test statistic for the null hypothesis of no jumps. On the other hand Podolskij and Ziggel (2010) combine the concepts truncated power variation and wild boostrap to propose a threshold-based jump test. In most of the above mentioned papers, the presence of realized jumps is tested over a "fixed time span". Corradi et al. (2014) and Corradi et al. (2018) proposed a "long time span" jump test instead, building on earlier work by Aït-Sahalia (2002). More related work on jump tests, self-excitation and mutual excitation in realized jumps can be found in Lee et al. (2013), Dungey et al. (2016), and Boswijk et al. (2018). In the next section we discuss six different jump tests which arise from different branches of the jump testing literature.

4.1 Barndorff-Nielsen and Shephard Test (BNS)

To test for the existence of jumps in the sample path of asset prices, Barndorff-Nielsen and Shephard (2006) propose non-parametric Hausman (1978) type tests using the difference between *Realized Quadratic Variation*, an estimator of integrated volatility which is not robust to jumps, and *Realized Bipower Variation*, which is a jump robust estimator of integrated volatility. *Realized Quadratic Variation* is considered to be the same as *Realized*

Volatility (RV). The adjusted jump ratio test statistic can be given as:

$$BNS = \frac{M^{1/2}}{\sqrt{\vartheta max(1, \frac{QPV}{(\mu_1^2 BPV)^2})}}(1 - \frac{BPV}{RV}) \xrightarrow{d} N(0,1) \tag{39}$$

where BPV is the same as in (13), RV is the same as in (12), $\vartheta =((\pi^2/4) + \pi$ - $5) \approx 0.6090$. The realized quadpower variation QPV is used to estimate integrated quarticity ($\int_0^t \sigma_s^4 ds$) and can be given as:

$$QPV = M \sum_{j=4}^{M} |\Delta_j X||\Delta_{j-1} X||\Delta_{j-2} X||\Delta_{j-3} X| \xrightarrow{d} \mu_1^4 \int_0^t \sigma_s^4 ds \tag{40}$$

The authors show that the null hypothesis of no jumps is rejected if the test statistic BNS is significantly positive.

4.2 Lee and Mykland Test (LM)

Lee and Mykland (2007) use the ratio of realized return to estimated instantaneous volatility, and further construct a nonparametric jump test to detect the exact timing of jumps at the intra-day level. The test statistic which identifies whether there is a jump during $(t+j/M, t+(j+1)/M]$ can be given as:

$$L_{(t+(j+1)/M)} = \frac{X_{t+(j+1)/M} - X_{t+j/M}}{\sigma_{\widehat{t+(j+1)/M}}} \tag{41}$$

where

$$\sigma_{\widehat{t+(j+1)/M}}^{\ 2} \equiv \frac{1}{K-2} \sum_{i=j-K+1}^{j-2} |X_{t+(i+1)/M} - X_{t+i/M} || X_{t+i/M} - X_{t+(i-1)/M}| \tag{42}$$

Here K is the window size of a local movement of the process. It is chosen in a way such that the effect of jumps on volatility estimation is eliminated. The authors suggest a value of K = 10 when the sampling frequency is 5-minute. Thus, it can be asymptotically shown that

$$\frac{max_{j \in \bar{A}_M} |L_{(t+(j+1)/M)}| - C_M}{S_M} \to \varepsilon, \quad as \;\; \Delta t \to 0, \tag{43}$$

where ε has a cumulative distribution function $P(\varepsilon \leq x) = exp(-e^{-x})$,

$$C_M = \frac{(2logM)^{1/2}}{c} - \frac{log\pi + log(logM)}{2c(2logM)^{1/2}} \quad and \quad s_M = \frac{1}{c(2logM)^{1/2}} \tag{44}$$

M is the number of intradaily observations, c ≈ 0.7979 and $\bar{A}_M$ is the set of $j \in \{0, 1, ..., M\}$ so that there are no jumps in $(t + j/M, t + (j+1)/M]$.

4.3 Jiang and Oomen Test (JO)

Jiang and Oomen (2008) compare a jump sensitive variance measure to realized volatility in order to test for jumps. Their idea is based on the fact that in the absence of jumps the accumulated difference between the simple return and log return (called the swap variance) captures one-half of the integrated volatility in the continuous time limit. Consequently it can be stated, in the absence of jumps the difference between swap variance and realized volatility should be zero, while in the presence of jumps the same difference reflects the replication error of variance swap thus detecting jumps. The swap variance can be given as

$$SV_{t,M} = 2\sum_{j=1}^{M-1}(\Delta_j P - \Delta_j X) \tag{45}$$

where $Y = log(P)$ and Y is the same as in (1). $\Delta_j P = \frac{P_{t+(j+1)/M}}{P_{t+j/M}} - 1$ and $\Delta_j X$ is the same as in (10). The three different swap variance tests proposed by the authors can be given as
(i) The difference test:

$$\frac{M}{\Omega_{SV}}(SV_{t,M} - RV_{t,M}) \xrightarrow{d} N(0,1) \tag{46}$$

(ii) The logarithmic test:

$$\frac{BPV_{t,M}M}{\Omega_{SV}}(log(SV_{t,M}) - log(RV_{t,M})) \xrightarrow{d} N(0,1) \tag{47}$$

(iii) The ratio test:

$$\frac{BPV_{t,M}M}{\Omega_{SV}}(1 - \frac{RV_{t,M}}{SV_{t,M}}) \xrightarrow{d} N(0,1) \tag{48}$$

where $\Omega_{SV} = \frac{\mu_6}{9}\frac{M^3\mu_{6/p}^{-p}}{M-p+1}\sum_{j=1}^{M-p}\prod_{k=0}^{p}|\Delta_{j+k}X|^{6/p}$ for $p \in \{1, 2, ...\}$, $\mu_z = E(|x|^z)$ for $z \sim N(0,1)$.

4.4 Aït-Sahalia and Jacod Test (ASJ)

In Aït-Sahalia et al. (2009), the authors develop a testing methodology for jumps in the (log) price process by comparing two higher order realized power variations with different sampling intervals, $k\Delta$ and Δ respectively. In this context $\Delta = \frac{1}{M}$, M is the number of intra-daily observations and k is a given integer. The pth order realized power variation can be given as

$$\widehat{B}(p,\Delta) = \sum_{j=1}^{M-1} |X_{t+(j+1)/M} - X_{t+j/M}|^p \tag{49}$$

The ratio of the two realized power variations with different sampling intervals takes the following form

$$\widehat{S}(p,k,\Delta) = \frac{\widehat{B}(p,k\Delta)}{\widehat{B}(p,\Delta)} \tag{50}$$

The corresponding jump test statistic can then be defined as,

$$ASJ = \frac{k^{(p/2)-1} - \widehat{S}(p,k,\Delta)}{\sqrt{V_{t,M}}} \xrightarrow{d} N(0,1) \tag{51}$$

where $V_{t,M}$ can be estimated using either a truncation technique as in

$$\widehat{V}_{t,M} = \Delta \frac{\widehat{A}(2p,\Delta)M(p,k)}{\widehat{A}(p,\Delta)^2} \tag{52}$$

where

$$\widehat{A}(2p,\Delta) = \frac{\Delta^{1-p/2}}{\mu_p} \sum_{j=1}^{M-1} |X_{t+(j+1)/M} - X_{t+j/M}|^p 1_{\{|X_{t+(j+1)/M} - X_{t+j/M}| \leq \alpha \Delta^{\varpi}\}} \tag{53}$$

or using multipower variation as in

$$\widehat{V}_{t,M} = \Delta \frac{M(p,k)\bar{A}(p/([p]+1), 2[p]+2, \Delta)}{\bar{A}(p/([p]+1), [p]+1, \Delta)^2} \tag{54}$$

where

$$\bar{A}(r,q,\Delta) = \frac{\Delta^{1-qr/2}}{\mu_r^q} \sum_{j=q}^{M-q+1} \prod_{i=0}^{q-1} |X_{t+(j+i)/M} - X_{t+(j+i-1)/M}|^r, \tag{55}$$

$$M(p,k) = \frac{1}{\mu_p^2}(k^{p-2}(1+k)\mu_{2p} + k^{p-2}(k-1)\mu_p^2 - 2k^{p/2-1} - \mu_{k,p}) \tag{56}$$

and $\mu_r = E(|U|^r)$ and $\mu_{k,p} = E(|U|^p|U + \sqrt{(}k-1)V|^p|)$ for $U,V \sim N(0,1)$. The null hypothesis of no jumps is rejected when the test statistic ASJ is significantly positive.

4.5 Podolskij and Ziggel Test (PZ)

In Podolskij and Ziggel (2010) the concept of truncated power variation is used to construct test statistics which diverge to infinity if jumps are present and have a normal distribution otherwise. The jump testing procedure in this paper is valid (under weak assumptions) for all semi-martingales with absolute continuous characteristics and general models for the noise processes. The methodology followed by the authors is a modification of that proposed in Mancini (2009). In particular they consider,

$$T(X,p) = M^{\frac{p-1}{2}} \sum_{j=1}^{M-1} |X_{t+(j+1)/M} - X_{t+j/M}|^p (1 - \eta_i 1_{\{|X_{t+(j+1)/M} - X_{t+j/M}| \leq \alpha\Delta^{\varpi}\}}) \tag{57}$$

where $\{\eta_i\}_{i\in[1,1/\Delta]}$ is a sequence of positive $i.i.d$ random variables. The test statistic has the following form

$$PZ = \frac{T(X,p)}{Var^*(\eta)\widehat{A}(2p,\Delta)} \xrightarrow{d} N(0,1) \tag{58}$$

where $\widehat{A}(2p,\Delta)$ is the same as in (53).

4.6 Corradi, Silvapulle and Swanson Test (CSS)

Building on previous work by Aït-Sahalia (2002), Corradi et al. (2018) design "long time span" jump tests based on realized third moments or "tricity" for the the null hypothesis that the probability of a jump is zero. This jump testing methodology is used to detect jumps by examining the "jump intensity" parameter in the data generating process rather than realized jumps over a "fixed time span". This test is of immense value when one is interested in using jump diffusion processes for valuation problems like options pricing and default modeling. Let,

$$\widehat{\mu}_{3,T,\Delta} = \frac{1}{T}\sum_{j=1}^{n-1}(X_{t+(j+1)/M} - X_{t+j/M} - \frac{X_{t+n/M} - X_{t+1/M}}{n})^3$$
$$- \frac{1}{T^+}\sum_{j=1}^{n^+-1}(X_{t+(j+1)/M} - X_{t+j/M} - \frac{X_{t+n^+/M} - X_{t+1/M}}{n^+})^3 1\{|X_{t+(j+1)/M} - X_{t+j/M}| \leq \tau(\Delta)\} \tag{59}$$

where we have n^+ observations over an increasing time span of T^+, a shrinking discrete sampling interval $\Delta = \frac{1}{M}$, so that $n^+ = \frac{T^+}{\Delta}$, $T^+ \to \infty$ and $\Delta \to 0$. $\tau(\Delta)$ is the truncation parameter and one example for the choice of such truncation can be given as follows. If σ_s as in (1) is a square root process, so that all moments exist, we can set $\tau(\Delta) = c\Delta^\eta$ with $\frac{2}{7} < \eta < \frac{1}{2}$. The authors define $n = \frac{T}{\Delta} = n^+ - \frac{T^+-T}{\Delta}$, with $T^+ > T$ and $\frac{T^+}{T} \to \infty$. Then, the test statistic for the null hypothesis of no jumps can be given as

$$CSS = \frac{T^{1/2}}{\Delta}\widehat{\mu}_{3,T,\Delta} \xrightarrow{d} N(0, \omega_0) \tag{60}$$

where ω_0 is defined in Corradi et al. (2018). Since, under the alternative hypothesis of positive jump intensity, the variance of the statistic is of larger order, it is difficult to construct a variance estimator which is consistent under all hypotheses. The authors use a threshold variance estimator, which removes the contribution of the jump component thus developing an estimator for the variance of CSS which is consistent under the null hypothesis of no jumps. Thus we have

$$\widehat{\sigma}^2_{CSS} = \frac{1}{\Delta^2}\sum_{j=0}^{n-1}(X_{t+(j+1)/M} - X_{t+j/M} - \frac{X_{t+n/M} - X_{t+1/M}}{n})^3 1\{|X_{t+(j+1)/M} - X_{t+j/M}| \leq \tau(\Delta)\} \tag{61}$$

Thus the t-statistic version of the jump test is

$$t_{CSS} = \frac{CSS}{\widehat{\sigma}_{CSS}} \tag{62}$$

5 Co-jump Testing

While univariate jump tests have been researched extensively, the study of co-jump tests has started growing only recently. One branch of literature proposes co-jump tests through identifying jumps in a portfolio. For example, Bollerslev et al. (2008) use observed return product to construct a test statistic for detecting co-jumps in an equiweighted index constructed from 40 stocks. Their co-jump test detects the modest-sized common jumps ignored in the Barndorff-Nielsen and Shephard (2004) jump test approach. Another branch uses univariate jump tests to identify co-jump among multivariate process. For example, Gilder et al. (2014) propose a co-exceedance rule to identify co-jumps by using univariate jump tests. Their Monte Carlo results show that the co-exceedance rule has similar power to the co-jump test proposed by Bollerslev et al. (2008). The third strand develops co-jump tests which can be directly applied to multiple price processes (see, e.g., Jacod and Todorov (2009), Bandi and Reno (2016), Bibinger and Winkelmann (2015) and Caporin et al. (2017)). Jacod and Todorov (2009) propose co-jump tests based on two null hypotheses: (i) there are common jumps in a bivariate process; (ii) there are disjoint jumps in a bivariate process. Mancini and Gobbi (2012) construct threshold estimators for integrated covariation from the realized covariation and show that the central limit theorem and robustness to nonsynchronous data still hold under different scenarios. Gnabo et al. (2014) propose a co-jump test based on bootstrapping methods. Bandi and Reno (2016) develop a nonparametric infinitesimal moments method to detect co-jumps between asset returns and volatilities. Bibinger and Winkelmann (2015) propose a spectral estimation method to detect co-jumps in multivariate high-frequency data in the presence of market microstructure noise and asynchronous observations. Caporin et al. (2017) build a co-jump test on the comparison between smoothed realized variance and smoothed random realized variation. More related literature about co-jumps can also be seen in Lahaye et al. (2011) and Dungey et al. (2011). In the following section, we discuss five most widely used co-jump tests in details. [3]

5.1 BLT Co-jump Testing

Bollerslev et al. (2008) propose a mcp test to detect co-jumps in a large ensemble of stocks. They develop a theoretical foundation which shows how only co-jumps (not idiosyncratic jumps) can be detected in a large equiweighted index. Let n denote the total number of

[3] We follow the notation and description as in Peng and Swanson (2018)

assets under co-jump detection. The mcp mean cross-product test statistic is defined as:

$$mcp_{t,j} = \frac{2}{n(n-1)} \sum_{i=1}^{n-1} \sum_{l=i+1}^{n} \Delta_j X^i \Delta_j X^l, j = 1, ..., M-1, t = 1, ..., T \tag{63}$$

where

$$\Delta_j X^i = X^i_{t+(j+1)/M} - X^i_{t+j/M}, \quad for \quad i = 1, ..., n \tag{64}$$

Since the mcp-statistic has nonzero mean and is analogous to a U-statistic, the studentized test statistic is:

$$z_{mcp,t,j} = \frac{mcp_{t,j} - \overline{mcp_t}}{s_{mcp,t}}, \quad for \quad j = 1, ..., M-1 \quad and \quad t = 1, ..., T. \tag{65}$$

where

$$\overline{mcp_t} = \frac{1}{M-1} mcp_t = \frac{1}{M-1} \sum_{j=1}^{M-1} mcp_{t,j} \tag{66}$$

and

$$s_{mcp,t} = \sqrt{\frac{1}{M-1} \sum_{j=1}^{M-1} (mcp_{t,j} - \overline{mcp_t})^2} \tag{67}$$

The null distribution under the null hypothesis of no jump is derived from bootstrapping the test statistics $z_{mcp,t,j}$ under Monte Caro simulations.

