# 3. MEASURING THE JUMP RISK CONTRIBUTION UNDER MARKET MICROSTRUCTURE NOISE – EVIDENCE FROM CHINESE STOCK MARKET[1]

**Chao YU**[2]
**Xujie ZHAO**[3]

## Abstract

*In this paper, we use the pre-averaging threshold method to measure the contribution of jump variation to the total price variation under the effect of market microstructure noise with financial high frequency data. We first show the advantages of our method by Monte Carlo simulation. Then, we apply the pre-averaging threshold estimator and bi-power variation estimator for comparison to the intraday data of Chinese stock market at different frequencies. The empirical results show that for the most stocks in our sample, the jump contribution estimated by noise-robust estimator at tick frequency is larger than the result at five-minute frequency, which is different from the result for US market that the jump variation is overestimated with lower-frequency data in Christensen et al. (2014). Moreover, jump jump component is an important contributor to the total risk in Chinese stock market.*

**Keywords:** financial high frequency data, jump risk contribution, market microstructure noise, pre-averaging

**JEL Classification:** C58, G17

## 1. Introduction

With the availability of reliable financial high frequency data over the last two decades, the dynamics of financial asset price can be analyzed closely. Many researches have demonstrated the presence of jumps in asset price process; see Barndorff-Nielsen and Shephard (2004, 2006), Huang and Tauchen (2005), Aït-Sahalia and Jacod (2009), Lee and Hannig (2010), Boudt, Croux and Laurent (2011), Lee and Mykland (2012), and so on. Jump component can account for the discontinuous movement in asset price, such as the 1987 crash in America, and is a very important source of financial risk.

---

[1] *This research was supported in part by NSFC (71601048), NSSFC (17CJY052), and the Fundamental Research Funds for the Central Universities in UIBE (CXTD9-07, CXTD10-10).*

[2] *School of Statistics, University of International Business and Economics, Beijing, P.R.China, E-mail: chaoyu@uibe.edu.cn*

[3] *School of International Trade and Economics, University of International Business and Economics, Beijing, P.R.China, E-mail: zhaoxujie@uibe.edu.cn. Corresponding author.*

Many researchers have focused on measuring the contribution of jump risk based on the realized measures with high frequency data. However, the empirical results from the literature show that jump risk contribution has large difference by using different frequencies data and measurement methods. For instance, Huang and Tauchen (2005) show that the proportion of jump variation over total price variation is 7.3% by using the 5-minute data of S&P500 from 1997 to 2002, whereas Andersen, Bollerslev and Diebold (2007) show that proportion of jump variation reaches up to 14.4% by using the same frequency data from 1990 to 2002. Tauchen and Zhou (2011) find that the estimation result of jump variation is only 5.4% by using the same frequency data from 1986 to 2005. This problem has attracted the researchers' attention. Christensen, Oomen and Podolskij (2014) apply the noise-robust jump variation measures to a set of equity and foreign exchange tick data in US market. They find that with lower-frequency data the proportion of jump variation is around 10%, while the jump variation with tick data is just over 1%. They obtain the conclusion that the jump variation measured with lower-frequency data tends to spuriously assign a burst of volatility to the jump component and then jump variation is overestimated.

Compared to the numerous researches related to the jump risk in developed markets, there are limited studies on the emerging markets. It is well known that stock returns in the emerging markets usually exhibit higher volatility, fatter tails, or more extreme fluctuations. Therefore, jumps in the emerging markets may be more frequent, and the ways in which they cause risks are likely to differ from the developed markets. In this paper we are concerned with the Chinese stock market, since it has become a more and more influencial emerging market around the world, with the rapid development in the past decades.. We use the intraday high frequency data of the Shanghai Stock Exchange (SSE) composite index and individual stocks to measure the contribution of jump variation over the total price variation in the Chinese stock market. By combining the pre-averaging method and threshold technique to overcome the effects of microstructure noise and jump, we use the pre-averaging threshold estimator to measure the jump contribution for the ultra high frequency data.