5.2 JT Co-jump Testing

Jacod and Todorov (2009) construct two test statistics to identify co-jumps under two different null hypothesis: i. There is at least one common jump under the null hypothesis; ii. There is at least one disjoint jump under the null hypothesis. The test statistics are proposed for detecting co-jumps on bivariate processes for the path of $s \longrightarrow X_s$ on $[0, t]$. Co-jumps among multivariate processes can be detected from the combination of bivariate processes. The test statistics of the common jump $\Phi_n^{(j)}$ and disjoint jump $\Phi_n^{(d)}$ are defined as:

$$\Phi_n^{(j)} = \frac{V(f, k\Delta_n)_t}{V(f, \Delta_n)_t} \tag{68}$$

$$\Phi_n^{(d)} = \frac{V(f, \Delta_n)_t}{\sqrt{V(g_1, \Delta_n)_t V(g_2, \Delta_n)_t}} \tag{69}$$

where k is an integer greater than 1, and $\Delta_n = \frac{t}{M}$ is the length of equispaced intra-daily time interval. $V(f, k\Delta_n)_t$ is defined as:

$$V(f, k\Delta_n)_t = \sum_{j=1}^{[t/k\Delta_n]} f(X_{(j+1)k/M} - X_{jk/M}) \tag{70}$$

Where the functions for $f(x)$, $g_1(x)$ and $g_2(x)$ are defined as:

$$f(x) = (x_1 x_2)^2, g_1(x) = (x_1)^4, g_2(x) = (x_2)^4 \tag{71}$$

They propose asymptotic properties and central limit theorems of these two test statistics when the mesh Δ_n approaches 0. They show that the test statistics for the null hypothesis with disjoint jumps $\Phi_n^{(d)}$ converges stably in law to 0 on $\Omega_T^{(d)}$ and the null hypothesis with common jumps $\Phi_n^{(j)}$ converges stably in law to 1 on $\Omega_T^{(d)}$. Here $\Omega_T^{(j)}$ and $\Omega_T^{(d)}$ are defined as:

$$\Omega_T^{(j)} = \{\omega\text{: on } [0, t] \text{ the process } \Delta_j X^1 \Delta_j X^2 \text{ is not identically } 0\} \tag{72}$$

$$\begin{aligned} \Omega_T^{(j)} = \{\omega\text{: on } [0, t] \text{ the processes } \Delta_j X^1 \text{ and } \Delta_j X^2 \text{ are} \\ \text{not identically 0, but the process } \Delta_j X^1 \Delta_j X^2 \text{ is}\} \end{aligned} \tag{73}$$

Where $\Delta_j X^i = X^i_{(j+1)/M} - X^i_{j/M}$, for $i = 1, 2$ and $j = 1, ..., M-1$. The authors construct critical regions of the two statistics as:

$$C_n^{(j)} = \{|\Phi_n^{(j)} - 1| \geq c_n^{(j)}\} \tag{74}$$

$$C_n^{(d)} = \{\Phi_n^{(d)} \geq c_n^{(d)}\} \tag{75}$$

5.3 MG Threshold Co-jump Test

Mancini and Gobbi (2012) use a threshold r_h to estimate each co-jump as:

$$\Delta_j X^1 \Delta_j X^2 - \Delta_j X^1 1_{\{(\Delta_j X^1)^2 \leq r_h\}} \Delta_j X^2 1_{\{(\Delta_j X^2)^2 \leq r_h\}}, \tag{76}$$

Where h is the length of observations interval and $h = \frac{t}{M}$ for every $j = 1, ..., M$. Threshold r_h is defined by a deterministic function from $h \to r_h$, with the following properties:

$$\lim_{h \to 0} r_n = 0 \quad and \quad \lim_{h \to 0} (hlog\frac{1}{h})/r_h = 0.$$

The threshold r_h depends on an unknown realized instantaneous volatility path. Monte Carlo simulations are used under different models to select a reasonable threshold. For example, in the model of stochastic volatility and finite compound Poisson jump part, the optimal choice of threshold is $r_h = 0.33\widehat{IC_{t,M}}h^{0.99}$, where the integrated covariation estimator $\widehat{IC_{t,M}}$ is derived by Mancini (2001):

$$\widehat{IC_{t,M}} = \tilde{v}_{1,1}^{(M)}(X^1, X^2)_t, \tag{77}$$

where

$$\tilde{v}_{1,1}^{(M)}(X^1, X^2)_t = h^{-3} \sum_{j:t_j \leq t} \Delta_j X^1 1_{\{(\Delta_j X^1)^2 \leq r_h\}} \Delta_j X^2 1_{\{(\Delta_j X^2)^2 \leq r_h\}}, \tag{78}$$

5.4 GST Co-exceedance Rule

Gilder et al. (2014) propose a co-exceedance based co-jump detection method by applying univariate jump tests to individual stocks to identify co-jumps. They select three univariate jump tests in Barndorff-Nielsen and Shephard (2006), Lee and Mykland (2007) and Andersen et al. (2010). The co-jumps are detected as intersection between ABD jump test results and BNS jump test results (ABD∩BNS), intersection between ABD jump test results and LM jump test results (ABD∩LM), intersection between BNS jump test results and LM jump test results (BNS∩LM), and the intersection among three jump tests results (ABD∩LM∩BNS).

The nonparametric BNS jump test and LM jump test have been discussed in subsection 4.1 and subsection 4.2 respectively. The ABD jump test in Andersen et al. (2010) is the sequential BNS test which first identifies jump days through BNS test and then calculates the maximum intra-day return as the jump level. Gilder et al. (2014) modified the maximum intra-day return during jump days into:

$$max\big(\mid \Delta_j X \mid / \sqrt{\widehat{s_{WSD,j}}^2 \cdot \Delta \cdot BPV_t}\big), \quad for j = 1, ..., M-1 \tag{79}$$

where $\Delta_j X = X_{t+(j+1)/M} - X_{t+j/M}$ for $t = 1, ..., T$ and $\Delta = \frac{1}{M}$. Here $\widehat{s_{WSD,j}}^2$ is the weighted standard deviation (WSD) estimator proposed by Boudt et al. (2011).

Comparisons between co-exceedance rule for co-jump detection and BLT co-jump test are made under extensive Monte Carlo simulations. The results show that intra-day co-exceedance based detection method has similar power to that of the BLT co-jump test both on large and small co-jumps.

5.5 CKR Co-jump Testing

The test statistics in Caporin et al. (2017) is derived from the difference between smoothed realized variance ($\widetilde{SRV}$) and smoothed randomized realized variance ($SRRV$). The SRRV is denoted as:

$$SRRV(X^i) = \sum_{j=1}^{M} \mid \Delta_j X^i \mid^2 \cdot K\left(\frac{\Delta_j X^i}{H^i_{\Delta_j,M}}\right) \cdot \eta^i_j, \tag{80}$$

$$i = 1, ..., n, \tag{81}$$

where $K(\cdot)$ is a differentiable kernel function with bounded first derivative almost everywhere in R having the following properties:

$$K(0) = 1, 0 \leq K(\cdot) \leq 1, \quad and \quad \lim_{x \to \infty} K(\mid x \mid) = 0 \tag{82}$$

And H is the bandwidth which is denoted as:

$$H^i_{\Delta_j,M} = h_M \cdot \widehat{\sigma_{\Delta_j}}^i \sqrt{\frac{t}{M}} \tag{83}$$

where h_M is the bandwidth parameter and $\widehat{\sigma_{\Delta_j}}^i$ is the point estimator of the local standard deviation of ith asset. η_j^i is an $n \times M$ matrix independent and identically distributed variable such that $E[\eta_j^i] = 1$ and $Var[\eta_j^i] = V_\eta \leq \infty$. V_η is set to 0.0025 in the application of the test.

Another estimator $(\widetilde{SRV})$ is written in the form as:

$$\widetilde{SRV}^n(X^i) = \sum_{j=1}^{M} \mid \Delta_j X^i \mid^2 \cdot \left(K\left(\frac{\Delta_j X^i}{H^i_{\Delta_j,M}}\right) + \pi^n_{k=1}\left(1 - K\left(\frac{\Delta_j X^k}{H^k_{\Delta_j,M}}\right)\right)\right) \tag{84}$$

The proposed test statistics takes the form:

$$S_{M,n} = \frac{1}{V_\eta}\sum_{i=1}^{M} \frac{\left(SRRV(X^i) - \widetilde{SRV}^n(X^i)\right)^2}{SQ(X^i)}, \tag{85}$$

where

$$SQ(X^i) = \sum_{j=1}^{M} \mid \Delta_j X^i \mid^4 \cdot K^2\left(\frac{\Delta_j X^i}{H^i_{\Delta_j,M}}\right), \quad i = 1, ..., n \tag{86}$$

The asymptotic behavior of the $S_{n,N}$ is described as:

$$S_{M,n} \xrightarrow{d} \chi^2(n), \quad on \quad \overline{\Omega}^n_T$$
$$S_{M,n} \xrightarrow{p} +\infty, \quad on \quad \overline{\Omega}^{MJ,n}_T$$

Where $\overline{\Omega}^n_T$ and $\overline{\Omega}^{MJ,n}_T$ is defined as:

$$\overline{\Omega}^{MJ,n}_T = \left\{ \omega \in \Omega \mid \Pi^n_{i=1}(\Delta X^j)_t \text{ is not identically } 0 \right\},$$
$$\overline{\Omega}^n_T = \Omega / \overline{\Omega}^{MJ,n}_T$$

6 Empirical Experiments

6.1 Data Description

The empirical experiments are conducted with six stocks and two ETFs. The six individual stocks which include the Boeing Company (BA), Exxon Mobile Corporation (XOM), Johnson & Johnson (JNJ), JPMorgan Chase & Co. (JPM), Microsoft Corporation (MSFT) and Walmart Inc.(WMT), have the highest weight in their corresponding SPDR market sector ETFs such as XLI (industrial sector), XLE (energy sector), XLV (healthcare sector), XLF (finance sector), XLK (technology sector) and XLP (consumer staples sector). The two SPDR sector ETFs chosen are the energy and technology sector ETFs, XLE & XLK. The dataset is obtained from the Trade and Quote Database (TAQ) of Wharton Research Data Service (WRDS) and it covers the period from January 1st, 2006 to December 31st, 2013 for a total of 2013 days. We select trade data ranging from 9:30 am to 4 pm on regular trading days. Overnight transactions are excluded from our dataset. We mainly use a 5-minute sampling frequency to eradicate the effect of market microstructure noise in the data which yields 78 total observations per day. We also use a 1-minute sampling frequency in specific cases which yields 390 observations per day. It should be noted that all empirical experiments are carried out on the logarithmic values of the stock and ETF prices.

6.2 Methodology

Our empirical experiment consists of three sections; (i) integrated volatility measures, (ii) jump tests and (iii) co-jump tests. For each of the different parts, we conduct analysis involving the most widely used measures and tests respectively. A detailed description of the different measures & tests used and the empirical methodologies thereof is given as follows.

Firstly we use six different measures to estimate Integrated Volatility for all the stocks and ETFs; (1) Realized Volatility (3.1), (2) Bipower Variation (3.2), (3) Tripower Variation (3.3), (4) Truncated Realized Volatility (3.7), (5) MedRV & (6) MinRV (3.11). Secondly to test for price jumps in the data three different jump tests are used; (1) ASJ jump test (4.4), (2) BNS jump test (4.1), (3) LM jump test (4.2). Lastly co-jump tests are carried out using (1) JT co-jump test (5.2), (2) BLT co-jump test (5.1) and (3) GST coexceedance rule (5.4).

Estimation of integrated volatility, BNS and LM jump tests as well as all the co-jump tests are carried out using 5 minute data where Δ is set to $\frac{1}{78}$. However for the ASJ jump test, both 1-minute ($\Delta = \frac{1}{390}$) and 5-minute frequencies are used as a basis for comparative study.

When conducting analysis using jump tests, we calculate the percentage of days identified as having jumps. For both the BNS and ASJ tests, it can be given as:

$$\text{Percentage of Jump Days} = \frac{100\sum_{i=1}^{T} I(Z_i > c_\alpha)}{T}\% \tag{87}$$

where $I(\cdot)$ is the jump indicator function, c_α is the critical value at α significance level and Z_i is the BNS or ASJ jump test statistics. For the LM jump test on the other hand it can be derived as:

$$\text{Percentage of Jump Days} = \frac{100\sum_{i=0}^{T} I(\exists t \in i, \mid L_t \mid > c_\alpha)}{T}\% \tag{88}$$

where L_t is the LM jump test statistic at the intra-day level within a particular day, t refers to the 78 intra-day intervals and c_α is the critical value at α significance level.

Once jumps are detected, we follow Andersen et al. (2007) and Duong and Swanson (2011) to construct risk measures by separating out the variation due to daily jump component and the continuous components. This is done by using volatility measures RV and TPV. It can be given as

$$\text{Variation due to Jump Component} = JV_t = max[RV_t - TPV_t, 0] * I_{jump,t} \tag{89}$$

Consequently the ratio of jump to total variation for all three jump tests can be calculated as,

$$\text{Ratio of Jump Varitaion to Total Variation} = \frac{JV_t}{RV_t} \tag{90}$$

For BLT co-jump test, the percentage of days identified as having co-jumps is calculated using,

$$\text{Percentage of Co-Jump Days} = \frac{100\sum_{i=0}^{T} I(\exists j, z_{mcp,i,j} < c_{mcp,\alpha,l} \cup z_{mcp,i,j} > c_{mcp,\alpha,r})}{T}\% \tag{91}$$

where $c_{mcp,\alpha,l}$ and $c_{mcp,\alpha,r}$ are left and right tail critical values derived from bootstrapping the null distribution. α is the significance level. For the JT co-jump test, the percentage of days identified as having co-jumps is calculated as:

$$\text{Percentage of Co-Jump Days} = \frac{100\sum_{i=0}^{T} I(\Phi_n^{(d)} \geq c_n^{(d)})}{T}\% \tag{92}$$

In the co-exceedance rule proposed by Gilder et al. (2014), we use the BNS jump test and the LM jump test to identify co-jumps. The percentage of days identified as having co-jumps can be given as:

$$\text{Percentage of Co-Jump Days} = \frac{100 \sum_{i=0}^{T} I(\mid Z_i \mid \geq \Phi_\alpha) * I(\exists t \in i, \mid L_t \mid > c_\alpha)}{T} \% \tag{93}$$

where Z_i is the BNS jump test statistic and L_t is the LM jump test statistic.

In addition to reporting the findings of our empirical experiment on the entire sample, we also conduct analysis after splitting the data set into two periods. The first sample consists of the period from January 2006 to June 2009 and the second sample consists of the period from July 2009 to December 2012. This is done to inspect whether the jump activity in the stocks and the ETFs changes considerably over time. The break date of our sample (June 2009) roughly corresponds to the end of the business cycle contraction after the financial crisis as given by NBER.

6.3 Findings

Table 1 gives the summary statistics for integrated volatility which is estimated using six volatility measures RV, BPV, TPV, $MedRV$, $MinRV$ and TRV. The sample period considered for the six stocks and the two ETFs is January 2006 - December 2013. The mean, standard deviation, minimum and maximum values are all in terms of 10^{-4}. Amongst all the stocks and ETFs, JPMorgan seems to have undergone maximum price fluctuations across the sample period as it displays the highest mean and max values across all the volatility measures. On the other hand Johnson & Johnson and XLK appear to be tied in terms of having undergone least amount of price fluctuations as they display the lowest mean and max volatility estimates. Amongst all the volatility measures, *Bipower Variation* reports the lowest mean volatility estimate while *Realized Volatility* reports the highest mean volatility estimate for any given stock or ETF. This can be explained by the fact that in the presence of frequent jumps, *Realized Volatility* overestimates integrated volatility. To get a clearer idea of how volatility differs across the stocks and ETFs, we turn to figures (1) - (2) which display the estimated volatility for the stocks Boeing and Exxon with respect to the six aforementioned volatility measures. Similar figures for 4 other stocks and 2 ETFs have not been given for the purpose of brevity and can be provided upon request. In general stocks and ETFs achieve their highest volatility in the fourth quarter of 2008 during the financial crisis with a few exceptions. For XLE, in case of all four volatility measures apart from TPV and TRV, volatility reaches its peak in the second quarter of 2009. For XLK on the other hand, only in case RV the volatility peak is reached in the first quarter of 2008 while for the other measures it is the fourth quarter of 2008.

We now look at tables (2 - 5) which display the descriptive statistics of the three jump tests. For the ASJ jump test we consider both 5-minute and 1-minute frequencies while for the BNS and the LM jump tests we only consider 5-minute frequency. Panel A in the tables refers to the pre-financial crisis sample period, January 2006 - June 2009 and Panel B refers to the post crisis period July 2009 - December 2012. In case of the ASJ jump tests, we find noticeable differences between 5-minute (table 2) and 1 minute (table 3) frequencies. Overall the mean value of the statistics is higher for the 1-minute data compared to the 5 minute frequency suggesting that more jumps would be identified in the 1-minute case. The skewness values are all negative irrespective of the sample period, type of stock and frequency of sampling suggesting that the ASJ test statistics are left-skewed. Panel A for both frequencies appear to have overall higher mean and max values again suggesting more jump activity in the financial crises period. In case of the BNS test (4) the skewness values are all positive, which suggests all BNS test statistics are right-skewed and have a long right tail. The kurtosis values are all above 3, which indicates the empirical distribution of BNS test statistics is leptokurtic. For the LM test (table 5), a window size of $k = 50$ is chosen. The mean of LM test statistics is around 0, while the max and min value of test statistics are far from 0, even reaching 1726.992 and -2124.609.

Tables (6 - 8) denote the percentage of days identified as having jumps for the ASJ, BNS and LM jump tests. For all jump tests $\alpha = 0.1$ and 0.05 significance levels are considered. In case of the ASJ jump test (table 6), it appears that Johnson & Johnson has the largest percentage of jump days for post-June 2009 period (panel B). However for the pre-June 2009 period (panel A), only with 5-minute frequency, Johnson & Johnson attains the highest jump day percentage. While for 1-minute frequency XLK seems to lead the race. XLE on the other hand has the lowest percentage of jump days across all significance levels, sample periods and sampling frequencies. In case of the BNS jump test (table 7) XLK has the largest percentage of jump days for the crisis (panel A) period while Microsoft displays the highest percentage in the post-crisis (panel B) period. Overall for all the stocks and

ETFs for both ASJ and BNS tests, panel A displays relatively higher jump activity than panel B which shows that jumps happen more frequently during financial crisis period, when compared with post financial crisis period. Table (8) shows the percentage of jump days and jump proportions for the LM test. The percentage of jump days is very large, reaching 90% in some cases. This is because LM jump test detects whether there is jump at each interval per day[4] and a day is classified as a jump day if a jump occurs on any of the 5-minute (78 observations) intervals. The jump proportion is calculated by the total number of test statistics which indicate jumps divided by the total number of test statistics across all intra-day intervals for the entire sample period. The jump proportions are much lower, close to 1%. It is noteworthy that both percentage of jump days and jump proportions are larger during post financial crisis period (panel B). One reason for this may be, the fact that the LM test detects more small and moderate jumps when compared with the ASJ and BNS tests and these types of jumps are more likely to happen during post financial crisis period.

To graphically illustrate the level of jump activity we turn to figures (3 - 5), which display the ASJ and BNS test statistic values for the days identified as having jumps for Boeing and Exxon across the sample period January 2006 - December 2013. Once again similar figures for other stocks and ETFs have not been given for the purpose of brevity and would be available upon request. The significance level considered is 5%. For the ASJ test, the analysis is carried out for both 5 (figure 3) and 1 minute (figure 4) frequencies. As is evident from the figures, with a higher sampling frequency of 1-minute, more jumps are detected across all stocks and ETFs in comparison to 5-minute frequency. In case of the BNS test both XLK and Johnson & Johnson appear to have the relatively higher degree of jump activity compared to the other stocks in the pre-June 2009 period, a result which evidently aligns with what we deduced from table (7).

Figures (6 - 8) contain the kernel density plots of ASJ, BNS and LM test statistics. In case of the ASJ test statistics (6), it appears that the distribution is left-tailed or negatively skewed. On the other hand the underlying distribution for the BNS test statistic (7) appears to be skewed right. The LM test statistics (8) display a high kurtosis and a long tail. All these results are consistent with what we found from tables (2-5).

When analyzing the average ratio of jump variation to total variation, we compare the results between the ASJ test, BNS test and LM test. For all three tests given in tables 9, 10

[4]Refer to equation (93)

and 11, ratio of jump variation to total variation is larger during financial crisis than post financial crisis and this result is robust across all significance levels. The tech sector ETF XLK has the largest jump variation ratio amongst all stocks. The BNS jump test is more likely to detect large jumps, especially during financial crisis period which is why the jump variation ratio reported by it is larger than the other tests.

Table (12) contains percentage of days identified as having co-jumps under both JT co-jump test and the co-exceedance rule between BNS jump test and LM jump test. Co-jumps are detected in case of each of the following pairwise stock combinations, including Exxon & JPMorgan, Exxon & Microsoft, Exxon & XLE, JPMorgan & Microsoft, Microsoft & XLK and XLE & XLK. The range of percentage of co-jump days in JT co-jump test is from 0.454% to 2.955%, while the range for the co-exceedance rule is from 2.838% to 9.545%. One reason for the larger percentage range in co-exceedance rule could be the fact, that the intersection results between two jump tests lead to a large false rejection rate. The percentage of co-jump days in JPMorgan & Microsoft, Microsoft & XLK and XLE & XLK is larger during the financial crisis period than post financial crisis period and this result is robust across the different significance levels and types of co-jump tests. In Table (13), we detect co-jumps among the six stocks (Boeing, Exxon, Johnson&Johnson, JPMorgan, Microsoft and Walmart), using the BLT co-jump test as in Bollerslev et al. (2008). As is clear from the table, the percentage of co-jump days as per the BLT test is small, ranging from 0.114% to 0.454%.

We now turn to discuss graphical representation of co-jumps. Figures have only been given for co-jumps between pairs Exxon & JPMorgan, Exxon & Microsoft. Figures involving co-jumps between other stock combinations are available upon request. Figure (9) denotes the kernel density plot of JT co-jump test. Overall the distribution of the test statistics appears to be heavily right tailed. Figure (11) shows the JT test statistics of co-jump days from year 2006 to 2013. It is clear that co-jumps are less densely populated when compared with jump days. When comparing how co-jumps are scattered between financial crisis period and post financial crisis period, there is no significant difference amongst Exxon & JPMorgan, Exxon & Microsoft. On the other hand more frequent co-jumps are visible during the financial crisis period in Microsoft & XLK and XLE & XLK. Figure (12) shows the days which have co-jumps as per the co-exceedance rule. The results show there is not much significant difference on how co-jumps are distributed between financial crisis and post financial crisis period.

Finally, figures (10) and (13) show the empirical findings from BLT co-jump tests. Figure (10) denotes the kernel density plot of empirical BLT test statistics. The distribution of the test statistics is evidently positively skewed. Figure (13) shows the daily return, daily closing price, realized variance, bipower variation and co-jump days for equi-weighted stock index. In Bollerslev et al. (2008), the authors show that detection of co-jumps among multiple stocks is equivalent to detecting co-jumps in an equi-weighted index composed by the same underlying stocks. Here we test co-jumps among six stocks, including Boeing, Exxon, Johnson&Johnson, JPMorgan, Microsoft and Walmart. The last panel of figure 13 shows the number of co-jump days at $\alpha = 10\%$ significance level. There are only 9 co-jump days among six stocks from year 2006 to 2013.

7 Conclusion

In this chapter, we review some of the most recent literature on integrated volatility measures, jump and co-jump tests. We then select a small subset of these measures and tests to conduct an empirical investigation with intra-day TAQ data of six individuals stocks and two ETFs. This study helps to reveal how the general volatility movement, jump and co-jump activity amongst the stocks vary across different types of tests and sampling frequencies.

We find that the occurrence of jumps is more frequent during and before the financial crisis period, i.e. January 2006 - June 2009 compared to the post financial crisis period, i.e. July 2009 - December 2013. All individual stocks apart from the ETFs reach their peak volatility in the fourth quarter of 2008. Overall, the incidence of co-jumps is lesser compared to jumps over the entire sample period i.e. January 2006 - December 2013. Additionally there is not much significant difference in terms of distribution of co-jumps between financial crisis and post financial crisis period.

References

Aït-Sahalia, Y., 2002. Telling from discrete data whether the underlying continuous-time model is a diffusion. The Journal of Finance 57, 2075–2112.

Aït-Sahalia, Y., Jacod, J., et al., 2009. Testing for jumps in a discretely observed process. The Annals of Statistics 37, 184–222.

Andersen, T.G., Benzoni, L., Lund, J., 2002. An empirical investigation of continuous-time equity return models. The Journal of Finance 57, 1239–1284.

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2003. Some like it smooth, and some like it rough: Untangling continuous and jump components in measuring, modeling, and forecasting asset return volatility. CFS Working Paper Series 2003/35, Center for Financial Studies (CFS) .

Andersen, T.G., Bollerslev, T., Diebold, F.X., 2007. Roughing it up: Including jump components in the measurement, modeling, and forecasting of return volatility. The review of economics and statistics 89, 701–720.

Andersen, T.G., Bollerslev, T., Diebold, F.X., Labys, P., 2001. The distribution of realized exchange rate volatility. Journal of the American statistical association 96, 42–55.

Andersen, T.G., Bollerslev, T., Frederiksen, P., Ørregaard Nielsen, M., 2010. Continuous-time models, realized volatilities, and testable distributional implications for daily stock returns. Journal of Applied Econometrics 25, 233–261.

Andersen, T.G., Dobrev, D., Schaumburg, E., 2012. Jump-robust volatility estimation using nearest neighbor truncation. Journal of Econometrics 169, 75–93.

Bandi, F.M., Reno, R., 2016. Price and volatility co-jumps. Journal of Financial Economics 119, 107–146.

Barndorff-Nielsen, O.E., Graversen, S.E., Jacod, J., Podolskij, M., Shephard, N., 2006. A central limit theorem for realised power and bipower variations of continuous semimartingales, in: From stochastic calculus to mathematical finance. Springer, pp. 33–68.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2008. Designing realized kernels to measure the ex post variation of equity prices in the presence of noise. Econometrica 76, 1481–1536.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A., Shephard, N., 2011. Subsampling realised kernels. Journal of Econometrics 160, 204–219.

Barndorff-Nielsen, O.E., Shephard, N., 2004. Power and bipower variation with stochastic volatility and jumps. Journal of financial econometrics 2, 1–37.

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

Barndorff-Nielsen, O.E., Shephard, N., et al., 2003. Realized power variation and stochastic volatility models. Bernoulli 9, 243–265.

Bibinger, M., Winkelmann, L., 2015. Econometrics of co-jumps in high-frequency data with noise. Journal of Econometrics 184, 361–378.

Bollerslev, T., Law, T.H., Tauchen, G., 2008. Risk, jumps, and diversification. Journal of Econometrics 144, 234–256.

Boswijk, H.P., Laeven, R.J., Yang, X., 2018. Testing for self-excitation in jumps. Journal of Econometrics 203, 256–266.

Boudt, K., Croux, C., Laurent, S., 2011. Robust estimation of intraweek periodicity in volatility and jump detection. Journal of Empirical Finance 18, 353–367.

Brownlees, C.T., Nualart, E., Sun, Y., 2016. A truncated two-scales realized volatility estimator .

Caporin, M., Kolokolov, A., Renò, R., 2017. Systemic co-jumps. Journal of Financial Economics 126, 563–591.