The main contribution of our paper contains the following two aspects. Firstly, we compare the sample properties of pre-averaging threshold estimator and pre-averaging bi-power variation estimator proposed by Christensen *et al.* (2014) by Monte Carlo simulation. The results show the advantages of pre-averaging threshold estimator in real application when the data frequency is very high. Secondly, we apply the pre-averaging threshold estimator to the tick data of Shanghai composite index and twenty individual stocks from various industries in Chinese stock market, and compare the performance with the bi-power variation (BV) estimator widely used in the senario of lower frequency data without consideration of noise. From the empirical analysis of the total sample and subsample including the trading days where jumps occurred, the results show that for most stocks, the jump contribution estimated by noise-robust estimator and the number of the days when jumps occurred obtained with tick data are larger than the results with five-minute data, which implies that there may exist frequent small jumps in the Chinese stock market. The results obtained are different from the results for the US market in Christensen *et al.* (2014). In addition, jumps still frequently happen and contribute more to the total risk, which should be paid more attention by the investors and regulators.

The rest of the paper is organized as follows. Section 2 describes the theoretical model and presents the pre-averaging threshold estimator to measure the jump contribution. Section 3 conducts the Monte Carlo simulation. Section 4 presents the empirical analysis. Section 5 concludes.

## 2. Measuring the Contribution of Jump Variation under Market Microstructure Noise

Assume that logarithmic price $X_t$ of a financial asset evolves as

$$dX_t = b_t dt + \sigma_t dW_t + dJ_t, \quad 0 \le t \le T, \tag{1}$$

where: $W_t$ is a standard Brownian motion. The drift $b_t$ and the volatility $\sigma_t$ are progressively measurable processes which guarantee that (1) has a unique, strong solution. $J_t$ is a Lévy jump process with a Lévy jump measure *v* and is independent of *W*. Note that $J_t$ can be written as the sum of "large" jump and "small" jump components:

$$J_t = \int_0^t \int_{\{|x|>1\}} x\mu(dx,ds) + \int_0^t \int_{\{|x|\le 1\}} x[\mu(dx,ds) - \nu(dx)ds] := J_{1t} + J_{2t}, \tag{2}$$

where: $\mu$ is the Poisson random measure of $J_t$. $J_{1t}$ is a compound Poisson process with finite activity of jump and can be further written as $J_{1t} = \sum_{i=1}^{N_t} Y_{\tau_i}$, where $N_t$ is a Poisson process with constant intensity $\lambda$ and $Y_{\tau_i}$ denotes the jump size at jump time $\tau_i$. $J_{2t}$ is a square integrable martingale with infinite activity of jump.

With the above setup, the total Quadratic Variation (QV) $[X]_T$ over interval [0, *T*] is

$$[X]_T = \int_0^T \sigma_s^2 ds + \sum_{s \le T} (\Delta J_s)^2, \tag{3}$$

where: $\int_0^T \sigma_s^2 ds$ is called Integrated Volatility (IV). Formula (3) shows that the total variation can be separated into two parts. One is continuous risk caused by ordinary volatility and the other is extreme risk caused by jumps. Our goal is to measure the contribution of jump component to the total price variation. Hence, we can define the jump contribution as

$$JC = \frac{[X]_T - \int_0^T \sigma_s^2 ds}{[X]_T}. \tag{4}$$

Based on (4), we can estimate jump contribution with discrete sampled observation by constructing the jump-robust estimators of IV and $[X]_T$, respectively. Assume that observations of $X_t$ are sampled at regularly spaced discrete times $0 = t_0 < t_1 < \cdots < t_n = T$ over a fixed time interval [0, *T*], *i.e.*, $t_i = i\Delta_n$ for $i = 0,1,\cdots,n$, where $\Delta_n = T/n$. A consistent estimator of $[X]_T$ is given by the Realized Variance (RV) (Banrndorff-Nielsen and Shephard, 2002):

$$RV = \sum_{i=1}^{n} (X_{t_i} - X_{t_{i-1}})^2 = \sum_{i=1}^{n} (\Delta_i^n X)^2. \tag{5}$$

For integrated volatility, the threshold estimator proposed by Mancini (2009) is given as:

$$TV = \sum_{i=1}^{n} (\Delta_i X)^2 I_{\{(\Delta_i X)^2 \le r(\Delta_n)\}}, \tag{6}$$

where threshold function $r(\Delta_n)$ is essential for identifying the intervals where no jumps occurred. Mancini (2009) proves that the threshold estimator is robust to the Lévy jumps.