Carr, P., Wu, L., 2003. What type of process underlies options? a simple robust test. The Journal of Finance 58, 2581–2610.

Chernov, M., Gallant, A.R., Ghysels, E., Tauchen, G., 2003. Alternative models for stock price dynamics .

Corradi, V., Distaso, W., Swanson, N.R., 2011. Predictive inference for integrated volatility. Journal of the American Statistical Association 106, 1496–1512.

Corradi, V., Silvapulle, M., Swanson, N., 2014. Consistent pretesting for jumps .

Corradi, V., Silvapulle, M.J., Swanson, N.R., 2018. Testing for jumps and jump intensity path dependence. Journal of Econometrics 204, 248–267.

Corsi, F., Pirino, D., Reno, R., 2010. Threshold bipower variation and the impact of jumps on volatility forecasting. Journal of Econometrics 159, 276–288.

Dungey, M., Henry, Ó.T., Hvozdyk, L., 2011. The impact of thin trading and jumps on realized hedge ratios .

Dungey, M.H., Erdemlioglu, D., Matei, M., Yang, X., 2016. Financial flights, stock market linkages and jump excitation .

Duong, D., Swanson, N.R., 2011. Volatility in discrete and continuous-time models: A survey with new evidence on large and small jumps, in: Missing Data Methods: Time-Series Methods and Applications, pp. 179–233.

Eraker, B., Johannes, M., Polson, N., 2003. The impact of jumps in volatility and returns. The Journal of Finance 58, 1269–1300.

Fan, J., Wang, Y., 2007. Multi-scale jump and volatility analysis for high-frequency financial data. Journal of the American Statistical Association 102, 1349–1362.

Gilder, D., Shackleton, M.B., Taylor, S.J., 2014. Cojumps in stock prices: Empirical evidence. Journal of Banking & Finance 40, 443–459.

Gnabo, J.Y., Hvozdyk, L., Lahaye, J., 2014. System-wide tail comovements: A bootstrap test for cojump identification on the s&p 500, us bonds and currencies. Journal of International Money and Finance 48, 147–174.

Hausman, J., 1978. Specification tests in econometrics. Econometrica 46, 1251–1271.

Hautsch, N., Podolskij, M., 2013. Preaveraging-based estimation of quadratic variation in the presence of noise and jumps: theory, implementation, and empirical evidence. Journal of Business & Economic Statistics 31, 165–183.

Huang, X., Tauchen, G., 2005. The relative contribution of jumps to total price variance. Journal of financial econometrics 3, 456–499.

Jacod, J., Li, Y., Zheng, X., 2017. Estimating the Integrated Volatility with Tick Observations. Technical Report.

Jacod, J., Rosenbaum, M., et al., 2013. Quarticity and other functionals of volatility: efficient estimation. The Annals of Statistics 41, 1462–1484.

Jacod, J., Todorov, V., 2009. Testing for common arrivals of jumps for discretely observed multidimensional processes. The Annals of Statistics 37, 1792–1838.

Jacod, J., Todorov, V., et al., 2014. Efficient estimation of integrated volatility in presence of infinite variation jumps. The Annals of Statistics 42, 1029–1069.

Jiang, G.J., Oomen, R.C., 2008. Testing for jumps when asset prices are observed with noise–a âĂIJswap varianceâĂİ approach. Journal of Econometrics 144, 352–370.

Jing, B.Y., Liu, Z., Kong, X.B., 2014. On the estimation of integrated volatility with jumps and microstructure noise. Journal of Business & Economic Statistics 32, 457–467.

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

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

Lahaye, J., Laurent, S., Neely, C.J., 2011. Jumps, cojumps and macro announcements. Journal of Applied Econometrics 26, 893–921.

Lee, S.S., Mykland, P.A., 2007. Jumps in financial markets: A new nonparametric test and jump dynamics. The Review of Financial Studies 21, 2535–2563.

Lee, T., Loretan, M., Ploberger, W., 2013. Rate-optimal tests for jumps in diffusion processes. Statistical Papers 54, 1009–1041.

Mancini, C., 2001. Disentangling the jumps of the diffusion in a geometric jumping brownian motion. Giornale dellâĂŹIstituto Italiano degli Attuari 64, 44.

Mancini, C., 2009. Non-parametric threshold estimation for models with stochastic diffusion coefficient and jumps. Scandinavian Journal of Statistics 36, 270–296.

Mancini, C., Gobbi, F., 2012. Identifying the brownian covariation from the co-jumps given discrete observations. Econometric Theory 28, 249–273.

Mukherjee, A., Swanson, N., 2018. New direction for volatility confidence interval prediction: Evidence from an experimental and empirical study. Working Paper, Rutgers University.

Mykland, P.A., Zhang, L., 2009. Inference for continuous semimartingales observed at high frequency. Econometrica 77, 1403–1445.

Pan, J., 2002. The jump-risk premia implicit in options: Evidence from an integrated time-series study. Journal of financial economics 63, 3–50.

Peng, W., Swanson, N.R., 2018. Co-jumps, co-jump tests and volatility prediction in the s&p 500 market. Working Paper .

Podolskij, M., Vetter, M., et al., 2009. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. Bernoulli 15, 634–658.

Podolskij, M., Ziggel, D., 2010. New tests for jumps in semimartingale models. Statistical inference for stochastic processes 13, 15–41.

Todorov, V., Tauchen, G., 2012. The realized laplace transform of volatility. Econometrica 80, 1105–1127.

Zhang, L., Mykland, P.A., Aït-Sahalia, Y., 2005. A tale of two time scales: Determining integrated volatility with noisy high-frequency data. Journal of the American Statistical Association 100, 1394–1411.

Zhang, L., et al., 2006. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. Bernoulli 12, 1019–1043.

Table 1: Descriptive Statistics of Integrated Volatility Measures: Sample period Jan 2006 - Dec 2013

	Volatility Measures	RV	BPV	TPV	MedRV	MinRV	TRV
Boeing	Mean	2.68	1.24	2.29	2.38	2.46	2.60
	Standard Deviation	4.35	2.10	3.95	4.10	4.29	4.32
	Min	0.21	0.10	0.15	0.19	0.20	0.21
	Max	60.05	28.47	58.80	62.83	65.26	60.05
	Skewness	6.50	6.66	6.81	7.13	7.00	6.63
	Kurtosis	61.54	62.81	66.26	73.91	70.90	63.35
Exxon	Mean	2.15	1.01	1.89	1.98	2.02	2.10
	Standard Deviation	5.30	2.65	4.94	5.47	5.75	5.08
	Min	0.10	0.04	0.07	0.08	0.09	0.10
	Max	131.00	72.39	140.95	156.49	166.62	131.00
	Skewness	12.17	14.05	15.01	15.47	15.91	12.89
	Kurtosis	223.33	301.42	346.86	357.84	374.71	254.66
Johnson & Johnson	Mean	1.05	0.47	0.85	0.90	0.93	1.00
	Standard Deviation	2.37	1.06	1.94	2.07	2.12	2.30
	Min	0.07	0.02	0.03	0.05	0.05	0.07
	Max	52.86	22.84	46.48	42.82	43.33	52.86
	Skewness	11.36	10.44	11.25	10.99	10.61	11.76
	Kurtosis	186.83	157.67	195.54	173.99	162.96	203.20
JP Morgan	Mean	5.85	2.76	5.11	5.26	5.41	5.80
	Standard Deviation	13.99	65.23	12.00	12.24	12.49	13.96
	Min	0.13	0.05	0.09	0.11	0.10	0.13
	Max	244.81	118.46	213.02	214.80	235.11	244.81
	Skewness	8.06	7.74	7.48	7.34	7.42	8.10
	Kurtosis	100.06	94.76	88.59	85.07	90.95	101.07
Microsoft	Mean	2.29	1.07	1.95	2.07	2.12	2.24
	Standard Deviation	3.71	1.82	3.49	3.62	3.67	3.69
	Min	0.16	0.06	0.10	0.13	0.14	0.14
	Max	62.08	34.69	66.05	54.42	65.94	62.08
	Skewness	7.32	7.98	8.14	7.44	7.76	7.44
	Kurtosis	82.10	99.39	10.17	79.17	91.30	83.96
Walmart	Mean	1.57	0.71	1.31	1.37	1.40	1.51
	Standard Deviation	2.95	1.32	2.53	2.53	2.48	2.86
	Min	0.13	0.57	0.10	0.12	0.11	0.13
	Max	71.09	31.41	63.15	60.95	53.54	71.09
	Skewness	10.74	10.07	10.93	10.19	8.76	11.02
	Kurtosis	190.00	172.43	205.05	180.79	130.01	204.82
XLE	Mean	2.82	1.35	2.49	2.70	2.74	2.72
	Standard Deviation	7.35	3.45	5.77	8.52	8.42	6.00
	Min	0.10	0.04	0.08	0.08	0.08	0.10
	Max	193.76	80.53	123.39	278.12	266.10	129.66
	Skewness	14.29	12.30	10.14	20.21	18.81	10.39
	Kurtosis	296.35	215.74	151.36	577.77	511.09	161.29
XLK	Mean	1.46	0.65	1.19	1.28	1.30	1.38
	Standard Deviation	3.18	1.41	2.72	2.69	2.78	2.81
	Min	0.07	0.02	0.04	0.06	0.05	0.07
	Max	61.21	24.11	49.50	40.78	41.31	49.62
	Skewness	8.88	7.92	8.53	7.52	7.72	7.83
	Kurtosis	117.31	87.29	103.04	76.99	79.80	89.51

*Notes: Table 1 gives the descriptive statistics of the different integarted volatility measures. Mean, standard deviation, min and max values are all in terms of 10^{-4}

Table 2: Descriptive Statistics of ASJ Jump Test: 5-minute sampling frequency

		Boeing	Exxcon	Johnson&Johnson	JP Morgan	Microsoft	Walmart	XLE	XLK
Panel A	Mean	0.042	0.177	0.256	0.088	0.149	0.185	0.016	0.164
	st.dev	1.239	1.118	1.166	1.229	1.189	1.188	1.243	1.113
	skewness	-1.129	-1.149	-1.293	-1.236	-1.150	-1.340	-1.182	-1.134
	kurtosis	4.120	4.459	4.739	4.814	4.253	4.752	4.612	4.583
	max	2.413	2.648	2.330	2.427	2.530	2.666	2.938	2.526
	min	-5.374	-4.383	-4.681	-5.833	-4.319	-4.363	-5.746	-4.361
Panel B	Mean	-0.053	0.093	-0.009	-0.056	0.109	0.0541	-0.018	0.063
	st. dev	1.291	1.168	1.302	1.273	1.182	1.146	1.192	1.158
	Skewness	-1.074	-1.216	-1.211	-1.009	-1.079	-0.840	-1.150	-1.195
	Kurtosis	3.743	4.465	4.342	3.636	4.046	3.402	4.535	4.808
	Max	2.210	2.301	2.431	2.300	2.563	2.367	2.783	2.495
	Min	-5.095	-4.495	-5.685	-4.624	-4.437	-3.963	-5.770	-5.658

*Notes: Table 2 gives the descriptive statistics of ASJ jump test at 5-minute frequency. Panel A covers the financial crisis period from Jan 2006 - June 2009 and Panel B covers the post fiancial crisis period from July 2009 - Dec 2012.

Table 3: Descriptive Statistics of ASJ Jump Test: 1-minute sampling frequency

		Boeing	Exxcon	Johnson&Johnson	JP Morgan	Microsoft	Walmart	XLE	XLK
Panel A	Mean	0.399	0.404	0.587	0.391	0.612	0.453	0.125	0.756
	st. dev	1.455	1.345	1.414	1.315	1.300	1.358	1.371	1.385
	Skewness	-0.896	-0.776	-1.627	-1.035	-1.05	-0.845	-0.347	-1.071
	Kurtosis	4.378	5.219	10.193	6.531	5.650	4.770	3.682	9.357
	Max	4.723	5.138	4.538	4.105	5.457	4.998	4.548	6.183
	Min	-6.235	-6.387	-11.076	-8.702	-5.554	-5.755	-4.864	-8.529
Panel B	Mean	0.236	0.348	0.430	0.249	0.532	0.317	0.066	0.643
	st. dev	1.400	1.281	1.504	1.335	1.247	1.503	1.321	1.111
	skewness	-1.100	-1.122	-0.884	-1.294	-1.029	-1.279	-0.711	-0.568
	Kurtosis	5.678	7.305	4.546	7.522	5.3471	7.099	3.819	4.707
	Max	4.244	4.867	5.120	4.198	4.684	3.505	4.265	5.429
	Min	-8.733	-9.165	-6.190	-9.994	-5.867	-10.974	-4.813	-3.816

*Notes: Table 3 gives the descriptive statistics of ASJ jump test at 1-minute frequency. See notes of Table 2.

Table 4: Descriptive Statistics of BNS Jump Test

		Boeing	Exxcon	Johnson&Johnson	JP Morgan	Microsoft	Walmart	XLE	XLK
Panel A	Mean	0.878	0.771	1.098	0.854	0.867	0.859	0.599	1.415
	st.dev	1.511	1.327	1.665	1.434	1.320	1.401	1.327	1.693
	Skewness	1.030	0.618	1.151	0.951	0.602	0.801	0.780	1.315
	Kurtosis	4.903	3.749	4.851	4.518	3.500	4.068	4.169	6.395
	Max	7.791	6.631	8.661	7.352	5.811	6.662	6.947	10.731
	Min	-2.513	-2.341	-2.693	-2.364	-2.299	-2.505	-2.640	-2.695
Panel B	Mean	0.688	0.667	0.876	0.558	0.880	0.918	0.572	0.814
	st.dev	1.406	1.328	1.450	1.249	1.406	1.535	1.228	1.275
	Skewness	0.983	0.572	0.837	0.574	0.647	0.925	0.615	0.525
	Kurtosis	5.400	3.557	3.646	3.652	3.428	4.215	3.711	3.530
	Max	9.340	6.964	7.128	5.486	7.008	7.549	5.807	6.909
	Min	-2.485	-2.546	-2.006	-2.750	-2.417	-2.137	-2.225	-2.244

*Notes: Table 4 gives the descriptive statistics of BNS jump test at 5-minute frequency. See notes of Table 2.