However, these estimation methods are only consistent under the assumption of semimartingale price process (1) in an idealized world. The real application encounters the well-known bias problem caused by market microstructure noise, when the data frequency is very high. Such kind of noises are usually caused by the frictions in actual trades, such as tick size, discrete observation, bid-ask spread, and other trading mechanics. Therefore, when we use the ultra high frequency data, we have to deal with the noise. There are some methods to overcome the effect of noise in estimating the integrated volatility, including the two time-scale approach (Zhang, 2005, 2006), pre-averaging method (Podolskij and Vetter, 2009; Jacod, Li, Mykland, 2009) and realized kernel method (Barndorff-Nielsen and Shephard, 2008). Since these three methods are asymptotically equivalent, in this paper we adopt pre-averaging method to construct noise-robust estimator.

With the presence of microstructure noise, at any given time $t_i$, the actually observed log-price is $Z_{t_i}$ other than $X_{t_i}$, which can be given as

$$Z_{t_i} = X_{t_i} + \varepsilon_{t_i}, \tag{7}$$

where: $\varepsilon_{t_i}$ is the noise term, with $E(\varepsilon_t) = 0$, $Var(\varepsilon_t) = \sigma_\varepsilon^2$. We combine the pre-averaging method and threshold technique to obtain the estimator of IV which is robust to both jumps and noises. Let $\bar{Z}_i^n = \sum_{j=1}^{k_n - 1} g_j^n \Delta_{i+j}^n Z$, which denotes the weighted average of $k_n$ observations of $Z_i^n, Z_{i+1}^n, \cdots, Z_{i+k_n-1}^n$ with weights $g_j^n = g(j/k_n)$. The weighting function $g(x)$ is required to be continuous on [0, 1], piecewise $C^1$ with a piecewise Lipschitz derivative $g'$, and satisfy $g(0) = g(1) = 0, \int_0^1 g^2(s)ds > 0$. Moreover, the integer sequence $k_n$ satisfies $k_n \sqrt{\Delta_n} = \theta + o(\Delta_n^{1/4})$ for some constant $\theta > 0$. Define $\varphi_1 = \int_0^1 (g'(s))^2 ds$, $\varphi_2 = \int_0^1 g^2(s)ds$, then the pre-averaging estimator of $[X]_T$ can be given as

$$PRV = \frac{n}{n - k_n + 2} \frac{1}{k_n \varphi_2} \sum_{i=0}^{n-k_n+1} (\bar{Z}_i^n)^2 - \frac{\varphi_1}{\theta^2 \varphi_2} \sigma_\varepsilon^2. \tag{8}$$

The pre-averaging threshold estimator of IV is given as

$$PTV = \frac{n}{n - k_n + 2} \frac{1}{k_n \varphi_2} \sum_{i=0}^{n-k_n+1} (\bar{Z}_i^n)^2 I_{\{(\bar{Z}_i^n)^2 \le r(\Delta_n)\}} - \frac{\varphi_1}{\theta^2 \varphi_2} \sigma_\varepsilon^2, \tag{9}$$

where: $\frac{\varphi_1}{\theta^2\varphi_2}\sigma_\varepsilon^2$ is a bias-correction. The variance $\sigma_\varepsilon^2$ of noise term can be estimated by $\hat{\sigma}_\varepsilon^2 = \frac{1}{2n}\sum_{i=1}^{n}(\Delta_i Z)^2$ , which is proposed by Bandi and Russell (2006). Furthermore, the threshold function $r(\Delta_n)$ is required to satisfy the following conditions.