Table 5: Descriptive Statistics of LM Jump Test

		Boeing	Exxcon	Johnson&Johnson	JP Morgan	Microsoft	Walmart	XLE	XLK
Panel A	Mean	0.595	4.436	-1.662	-0.676	3.987	0.181	2.202	2.347
	st.dev	42.865	50.142	101.707	47.879	66.636	52.002	37.687	72.997
	Skewness	0.911	2.909	0.675	0.850	1.267	-0.146	2.714	0.316
	Kurtosis	15.591	31.900	16.875	25.186	13.907	10.491	41.367	14.882
	Max	386.170	602.678	1008.160	525.995	585.419	314.852	472.919	509.147
	Min	-309.16	-185.941	-615.524	-354.424	-289.795	-343.618	-189.563	-570.747
Panel B	Mean	-0.110	0.295	0.076	0.137	0.097	0.636	-0.051	0.042
	st.dev	42.450	47.158	67.736	33.483	42.875	61.795	43.314	54.955
	Skewness	-0.465	-0.251	-0.635	1.958	1.035	0.138	-0.451	0.128
	Kurtosis	67.346	29.672	49.711	165.139	92.97	76.991	48.134	30.865
	Max	1047.900	1192.159	1455.508	1320.357	1726.992	1529.471	1288.848	1571.141
	Min	-1275.319	-764.818	-1789.208	-1399.492	-1064.302	-2124.609	-881.903	-881.418

*Notes: Table 5 gives the descriptive statistics of LM jump test at 5-minute frequency. See notes of Table 2.

Table 6: Percentage of days identified as having jumps - ASJ Jump Test

Name	Panel A: Jan 2006 - Jun 2009				Panel B: July 2009 - Dec 2012			
	1 min		5 min		1 min		5 min	
Significance level	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Boeing	15.79	27.95	3.29	10.79	10.89	23.38	2.27	9.98
Exxon	12.61	22.84	3.63	11.47	9.76	21.11	2.72	8.74
Johnson & Johnson	19.88	32.04	4.77	14.09	16.00	27.80	3.40	10.44
JP Morgan	12.38	24.20	3.52	11.36	10.32	18.84	3.06	8.05
Microsoft	17.04	31.70	3.52	11.36	14.18	25.42	4.08	11.12
Walmart	15.11	27.61	3.52	11.59	13.16	24.85	3.74	9.53
XLE	10.22	17.84	3.06	9.43	7.60	13.96	2.27	7.94
XLK	20.45	34.31	3.63	10.68	15.20	26.44	2.83	9.19

*Notes: Table 6 gives the percentage of days identified as having jumps by the ASJ test at both 5 minute and 1 minute sampling frequencies. Jumps are tested at $\alpha = 0.05$ and $\alpha = 0.1$ significance level. Percentage of days is calculated using the equation 87 in section 6.2. Panel A covers the financial crisis period from Jan 2006 - June 2009 and Panel B covers the post fiancial crisis period from July 2009 - Dec 2012.

Table 7: Percentage of Days Identified as having Jumps - BNS Jump Test

Name	Panel A: Jan 2006 - Jun 2009		Panel B: July 2009 - Dec 2012	
Signifiance Level	0.05	0.10	0.05	0.10
Boeing	20.68	27.73	16.69	23.04
Exxcon	17.95	26.25	15.78	23.27
Johnson&Johnson	24.09	30.23	20.43	25.88
JP Morgan	19.89	25.23	12.83	19.64
Microsoft	18.86	25.23	21.45	27.70
Walmart	19.77	27.38	20.66	26.67
XLE	15.11	21.36	12.71	19.41
XLK	30.00	38.18	17.93	25.54

*Notes: Table 7 shows percentage of days identified as having jumps by BNS test at 5 minute sampling frequency. See notes of Table 6.

Table 8: Percentage of days identified as having jumps and Jump Proportion - LM Jump Test

Name	Panel A: Jan 2006 - Jun 2009				Panel B: July 2009 - Dec 2012			
	Jump proportion		% of Jump days		Jump proportion		% of Jump days	
Signifiance Level	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Boeing	1.06	1.07	82.39	83.41	1.14	1.15	88.99	89.56
Exxcon	1.10	1.11	85.45	86.36	1.17	1.17	90.92	91.49
Johnson&Johnson	1.15	1.15	89.55	89.89	1.21	1.21	94.32	94.55
JP Morgan	1.03	1.04	80.45	80.91	1.11	1.12	86.83	87.29
Microsoft	1.15	1.15	89.66	89.66	1.16	1.17	90.47	91.15
Walmart	1.11	1.12	86.36	87.05	1.19	1.19	92.51	92.74
XLE	1.03	1.05	80.45	81.93	1.17	1.18	91.60	92.30
XLK	1.06	1.07	82.84	83.41	1.14	1.14	88.88	89.22

*Notes: Table 8 shows percentage of days identified as having jumps and jump proportion as per the LM test at 5 minute sampling frequency. Percentage of days is calculated using the equation 88 in section 6.2. See notes of Table 6.

Table 9: Average Ratio of Jump Variation to Total Variation - ASJ Jump Test

Name	Panel A: Jan 2006 - June 2009				Panel B: July 2009 - Dec 2012			
	1 min		5 min		1 min		5 min	
Significance level	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Boeing	3.74	6.14	0.70	2.38	2.14	4.08	0.32	1.68
Exxon	2.86	4.28	0.44	1.63	1.55	3.02	0.26	1.03
Johnson & Johnson	5.21	8.09	0.99	2.87	3.49	5.96	0.64	2.17
JP Morgan	2.97	5.21	0.55	2.16	1.56	2.90	0.42	1.14
Microsoft	4.59	8.05	0.50	2.07	3.30	5.47	0.96	2.20
Walmart	3.50	6.30	0.53	2.18	2.93	5.33	0.86	1.81
XLE	2.08	3.35	0.34	1.39	1.44	2.40	0.24	1.05
XLK	10.78	16.96	1.27	3.40	4.69	7.88	0.54	1.74

Notes*: Table 9 gives the average ratio of jump varaition to total variation as per the ASJ test using both 5 minute and 1 minute sampling frequencies. Jump ratio is calculated using the equation 90 in section 6.2. Panel A covers the financial crisis period from Jan 2006 - June 2009 and Panel B covers the post fiancial crisis period from July 2009 - Dec 2012.

Table 10: Average Ratio of Jump Variation to Total Variation - BNS Jump Test

Name	Panel A: Jan 2006 - Jun 2009		Panel B: July 2009 - Dec 2012	
Signifiance Level	0.05	0.10	0.05	0.10
Boeing	37.87	34.36	6.10	8.58
Exxcon	32.00	28.44	6.93	7.67
Johnson&Johnson	42.46	39.93	9.30	11.09
JP Morgan	36.81	33.75	5.99	8.13
Microsoft	36.44	33.26	8.92	9.76
Walmart	36.60	33.43	7.91	8.87
XLE	33.00	30.08	4.37	5.64
XLK	44.63	41.78	12.75	14.70

*Notes: Table 10 shows average ratio of jump variation to total variation as per the BNS test using 5 minute frequency. See notes of Table 9.

Table 11: Average Ratio of Jump Variation to Total Variation - LM Jump Test

Name	Panel A: Jan 2006 - Jun 2009		Panel B: July 2009 - Dec 2012	
Signifiance Level	0.05	0.10	0.05	0.10
Boeing	17.72	17.79	14.59	14.87
Exxcon	14.65	14.58	12.59	12.62
Johnson&Johnson	21.53	21.53	19.34	19.4
JP Morgan	17.41	17.36	13.92	13.93
Microsoft	17.57	17.57	15.59	15.60
Walmart	17.68	17.56	15.22	15.23
XLE	13.58	13.56	10.98	11.17
XLK	25.45	25.35	21.33	21.41

*Notes: Table 11 shows the average ratio of jump variation to total variation as per the LM test using 5 minute frequency. See notes of Table 9.

Table 12: Percentage of Days Identified as having Co-jumps

	Panel A: Jan 2006 - Jun 2009				Panel B: July 2009 - Dec 2012			
	JT Test		LM&BNS Test		JT Test		LM&BNS Test	
Significance Level	0.05	0.10	0.05	0.10	0.05	0.10	0.05	0.10
Exxon&JPMorgan	1.364	1.818	3.864	7.159	1.703	1.816	2.838	5.335
Exxon & Microsoft	0.795	0.909	3.295	6.818	1.022	1.476	3.519	5.675
Exxon & XLE	1.136	1.59	3.977	7.386	1.93	2.497	3.746	8.059
JPMorgan & Microsoft	1.136	1.363	3.409	5.682	1.135	1.249	2.951	5.335
Microsoft & XLK	2.500	2.955	6.136	9.545	1.022	1.362	4.767	7.491
XLE & XLK	2.045	2.386	4.773	8.295	0.454	0.568	3.065	6.129

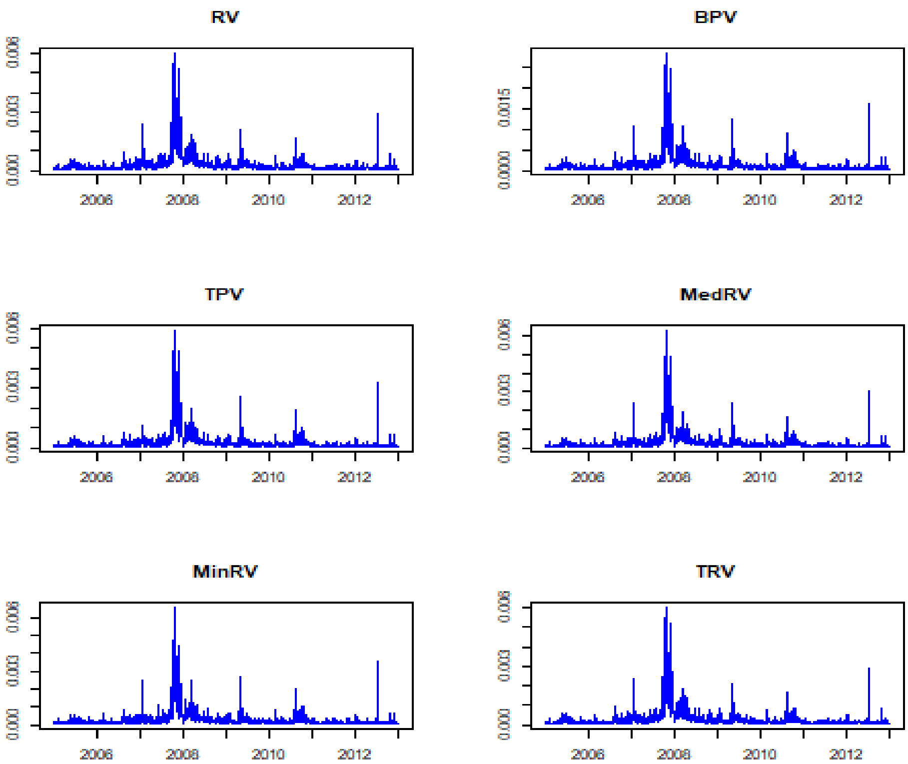
*Notes: Table 12 shows the percentage of days identified as having co-jumps. Co-jumps are detected at $\alpha = 0.05$ and $\alpha = 0.1$ significance level. Both JT co-jump test and co-exceedance rule between BNS test and LM test are used to test co-jumps. Panel A covers the financial crisis period from Jan 2006 - June 2009 and Panel B covers the post fiancial crisis period from July 2009 - Dec 2012. The test statistics are calculated at 5-minute frequency.

Table 13: Percentage of Days Identified as having Co-jumps

	Panel		Panel B	
Significance Level	0.05	0.1	0.05	0.1
BLT Test	0.227	0.341	0.114	0.454

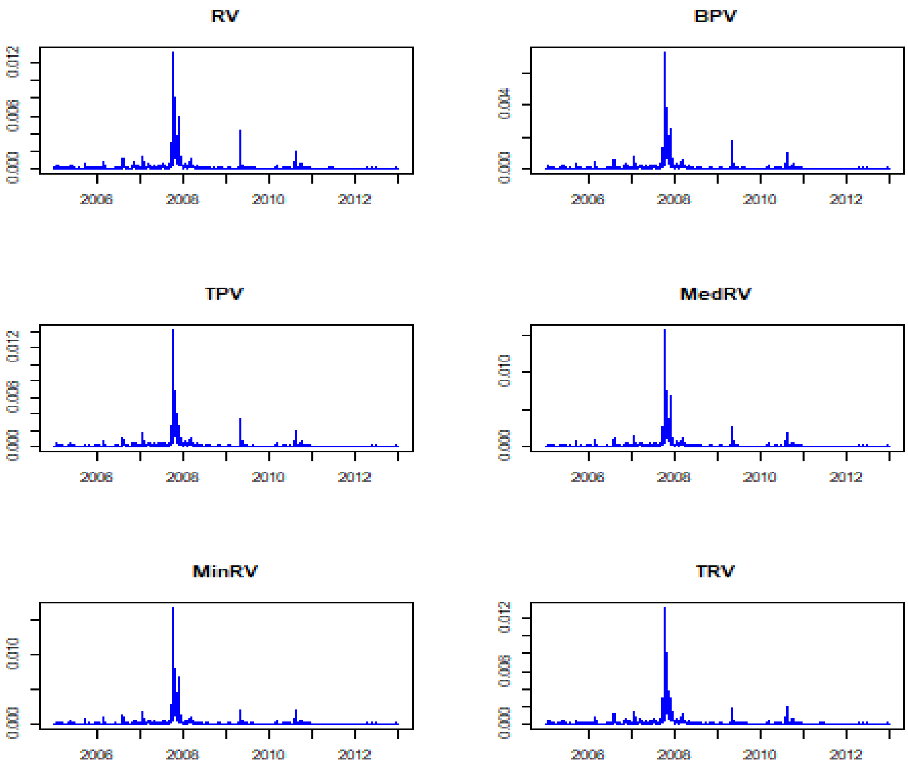
*Notes: See notes to Table 12. Table 13 shows the percentage of days identified as having co-jumps from BLT co-jump tests. Co-jumps are detected among six stocks: Boeing, Exxon, JohnsonJohnson, JPMorgan, Microsoft and Walmart.

Figure 1: Integrated Volatility Measures - Boeing



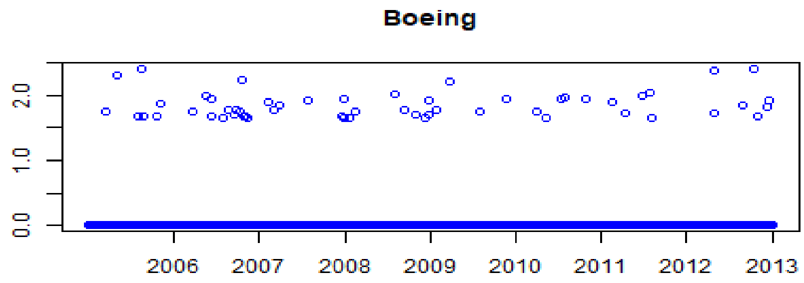
Notes*: Figure 1 displays volatility of Boeing across the sample period Jan 2006 - Dec 2013 using 5-minute sampling frequency with respect to six different integrated volatility measures which include RV, BPV, TPV, MedRV, MinRV and TRV.