**Assumption 1**: Threshold function $r(\Delta_n)$ is a deterministic function of the step length $\Delta_n$ such that

(a) $\lim_{\Delta_n \to 0} r(\Delta_n) = 0$ ; (b) $\lim_{\Delta_n \to 0} \frac{\Delta_n^{1/2}(\log\frac{1}{\Delta_n})^2}{r(\Delta_n)} = 0$ .

Power functions $r(\Delta_n) = \beta\Delta_n^\alpha$ for any $\alpha \in (0,1/2)$ and $\beta \in R$ are possible choices. The threshold function $r(\Delta_n)$ satisfying Assumption 1 can be used to asymptotically identify the intervals where no jump occurred; also see the literature on the noise- and jump-robust volatility estimation (Jing, Liu and Kong, 2014).

Jacod *et al.* (2009) show that $PRV \xrightarrow{P} [X]_T$ , and Jing *et al.* (2014) prove that $PTV \xrightarrow{P} \int_0^T \sigma_s^2 ds$ , Therefore, $PRV - PTV \xrightarrow{P} \sum_{s \le T}(\Delta J_s)^2$ . Based on these results, we can obtain the estimator of jump contribution in the presence of Lévy jump and microstructure noise:

$$\hat{JC} = \frac{PRV - PTV}{PRV} \quad (10)$$

Christensen *et al.* (2014) also use the pre-averaging method to deal with the noise, but they use the bi-power variation to diminish the effect of jumps in the case of finite activity jump and propose the pre-averaging bi-power variation (PBV) to estimate IV. In Section 3, we will show the advantages of PTV over PBV by simulation study.

## 3. The Monte Carlo Simulation

In this section, we study the finite sample properties of pre-averaging threshold volatility (PTV) estimator in three scenarios of no jumps, jumps with finite activity, and jumps with infinite activity respectively by Monte Carlo simulation. In order to show the advantages of our estimator, we report the results of pre-averaging bi-power volatility (PBV) estimator proposed by Christensen *et al.* (2014) for comparison.

We consider the following three models of asset price.

Model I: Brownian motion:

$$dX_t = \sigma dW_t, \text{ where } \sigma^2 = 0.09. \quad (11)$$

Model II: Brownian motion + finite activity jumps (compound Poisson process):

$$dX_t = \sigma dW_t + \sum_{i=1}^{N_t} Y_{\tau_i} , \quad (12)$$

where: $N_t$ is a Poisson process with intensity $\lambda$, and jump size $Y_{\tau_i} \sim N(0,\sigma_Y^2)$. We set $\lambda$=3, $\sigma_Y$=0.15.

Model III: Brownian motion + infinite activity jumps (Variance-Gamma process):

$$dX_t = \sigma dW_t + cG_t + \eta W_{G_t} \tag{13}$$

where: $G_t \sim Gamma(t/b, b)$, $c$ and $\eta$ are constants. Following Mancini (2009), set $b$ =0.23, $c$ =-0.2, $\eta$ =0.2.

Let T=1. We simulate 1000 price paths for each model using first-order Euler discretisation scheme with n=40000. We further consider the effect of microstructure noises. Following Christensen *et al.* (2014), we assume that the noise series follows an AR(1) model:

$$\varepsilon_i = \beta\varepsilon_{i-1} + u_i \tag{14}$$

where: $u_i$ is i.i.d. and follows $N(0,\sigma_\varepsilon^2(1-\beta^2))$. From this assumption, if $\beta = 0$, the noises are i.i.d., while if $\beta \neq 0$, the noises are correlated. Furthermore, we select $\sigma_\varepsilon^2$ by fixing the noise ratio (Oomen, 2006) $\gamma = \sqrt{N\sigma_\varepsilon^2 / \int_0^1 \sigma_s^2 ds} = 0.5$. This ensures that the magnitude of the noise is in proportion to the efficient price variation. Then we can obtain the price paths contaminated with noises.

Following Jacod *et al.* (2009), we choose $g(x) = x \wedge (1-x)$ and set $\theta$ = [0.1, 0.3, 0.5]. In order to easily compare the estimation performance of PTV and PBV, we calculate the estimation results of the ratios over the true value IV: PTV/IV and PBV/IV. Table 1 and Table 2 report the results of the bias and variance of both ratios in the case of i.i.d. ($\beta = 0$) and correlated noises ($\beta = 0.3$) respectively.

From Table 1 and Table 2, one may see that whether the noise is independent or correlated, the performances of PTV and PBV have little difference for Model I where there are no jumps. However, for model II, where the price process contains jumps of finite activity, it's obviously that the PTV performs better than the PBV. In particular, the variance of PTV is much smaller than the ones of PBV, which suggests that estimation by PTV is more stable than PBV. For model III, where the price process contains jumps of infinite activity, PTV performs much better than PBV from both aspects of bias and variance. These results demonstrate the advantages of the estimator (10) for measuring the jump contribution in real applications.

**Table 1**

**Estimation Results of PTV/IV and PBV/IV ( $\beta = 0$ ) ($\times 10^{-2}$)**

| Model | PBV/IV | | | | | | PTV/IV | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | θ=0.1 | | θ=0.3 | | θ=0.5 | | θ=0.1 | | θ=0.3 | | θ=0.5 | |
| | Bias | Var. | Bias | Var. | Bias | Var. | Bias | Var. | Bias | Var. | Bias | Var. |
| I | 2.35 | 8.28e-4 | 4.92 | 2.19e-4 | 1.33 | 1.41e-4 | -0.94 | 6.18e-4 | 1.14 | 1.94e-4 | -0.44 | 2.60e-4 |
| II | 2.36 | 0.13 | 1.82 | 0.3 | 6.18 | 0.49 | -3.79 | 1.55e-3 | -7.18 | 4.32e-3 | -4.71 | 8.16e-3 |
| III | 8.64 | 0.14 | 12.09 | 0.37 | 11.84 | 0.58 | -2.16 | 2.43e-3 | 0.65 | 9.35e-3 | 2.07 | 0.02 |

**Table 2**

**Estimation Results of PTV/IV and PBV/IV ( $\beta = 0.3$ ) (×$10^{-2}$)**

| Model | PBV/IV | | | | | | PTV/IV | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | θ=0.1 | | θ=0.3 | | θ=0.5 | | θ=0.1 | | θ=0.3 | | θ=0.5 | |
| | Bias | Var. | Bias | Var. | Bias | Var. | Bias | Var. | Bias | Var. | Bias | Var. |
| I | -3.03 | 1.15e-3 | -8.9 | 3.93e-4 | -5.06 | 2.74e-4 | -5.33 | 1.01e-3 | -9.74 | 3.74e-4 | -8.19 | 1.72e-4 |
| II | 9.65 | 0.09 | 4.31 | 0.23 | 4.60 | 0.4 | 2.33 | 1.40e-3 | 0.04 | 1.97e-3 | -2.12 | 5.37e-3 |
| III | 6.99 | 0.11 | 7.28 | 1.97 | 7.48 | 1.13 | -3.13 | 2.67e-3 | -6.45 | 0.02 | -6.32 | 0.03 |

# 4. Empirical Study

In this section, we apply the PTV and BV (usually used in the case of lower frequency data, *e.g.*, of an order of minutes) estimators to the high frequency data to measure the contribution of jump risk in Chinese stock market. Our sample includes one market index and twenty individual stocks. In particular, we choose the composite index of Shanghai Stock Exchange and twenty individual stocks from ten industries, which are composed of the top two weighted stocks available of each industry index developed by China Securities Index Company. The sample period is from Jan. 4th, 2013 to Dec. 31st, 2013. If the data of one stock is unavailable or incomplete in 2013, the next stock would be chosen. The final selected index and stocks are listed in Table 3. We collect the tick transaction price data and five-minute transaction price data from 9:30 am to 15:00 pm each day for each one of these securities, which are from the RESSET high frequency database. Note that the so-called tick data here are factually recorded at frequency of around six seconds, which is the highest frequency of the data available in the database. There are totally 238 trading days in 2013, but for some stocks, the number of valid trading days is less than 238 after deleting the days with errors (incorrect date, etc.). Our study is designed to show the distinctions of jump variation contribution estimated by using different measuring methods and data at different frequencies, and then to reveal the jump contribution to the total risk in Chinese stock market under different resolutions.