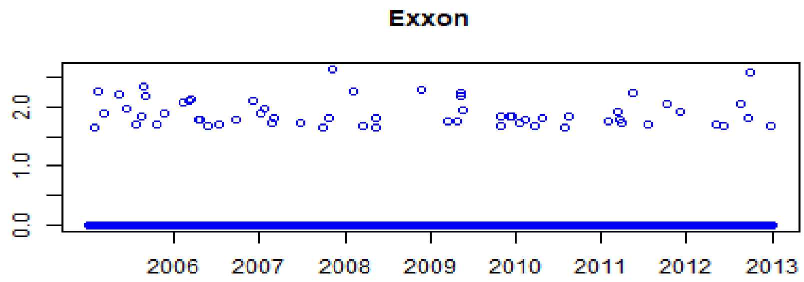
Figure 2: Integrated Volatility Measures - Exxon



Notes*: Figure 2 displays volatility movement of Exxon. See notes of Figure 1.

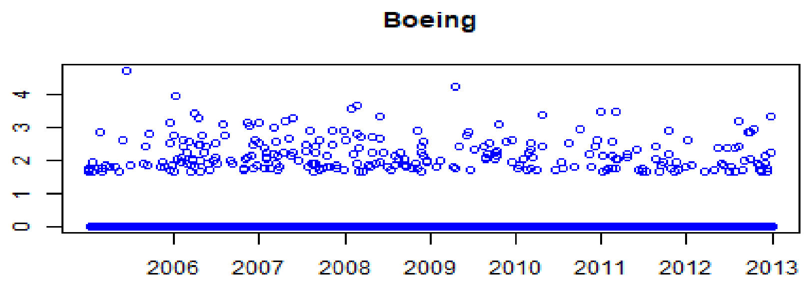
Figure 3: ASJ Jump Test Statistics of Days Identified as having Jumps: 5-minute sampling frequency

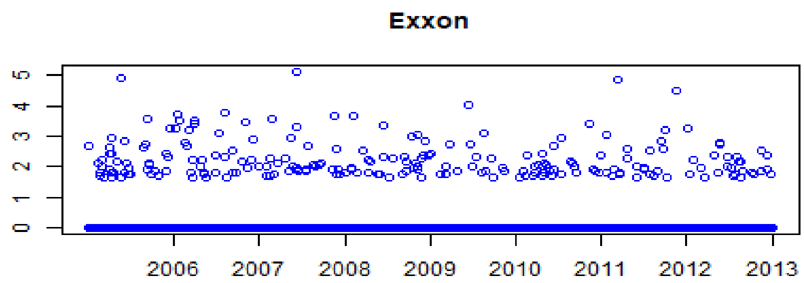




*Notes: Figure 3 displays the scatter plot for ASJ test statistics for days identified as having jumps using 5-minute sampling frequency. We consider the following stocks and ETFs: Boeing & Exxon for the sample period Jan 2006 to Dec 2013.

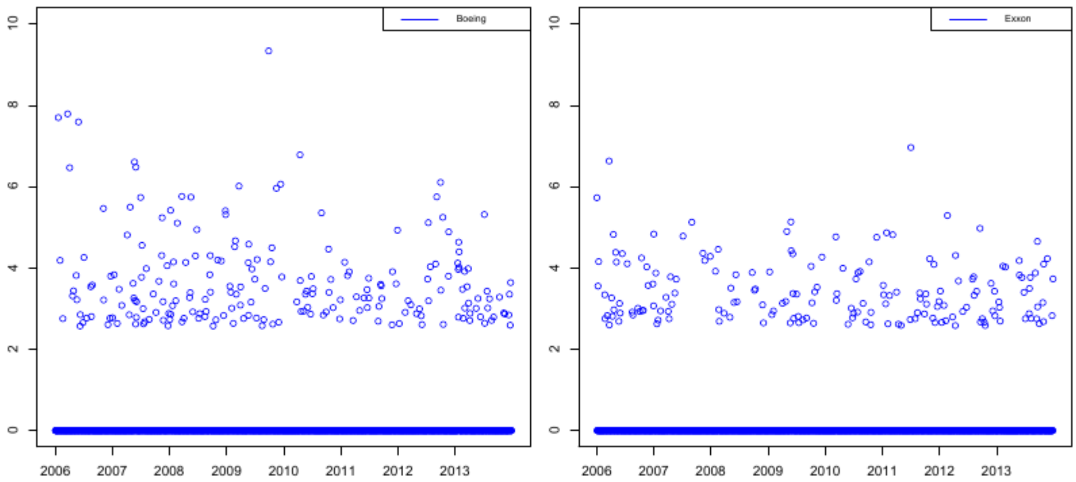
Figure 4: ASJ Jump Test Statistics of Days Identified as having Jumps: 1-minute sampling frequency





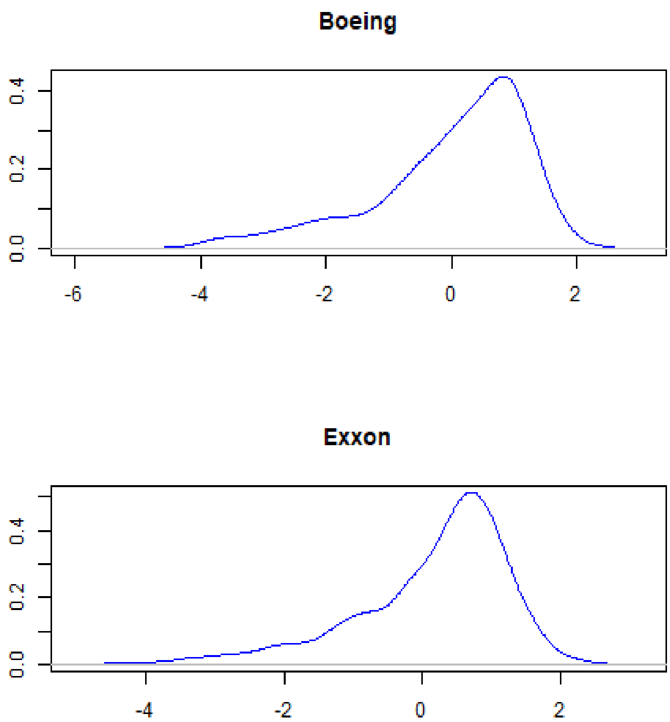
*Notes: Figure 4 displays the scatter plot for ASJ test statistics for days identified as having jumps using 1-minute sampling frequency. See notes of figure 3.

Figure 5: BNS Jump Test Statistics of Days Identified as having Jumps



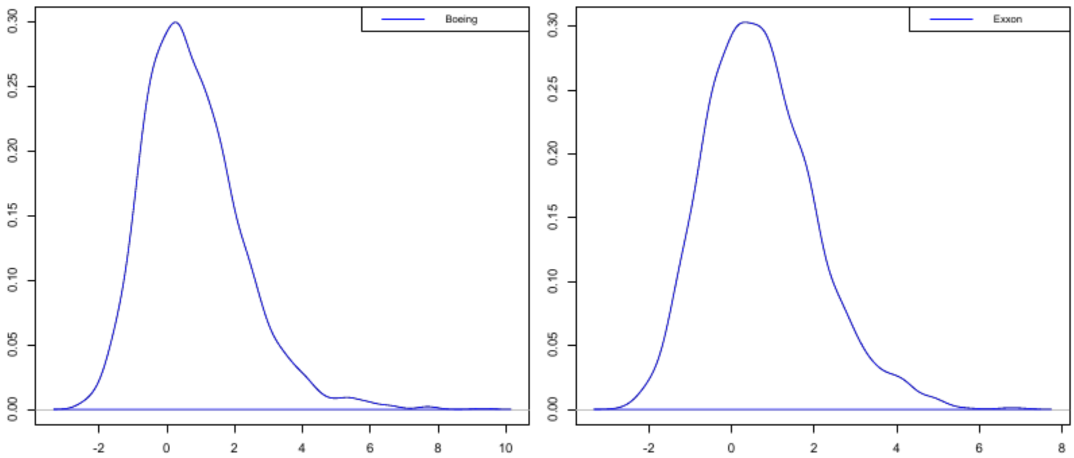
*Notes: Figure 5 displays the scatter plot for BNS test statistics for days identified as having jumps using 5-minute sampling frequency. See notes of figure 3.

Figure 6: Kernel Density Plots for ASJ Test Statistics



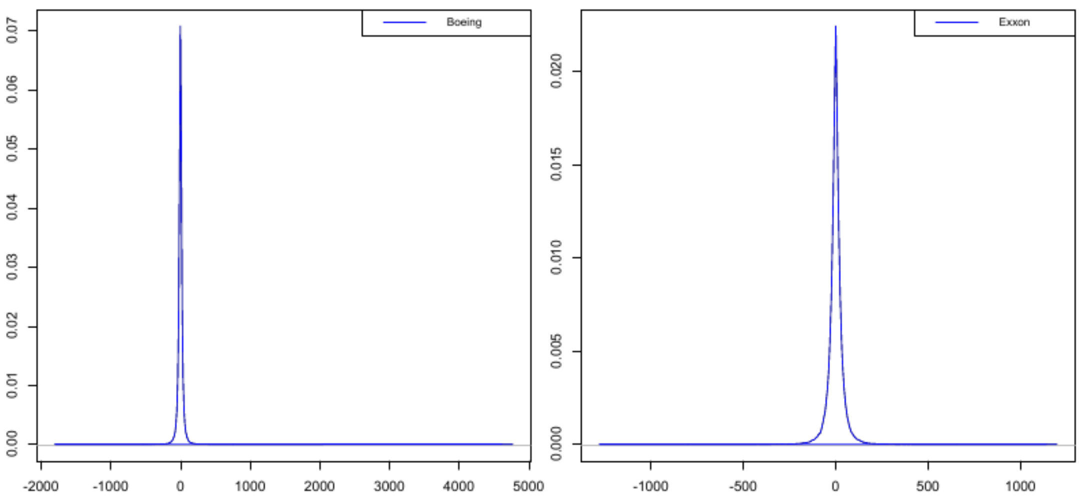
Notes*: Figure 6 displays the kernel density plot of ASJ jump test statistics using 5-minute sampling frequency. See notes of figure 3.

Figure 7: Kernel Density Plot of BNS Jump Test Statistics



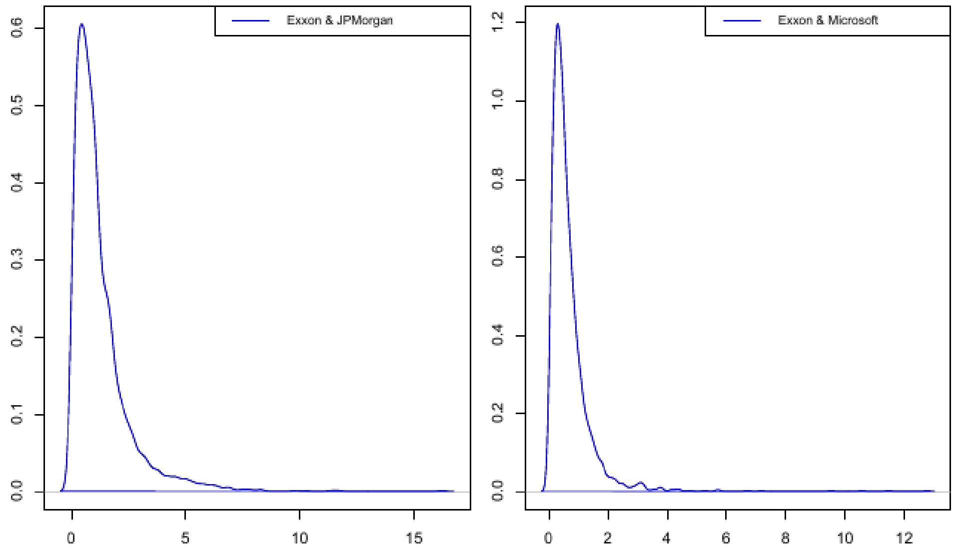
Notes*: Figure 7 displays the kernel density plot of BNS jump test statistics using 5-minute sampling frequency. See notes of figure 3.

Figure 8: Kernel Density Plot of LM Jump Test Statistics



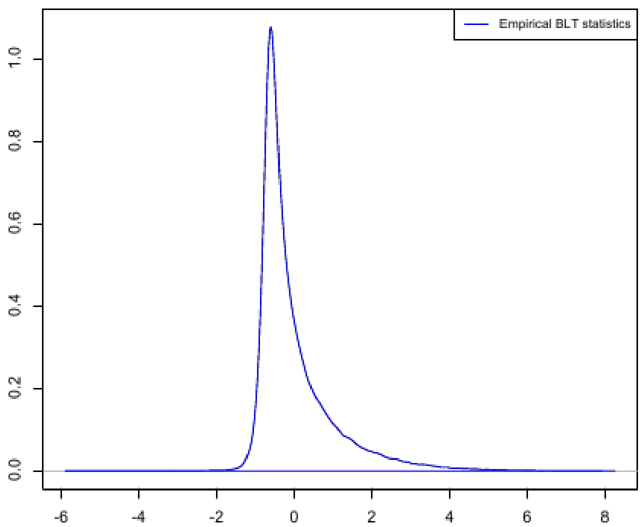
Notes*: Figure 8 displays the kernel density plot of LM jump test statistics using 5-minute sampling frequency. See notes of figure 3.

Figure 9: Kernel Density Plot of JT Co-jump Test Statistics



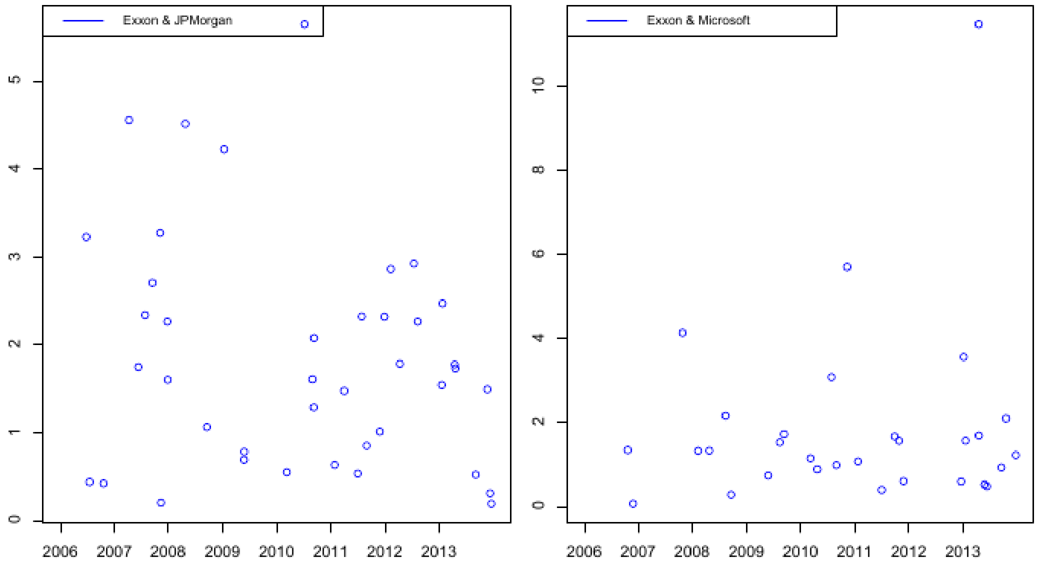
*Notes: Figure 9 displays the kenel density plot of JT co-jump test using 5-minute sampling frequency. The co-jumps are tested for the pairs Exxon & JPMorgan and Exxon & Microsft for the sample period Jan 2006-Dec 2013.