**Table 3**

**The Information of the Market Index and Individual Stocks**

| Symbol | Stock Name | Sector |
|---|---|---|
| SH_INDEX | Composite index of Shanghai Stock Exchange | Market index |
| SINOPEC | Sinopec Group | Energy |
| SHENHUA | China Shenhua Energy | Energy |
| WANHUA | Wanhua Chemical Group | Raw material |
| ZIJIN | Zijin Mining Group | Raw material |
| LONGI | Xi'an Longi Silicon Materials | Industry |
| SANY | Sany Heavy Industry | Industry |
| CTGDF | China Tourism Group Duty Free | Optional consumption |
| GREE | Gree Electric Appliances | Optional consumption |
| YILI | Inner Mongolia Yili Industrial Group | Main consumption |
| MOUTAI | Kweichow Moutai | Main consumption |
| HENGRUI | Jiangsu Hengrui Medicine | Medicine and healthcare |
| CHNT | Changchun High & New Technology | Medicine and healthcare |
| PING AN | Ping An Insurance | Finance and real estate |
| CMB | China Merchants Bank | Finance and real estate |

| Symbol | Stock Name | Sector |
|---|---|---|
| LUXSHARE | Shenzhen Luxshare Precision Industry | Information technology |
| HIKVISION | Hangzhou Hikvision Digital Technology | Information technology |
| ZTE | Zhongxing Telecom Equipment | Telecom service |
| WINGTECH | Wingtech Technology | Telecom service |
| CYP | China Yangtze Power | Public utility |
| GDPD | GD Power Development | Public utility |

Let $P_{i,t}$ denote the *i*th price observation on day *t*. We calculate the *i*th return on day *t* as $r_{i,t} = [\ln(P_{i,t}) - \ln(P_{i-1,t})]\times 100$. Table 4 and Table 5 give the descriptive statistics of the return series computed by tick data and five-minute data respectively. From Table 4 and Table 5, one may see that for the twenty one assets, the sample means of return series are almost zero and the means of tick data are smaller. Meanwhile, the results of kurtosis for return series at tick frequency are much bigger than the results at five-minute frequency, which implies that the tick returns have thicker tails.

**Table 4**

**Descriptive Statistics for Returns of Tick Data**

| Symbol | Mean | Std. | Min. | Max. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| SH_INDEX | 2.1684e-5 | 0.0250 | -1.8295 | 2.0301 | 0.8179 | 242.0506 |
| SINOPEC | 2.1048e-5 | 0.1237 | -8.9612 | 9.1847 | 0.7559 | 215.8919 |
| SHENHUA | -8.4292e-05 | 0.0612 | -4.9199 | 5.2438 | 1.2578 | 403.5770 |
| WANHUA | 6.3781e-5 | 0.0825 | -2.6609 | 2.6384 | 0.1598 | 29.5153 |
| ZIJIN | 1.5890e-06 | 0.2064 | -3.6769 | 5.1960 | 0.0494 | 5.6237 |
| LONGI | 2.8309e-04 | 0.1236 | -6.2304 | 5.1508 | -0.1304 | 54.0303 |
| SANY | -4.3658e-05 | 0.0932 | -6.6215 | 9.3839 | 1.4403 | 504.7170 |
| CTGDF | 7.7804e-05 | 0.0888 | -8.5333 | 9.9640 | 0.7280 | 736.9451 |
| GREE | 9.5995e-06 | 0.0661 | -2.8059 | 4.0671 | 0.2356 | 51.9082 |
| YILI | 1.0394e-04 | 0.0871 | -4.1158 | 2.9653 | 0.0804 | 46.1443 |
| MOUTAI | -1.2428e-05 | 0.0519 | -2.0461 | 2.1955 | 0.1532 | 60.1934 |
| HENGRUI | 1.3281e-04 | 0.0837 | -4.0526 | 5.3008 | 0.2962 | 85.6976 |
| CHNT | 1.3504e-04 | 0.0802 | -3.3842 | 3.3114 | -0.1011 | 78.1946 |
| PING AN | 1.1933e-06 | 0.0512 | -2.7763 | 3.0199 | 0.4763 | 93.2065 |
| CMB | -1.6606e-05 | 0.0859 | -11.2341 | 11.1419 | 2.3630 | 1.9462e03 |
| LUXSHARE | 2.4419e-04 | 0.1231 | -4.3276 | 3.9671 | 0.2370 | 97.3585 |
| HIKVISION | 1.3685e-05 | 0.0863 | -2.9255 | 2.3878 | 0.0011 | 39.2971 |
| ZTE | 3.8883e-06 | 0.0786 | -2.5068 | 2.7764 | 0.0786 | 20.4009 |
| WINGTECH | 3.1841e-04 | 0.1330 | -6.9239 | 7.0718 | 0.1336 | 168.6446 |
| CYP | -2.6388e-05 | 0.1040 | -9.1476 | 9.5840 | 0.1562 | 818.7217 |
| GDPD | -5.2521e-05 | 0.2173 | -9.3755 | 8.9491 | -0.0973 | 24.8334 |

## *4.1 Jump Contribution for Total Sample*

In this subsection, we apply the two measuring methods of jump contribution to the data with different frequencies for all the valid trading days of each asset. The following estimators are used, which are constructed based on different estimators of QV and IV:

$$\hat{JC}_1 = 1 - \frac{PTV}{PRV}, \qquad \hat{JC}_2 = 1 - \frac{BV}{RV}. \tag{15}$$

$\hat{JC}_1$ estimator is applicable to the tick data, and $\hat{JC}_2$ estimator is applicable to the five-minute data, where BV is a widely used jump-robust estimator for IV at the spase sampling freqeuency (e.g. five minutes). For each trading day, we calculate the jump contribution by using $\hat{JC}_1$ and $\hat{JC}_2$, and then compute the values of the mean and standard deviation for the daily series of jump contribution. Table 6 presents the estimation results of $\hat{JC}_1$ for tick data and $\hat{JC}_2$ for five-minute data. Note that in the real application, due to the construction of the estimators, the case that the numerator BV is larger than the denominator RV in (15) would happen, which would result in the negative results of $\hat{JC}_2$. Since the negative jump contribution is not reasonable from the theoretical point, the statistics of $\hat{JC}_2$ in the table are calculated by discarding the negative values.

**Table 5**

**Descriptive Statistics for Returns of Five-minute Data**

| Symbol | Mean | Std. | Min. | Max. | Skewness | Kurtosis |
|---|---|---|---|---|---|---|
| SH_INDEX | 0.0014 | 0.1363 | -1.7643 | 2.0777 | 0.2158 | 13.8174 |
| SINOPEC | 0.0021 | 0.2509 | -1.4493 | 2.2076 | 0.4783 | 7.3368 |
| SHENHUA | -0.0033 | 0.2016 | -3.0405 | 2.3558 | 0.3161 | 14.4119 |
| WANHUA | 0.0036 | 0.3130 | -2.7771 | 2.6827 | 0.2207 | 6.9106 |
| ZIJIN | 0.0025 | 0.3042 | -2.1506 | 3.8319 | 0.6915 | 10.4502 |
| LONGI | 0.0112 | 0.5240 | -3.7462 | 4.9808 | 0.6599 | 9.6053 |
| SANY | -0.0023 | 0.3072 | -4.4736 | 4.1932 | 0.7139 | 21.1319 |
| CTGDF | 0.0032 | 0.3199 | -2.3870 | 2.9999 | 0.4245 | 8.6626 |
| GREE | 0.0010 | 0.3024 | -3.2261 | 2.6440 | 0.1028 | 8.8512 |
| YILI | 0.0041 | 0.3865 | -2.4701 | 3.1449 | 0.3225 | 7.2329 |
| MOUTAI | -0.0013 | 0.2360 | -3.3336 | 2.2347 | -0.0334 | 14.0502 |
| HENGRUI | 0.0053 | 0.3160 | -1.7939 | 2.7737 | 0.3974 | 6.6028 |
| CHNT | 0.0026 | 0.3856 | -2.5906 | 3.4717 | 0.4296 | 7.6771 |
| PING AN | -0.0009 | 0.2634 | -3.0318 | 2.7446 | 0.3519 | 10.7043 |
| CMB | 0.0002 | 0.2850 | -3.2277 | 3.7876 | 0.4421 | 13.6735 |
| LUXSHARE | 0.0096 | 0.4670 | -6.2889 | 4.4143 | 0.2920 | 11.8826 |
| HIKVISION | 0.0009 | 0.4042 | -3.6490 | 3.1394 | 0.2595 | 7.6925 |
| ZTE | 0.0001 | 0.3980 | -4.3582 | 4.2721 | 0.3526 | 12.3015 |
| WINGTECH | 0.0078 | 0.4536 | -4.4568 | 4.2090 | 0.2684 | 10.8524 |
| CYP | -0.0013 | 0.2134 | -1.5083 | 2.3563 | 0.3998 | 7.3904 |
| GDPD | -0.0014 | 0.3263 | -1.6598 | 2.6938 | 0.2876 | 4.8783 |