Figure 10: Kernel Density Plot of empirical observed BLT Statistics



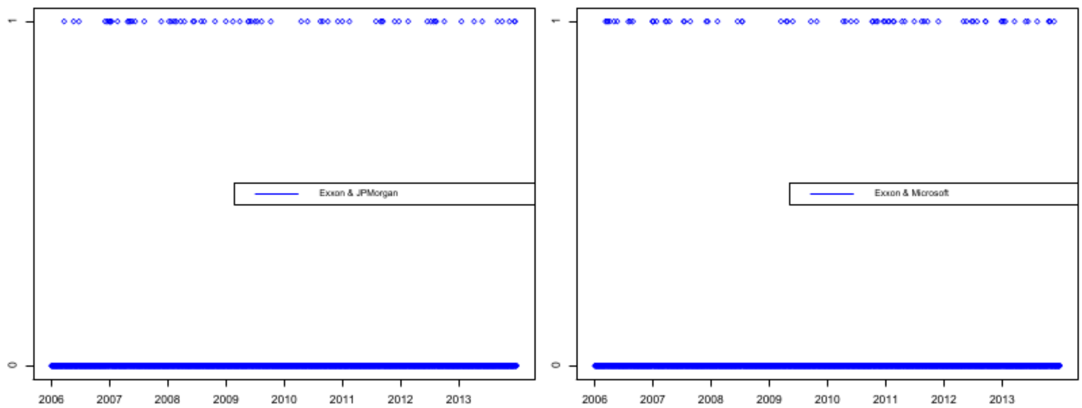
*Notes: Figure 10 displays the kernel density plot of the empirical observed BLT co-jump test statistics using 5-minute sampling frequency for the sample period Jan 2006-Dec 2013.

Figure 11: JT Co-jump Test Statistics of Days Identified as Having Co-jumps



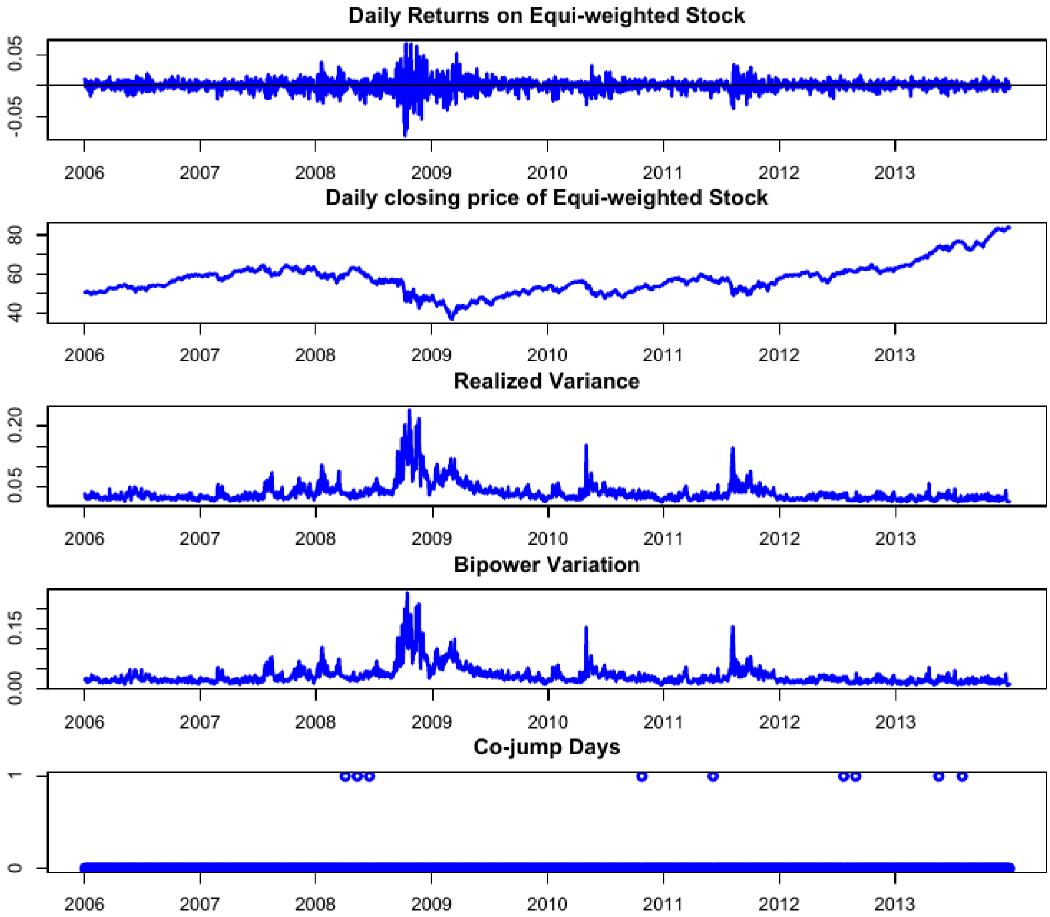
*Notes: Figure 11 displays the co-jump days test statistics of JT test for the sample period Jan 2006 to Dec 2013 using sampling frequency of 5-minute.

Figure 12: LM & BNS Test Statistics for Days Having Co-jumps



*Notes: Figure 12 displays the co-jump days identified from co-exceedance rule between LM jumps test and BNS jump test for the pairs Exxon & JPMorgan and Exxon & Microsoft. 5-minute sampling frequency is considered for sample period Jan 2006-Dec 2013.

Figure 13: Daily Return, Daily Closing Price, Realized Variance, Bipower Variation and Co-jump Days for Equi-weighted Stock Index



*Notes: Figure 13 displays the daily return, daily closing price, realized variance, bipower variation and co-jump days for equi-weighted stock index. The co-jump days in the last panel are detected through BLT co-jump tests at $\alpha = 0.1$ significance level from Jan 2006 to Dec 2013. The equi-weighted stock index is composed of six stocks (Boeing, Exxon, JohnsonJohnson, JPMorgan, Microsoft and Walmart) with equal weights.

8 Appendix: R Code

Please find below, R (statistical software) codes for the tests and measures which have been used in the empirical section of this paper. This includes the volatility measures RV, BPV, TPV, MinRV, MedRV, TRV; ASJ, BNS jump tests; BLT and JT co-jump tests. The format of input data can be given as follows: rows signify the trading days and the columns signify the intra-day intervals.

```
#_____________________________________________________________________________________________________________________________________#
                                                ### Integrated Volatility Measures ###
#_____________________________________________________________________________________________________________________________________#

  ##--------------------##
  ## Data preparation 1 ##
  ##--------------------##
  data <- read.csv(file="asset1_1min_data.csv", header=FALSE, sep=",")
  mat <- data.matrix(data)
  fin_data <- t(mat)
  days <- nrow(data)
  freq <- 79
  data_5 <- matrix(0, freq, days)
  for (j in 1:days){for(i in 1:freq){data_5[i,j] <- fin_data[(i-1)*5+1,j]}}

  ##--------------------------##
  ## Realized Volatility - RV ##
  ##--------------------------##
  dif_data <- diff(data_5)
  RV <- colSums((dif_data)^2)


  ##-------------------------##
  ## Bipower Variation - BPV ##
  ##-------------------------##
  dif1 <- dif_data[1:freq-2,1:days]
  dif2 <- dif_data[2:freq-1,1:days]
  BPV <-  (sqrt(2)/sqrt(pi))*colSums(abs(dif1)*abs(dif2))


  ##--------------------------##
  ## Tripower variation - TPV ##
  ##--------------------------##
  dif11 <- dif_data[1:freq-3, 1:days]
  dif22 <- dif_data[2:freq-2, 1:days]
  dif33 <- dif_data[3:freq-1, 1:days]
  cons <- (((2^(1/3))*gamma(5/6))/gamma(1/2))^(-3)
  TPV <- cons*colSums((abs(dif11)^(2/3))*(abs(dif22)^(2/3))*(abs(dif33)^(2/3)))


  ##--------##
  ## MinRV  ##
  ##--------##
  minvec <- matrix(0, freq-2, days)
  for (j in 1:days){for (i in 1:freq-2){minvec[i,j] <- (min( abs(dif1[i,j]), abs(dif2[i,j]) ))^2}}
  MinRV <- (pi/(pi-2))*(freq/freq-1)*colSums(minvec)


  ##-------##
  ## MedRV ##
  ##-------##
  medvec <- matrix(0, freq-3, days)
  for (j in 1:days){for (i in 1:freq-3){medvec[i,j] <- (median(c(abs(dif11[i,j]),abs(dif22[i,j]),abs(dif33[i,j]))))^2}}
  MedRV <- (pi/(6 - 4*sqrt(3) + pi))*(freq/freq-2)*colSums(medvec)


  ##---------------------------------------##
  ## Truncated Realized Volatility - TRV  ##
  ##---------------------------------------##
  delta <- 1/freq
  omega <- 0.47
  alpha <- matrix(0, freq-1, days)
  for (j in 1: days){
      for (i in 1: freq-1){if (abs(dif_data[i,j]) <= sqrt(delta)){alpha[i,j] <- abs((dif_data[i,j]))^2} else {alpha[i,j] <- 0}}}
  alph_fin <- 5*sqrt(colSums(alpha))
  trun <- matrix(0, freq-1, days)
  for (j in 1: days){for (i in 1: freq-1){
      if (abs(dif_data[i,j]) <= alph_fin[j]*delta^omega){trun[i,j] <- abs((dif_data[i,j]))^2} else {trun[i,j] <- 0}}}
  TRV <- colSums(trun)
```

```
##____________________________________________________________________________________________________________________________________##
                                                            ### Jump Tests ###
##____________________________________________________________________________________________________________________________________##


    ##---------------##
    ## ASJ Jump Test ##
    ##---------------##
    BPD <- colSums((abs(dif_data))^4)
    kfreq = (freq+1)/2
    data_10 <- matrix(0, kfreq, days)
    for (j in 1: days){for (i in 1:kfreq){data_10[i,j] <- data_5[(i-1)*2+1,j]}}
    BPK <- colSums((abs(diff(data_10)))^4)
    SPK <- BPK/BPD
    trun_4 <- matrix(0,freq-1,days)
    for (j in 1: days){for (i in 1:freq-1){
        if (abs(dif_data[i,j]) <= alph_fin[j]*delta^omega){trun_4[i,j] <- abs((dif_data[i,j]))^4}else {trun_4[i,j] <- 0}}}
    mp <- pi^(-0.5)*4*gamma(5/2)
    AP <- (delta^(-1)/mp)*colSums(trun_4)
    trun_8 <- matrix(0, freq-1, days)
    for (j in 1: days){for (i in 1: freq-1){
        if (abs(dif_data[i,j]) <= alph_fin[j]*delta^omega){trun_8[i,j] <- abs((dif_data[i,j]))^8} else {trun_8[i,j] <- 0}}}
    mp_8 <- pi^(-0.5)*16*gamma(9/2)
    AP_8 <- (delta^(-3)/mp_8)*colSums(trun_8)
    Var <- (delta* AP_8*160)/(3*AP^2)
    ASJ <- (2 - SPK)/sqrt(Var)


    ##---------------------##
    ## Data preparation 2  ##
    ##---------------------##
    data_asset1 <- read.table("asset1_1min_data.csv",header=FALSE, sep=",");
    data_asset2 <- read.table("asset2_1min_data.csv",header=FALSE, sep=",");
    data_asset3 <- read.table("asset3_1min_data.csv",header=FALSE, sep=",");
    col_diff_asset1 <- apply(data_asset1, 1, diff);
    col_diff_asset2 <- apply(data_asset2, 1, diff);
    col_diff_asset3 <- apply(data_asset3, 1, diff);
    vec_asset1=as.vector(col_diff_asset1);
    vec_asset2=as.vector(col_diff_asset2);
    vec_asset3=as.vector(col_diff_asset3);
    vec_asset1_5min=colSums(matrix(vec_asset1, nrow=5));
    vec_asset2_5min=colSums(matrix(vec_asset2, nrow=5));
    vec_asset3_5min=colSums(matrix(vec_asset3, nrow=5));
    matrix_5min_all=cbind(vec_asset1_5min, vec_asset2_5min, vec_asset3_5min);


    ##-------------------##
    ## BNS Jump Test     ##
    ##-------------------##
    simu_n <- days;
    N=ncol(matrix_5min_all);
    freq=nrow(matrix_5min_all)/simu_n;
    RVdaily=matrix(0,nrow=simu_n,ncol=N);
    matrix_5min_all_sqr=matrix_5min_all*matrix_5min_all;
    for (j in 1:N){for (i in 1:simu_n){RVdaily[i,j]=sum(matrix_5min_all_sqr[((i-1)*freq+1):(i*freq),j])}}
    BV=matrix(0,nrow=simu_n,ncol=N);
    mu1=sqrt(2/pi);
    for (k in 1:N){ for (i in 1:simu_n){for (j in 2:freq){BV[i,k]=BV[i,k]+abs(matrix_5min_all[
        ((i-1)*freq+j),k])*abs(matrix_5min_all[((i-1)*freq+j-1),k])}BV[i,k]=(mu1)^(-2)*(freq/(freq-1))*BV[i,k]}}
    QP=matrix(0,nrow=simu_n,ncol=N);
    for (k in 1:N){for (i in 1:simu_n){for (j in 4: freq){
        QP[i,k]=QP[i,k]+(abs(matrix_5min_all[((i-1)*freq+j),k]))*(abs(matrix_5min_all[((i-1)*freq+j-1),k]))*
        (abs(matrix_5min_all[((i-1)*freq+j-2),k]))*(abs(matrix_5min_all[((i-1)*freq+j-3),k]));}QP[i,k]=freq*QP[i,k]}}
    Vqq=2;
    Vbb=(pi/2)^2+pi-3;
    Zqp=matrix(0,nrow=simu_n,ncol=N);
    for (k in 1:N){for (i in 1:simu_n){
         Zqp[i,k]=-(((BV[i,k]/RVdaily[i,k])-1)*(freq^0.5))/(sqrt((Vbb-Vqq)*max(1,QP[i,k]/(BV[i,k]^2*mu1^4)))) }}
    detect_BNS=matrix(0,nrow=simu_n,ncol=N);
    for (k in 1:N){for (j in 1:simu_n){detect_BNS[j,k]=ifelse(abs(Zqp[j,k])>1.96,1,0)}}
```

```
##____________________________________________________________________________________________________________________________________##
                                                        ### Co-Jump Tests ###
##____________________________________________________________________________________________________________________________________##


    ##------------------##
    ## BLT Co-Jump Test ##
    ##------------------##
    library(MASS)
    Input1=matrix_5min_all
    cov_cof=cov(matrix_5min_all)
    N=ncol(matrix_5min_all)
    simu_n<-days;
    mu=rep(0,N);
    freq=nrow(matrix_5min_all)/simu_n;
    set.seed(123)
    bpath=mvrnorm(n = simu_n*freq, mu, cov_cof,  empirical = FALSE)
    dt=1/freq
    sigma<-10
    bdw=sqrt(dt)*bpath*sigma
    bz=matrix(0,nrow=simu_n*freq,ncol=N-1)
    for (i in 1:(N-1)){for (l in (i+1):N){ bz[,i]=bz[,i]+bdw[,i]*bdw[,l]}}
    bmcp=rowSums(bz)
    bmcptj=(2/((N-1)*N))*bmcp
    bmcp_bar=matrix(0,nrow=simu_n,ncol=1)
    for (i in 1:simu_n){bmcp_bar[i]=mean(bmcptj[((i-1)*freq+1):(i*freq)])}
    mcp_bar1=matrix(0,nrow=simu_n*freq,1)
    for (i in 1:simu_n){mcp_bar1[((i-1)*freq+1):(i*freq),]=bmcp_bar[i]}
    mcp_sqr=(bmcptj-mcp_bar1)*(bmcptj-mcp_bar1)/(freq-1)
    sd_mcpt=matrix(0,nrow=simu_n,ncol=1)
    for (i in 1:simu_n){sd_mcpt[i]=sqrt(sum(mcp_sqr[((i-1)*freq+1):(i*freq)]))}
    mcp_sqr1=matrix(0,nrow=freq*simu_n,ncol=1)
    for (i in 1:simu_n){mcp_sqr1[((i-1)*freq+1):(i*freq)]=sd_mcpt[i]}
    zmcptj=(bmcptj-mcp_bar1)/mcp_sqr1
    sortnul=sort(zmcptj, decreasing = TRUE)
    rightside=matrix(0,nrow=2,1)
    rightside[1]=sortnul[round(freq*simu_n*0.05, digits = 0)]
    rightside[2]=sortnul[round(freq*simu_n*0.025, digits = 0)]
    sortnul1=sort(zmcptj, decreasing = FALSE)
    leftside=matrix(0,nrow=2,1)
    leftside[1]=sortnul1[round(freq*simu_n*0.05, digits = 0)]
    leftside[2]=sortnul1[round(freq*simu_n*0.025, digits = 0)]
    z_obs=matrix(0,nrow=simu_n*freq,ncol=N-1)
    for (i in 1:(N-1)){for (l in (i+1):N){ z_obs[,i]=bz[,i]+Input1[,i]*Input1[,l]}}
    mcp_obs=rowSums(z_obs)
    mcptj_obs=(2/((N-1)*N))*mcp_obs
    mcp_obsbar=matrix(0,nrow=simu_n,ncol=1)
    for (i in 1:simu_n){mcp_obsbar[i]=mean(mcptj_obs[((i-1)*freq+1):(i*freq)])}
    mcp_obsbar1=matrix(0,nrow=freq*simu_n,ncol=1)
    for (i in 1:simu_n){mcp_obsbar1[((i-1)*freq+1):(i*freq),]=mcp_obsbar[i]}
    mcp_obssqr=(mcptj_obs-mcp_obsbar1)*(mcptj_obs-mcp_obsbar1)/(freq-1)
    Sd_obsmcpt=matrix(0,nrow=simu_n,ncol=1)
    for (i in 1:simu_n){Sd_obsmcpt[i]=sqrt(sum(mcp_obssqr[((i-1)*freq+1):(i*freq)]))}
    mcp_obssqr1=matrix(0,nrow=freq*simu_n,ncol=1)
    for (i in 1:simu_n){mcp_obssqr1[((i-1)*freq+1):(i*freq)]=Sd_obsmcpt[i]}
    zmcptj_obs=(mcptj_obs-mcp_obsbar1)/mcp_obssqr1
    detectr=matrix(0,nrow=freq*simu_n,ncol=4)
    for (i in 1:4){ for (j in 1:freq*simu_n){detectr[j,i]=ifelse(zmcptj_obs[j]>rightside[i], 1, 0)}}
    detectl=matrix(0,nrow=freq*simu_n,ncol=4)
    for (i in 1:4){ for (j in 1:freq*simu_n){detectl[j,i]=ifelse(zmcptj_obs[j]<leftside[i], 1, 0)}}


    ##------------------##
    ## JT Co-Jump Test  ##
    ##------------------##
    library(MASS)
    combos=combn(1:ncol(matrix_5min_all), 2, FUN = NULL, simplify = TRUE);
    n1=ncol(combos);
    V=matrix(0,nrow=simu_n,ncol=n1);
    Input1=matrix_5min_all;
    freq=nrow(matrix_5min_all)/simu_n;
    for (m in 1:n1){for (j in 1:simu_n){
        for (i in 1: freq){V[j,m]=V[j,m]+(Input1[((j-1)*freq)+i,combos[1,m]]*Input1[(j-1)*freq+i,combos[2,m]])^2}}};
    V_g1=matrix(0,nrow=simu_n,ncol=n1);
    for (m in 1:n1){for (j in 1:simu_n){for (i in 1: freq){V_g1[j,m]=V_g1[j,m]+(Input1[(j-1)*freq+i,combos[1,m]])^4}}};
    V_g2=matrix(0,nrow=simu_n,ncol=n1);
    for (m in 1:n1){for (j in 1:simu_n){for (i in 1: freq){V_g2[j,m]=V_g1[j,m]+(Input1[(j-1)*freq+i,combos[2,m]])^4}}};
```

```
        phi_d=V/sqrt(V_g1*V_g2);
        delta=1/freq;
        A_hatn=matrix(0,nrow=simu_n,ncol=n1);
        for (m in 1:n1){for (j in 1:simu_n){for (i in ((j-1)*freq+1):((j-1)*freq+freq-3)){A_hatn[j,m]=A_hatn[j,m]+abs(Input1[i,combos[1,m]]
                            *Input1[i+1,combos[1,m]]*Input1[i+2,combos[2,m]]*Input1[i+3,combos[2,m]])
                            +(1/8)*abs((Input1[i,combos[1,m]]+Input1[i,combos[2,m]])*(Input1[i+1,combos[1,m]]
                            +Input1[i+1,combos[2,m]])*(Input1[i+2,combos[1,m]]+Input1[i+2,combos[2,m]])
                            *(Input1[i+3,combos[1,m]]+Input1[i+3,combos[2,m]]));
                            +(1/8)*abs((Input1[i,combos[1,m]]-Input1[i,combos[2,m]])
                            *(Input1[i+1,combos[1,m]]-Input1[i+1,combos[2,m]])
                            *(Input1[i+2,combos[1,m]]-Input1[i+2,combos[2,m]])
                            *(Input1[i+3,combos[1,m]]-Input1[i+3,combos[2,m]]));
                            -(1/4)*abs((Input1[i,combos[1,m]]+Input1[i,combos[2,m]])
                            *(Input1[i+1,combos[1,m]]+Input1[i+1,combos[2,m]])*(Input1[i+2,combos[1,m]]
                            -Input1[i+2,combos[2,m]])*(Input1[i+3,combos[1,m]]-Input1[i+3,combos[2,m]]));
        }}};
        A_hatn=A_hatn*(pi)^2/(4*delta);
        BV=matrix(0,nrow=simu_n,ncol=ncol(matrix_5min_all));
        mu1=sqrt(2/pi);
        for (k in 1:5){for (i in 1:simu_n){ for (j in 2:freq){BV[i,k]=BV[i,k]+abs(Input1[((i-1)*freq+j),k])*abs(Input1[((i-1)*freq+j-1),k])}}}
        BV=(mu1)^(-2)*(freq/(freq-1))*BV
        trunc=1*sqrt(BV)*delta^(0.49);
        mu=c(0,0);
        sigma=matrix(c(1, 0, 0,1), nrow=2, ncol=2);
        Nn=20
        kn=1/sqrt(delta)
        Z_alphad_10=matrix(0,nrow=simu_n,ncol=n1);
        Z_alphad_5=matrix(0,nrow=simu_n,ncol=n1);
        for (m in 1:n1){
                for (kk in 1:simu_n){
                    Dhatn=matrix(0,nrow=Nn,ncol=1)
                    chatp1=matrix(0,nrow=freq-2*round(kn)-1,ncol=1)
                    for (i in ((kk-1)*freq+1+round(kn)):((kk-1)*freq+freq-round(kn)-1)){
                        for (j in (i+1):(i+round(kn)+1)){if (abs(Input1[j,combos[1,m]])<=trunc[kk,combos[1,m]]){
                        chatp1[i-round(kn)-(kk-1)*freq,]=chatp1[i-round(kn)-(kk-1)*freq,]
                        +Input1[j,combos[1,m]]*Input1[j,combos[1,m]]}
                        else{chatp1[(i-round(kn)-(kk-1)*freq),]=chatp1[i-round(kn)-(kk-1)*freq,]}}
                        chatp1[i-round(kn)-(kk-1)*freq,]=(1/(kn*delta))*chatp1[i-round(kn)-(kk-1)*freq,]}
                        chatm1=matrix(0,nrow=freq-2*round(kn)-1,ncol=1)
                    for (i in ((kk-1)*freq+1+round(kn)):((kk-1)*freq+freq-round(kn)-1)){
                        for (j in (i-round(kn)):(i-1)){if (abs(Input1[j,combos[1,m]])<=trunc[kk,combos[1,m]]){
                        chatm1[i-round(kn)-(kk-1)*freq,]=chatm1[i-round(kn)-(kk-1)*freq,]
                        +Input1[j,combos[1,m]]*Input1[j,combos[1,m]]}
                        else{chatm1[(i-round(kn)-(kk-1)*freq),]=chatm1[i-round(kn)-(kk-1)*freq,]}}
                        chatm1[i-round(kn)-(kk-1)*freq,]=(1/(kn*delta))*chatm1[i-round(kn)-(kk-1)*freq]}
                        chatp2=matrix(0,nrow=freq-2*round(kn)-1,ncol=1)
                    for (i in ((kk-1)*freq+1+round(kn)):((kk-1)*freq+freq-round(kn)-1)){
                         for (j in (i+1):(i+round(kn)+1)){if (abs(Input1[j,combos[2,m]])<=trunc[kk,combos[2,m]]){
                        chatp2[i-round(kn)-(kk-1)*freq,]=chatp1[i-round(kn)-(kk-1)*freq,]
                        +Input1[j,combos[2,m]]*Input1[j,combos[2,m]]}
                        else{chatp2[(i-round(kn)-(kk-1)*freq),]=chatp2[i-round(kn)-(kk-1)*freq,]}}
                        chatp2[i-round(kn)-(kk-1)*freq,]=(1/(kn*delta))*chatp2[i-round(kn)-(kk-1)*freq,]}
                        chatm2=matrix(0,nrow=freq-2*round(kn)-1,ncol=1)
                    for (i in ((kk-1)*freq+1+round(kn)):((kk-1)*freq+freq-round(kn)-1)){
                        for (j in (i-round(kn)):(i-1)){if (abs(Input1[j,combos[2,m]])<=trunc[kk,combos[2,m]]){
                        chatm2[i-round(kn)-(kk-1)*freq,]=chatm2[i-round(kn)-(kk-1)*freq,]
                        +Input1[j,combos[2,m]]*Input1[j,combos[2,m]]}
                        else{chatm2[(i-round(kn)-(kk-1)*freq),]=chatm2[i-round(kn)-(kk-1)*freq,]}}
                        chatm2[i-round(kn)-(kk-1)*freq,]=(1/(kn*delta))*chatm2[i-round(kn)-(kk-1)*freq,]}
                    for (s in 1:Nn){set.seed(s);
                        U=matrix(0,nrow=freq-2*round(kn)-1,ncol=2);
                        for (i in 1:(freq-2*round(kn)-1)){U[i,]=mvrnorm(n = 1, mu, sigma)};
                        Up=matrix(0,nrow=freq-2*round(kn)-1,ncol=2)
                        for (i in 1:(freq-2*round(kn)-1)){Up[i,]=mvrnorm(n = 1, mu, sigma)}
                        K=runif(freq-2*round(kn)-1);
                        R1=sqrt(K)*sqrt(chatm1)*U[,1]+sqrt(1-K)*sqrt(chatp1)*Up[,1];
                        R2=sqrt(K)*sqrt(chatm2)*U[,1]+sqrt(1-K)*sqrt(chatp2)*Up[,1];
                        for (i in ((kk-1)*freq+1+round(kn)):((kk-1)*freq+freq-round(kn)-1)){
                        if (abs(Input1[i,combos[1,m]])>trunc[kk,combos[1,m]]
                        && abs(Input1[i,combos[2,m]])>trunc[kk,combos[2,m]]){
                        Dhatn[s]=Dhatn[s]+(Input1[i,combos[1,m]]*R2[i-round(kn)-(kk-1)*freq])^2
                        +(Input1[i,combos[2,m]]*R1[i-round(kn)-(kk-1)*freq])^2}else{Dhatn[s]=Dhatn[s]}}}
                        Dhatns=sort(Dhatn, decreasing = TRUE);
                        Z_alphad_10[kk,m]=Dhatns[Nn*0.1];
                        Z_alphad_5[kk,m]=Dhatns[round(Nn*0.05)];
                }}
        cnd_10=sqrt(delta)*(Z_alphad_10+A_hatn)/sqrt(V_g1*V_g2);
        cnd_5=sqrt(delta)*(Z_alphad_5+A_hatn)/sqrt(V_g1*V_g2);
##----------------------------------------------------------------------------------------------------------------------------##
```