Table 6 shows that for most stocks (19 stocks among 21 stocks), the results of average daily jump contribution estimated by $\hat{JC}_1$ using tick data are larger than the results estimated by $\hat{JC}_2$ using five-minute data, which is obviously supported by the results of kurtosis above. From the theory of high frequency data analysis, on one hand, the BV estimator is robust to the large jumps and the PTV estimator is robust to the noise and Lévy jumps. On the other hand, the small jumps would be more likely to be captured at higher frequency than sparse sampling frequency. Therefore, the results of $\hat{JC}_2$ mainly reflect the contribution of large jump components at five-minute frequency and the results of $\hat{JC}_1$ may consist of both large jump and small jump components at tick frequency. Hence, the result that $\hat{JC}_1$ is much larger than $\hat{JC}_2$ may imply the exsitence of frequent small jump components in the asset price. It suggests that there are frequent discontinouous movements for the most assets, which is in accordance with the charactersitics of emerging markets. Furthermore, the two exceptions (ZIJIN and GDPD, marked in bold in Table 6) that estimated by $\hat{JC}_1$ is less than the average jump contribution estimated by $\hat{JC}_2$, may be attributed to the situations that a series of small continuous movements in the same direction at tick frequency is finally aggregated into a large jump captured at five-minute frequency. In general, the results are different from the findings of US market obtained in Christensen *et al.* (2014).

**Table 6**

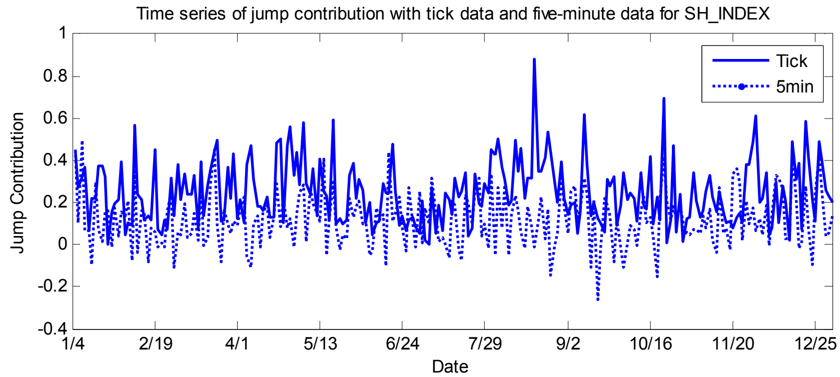
**Results of $\hat{JC}_1$ for Tick Data and $\hat{JC}_2$ for Five-minute Data with Total Sample**

| Symbol | Tick data | | Five-minute data | |
|---|---|---|---|---|
| | Mean | Std. | Mean | Std. |
| SH_INDEX | 0.2479 | 0.1499 | 0.1571 | 0.1056 |
| SINOPEC | 0.2138 | 0.1199 | 0.1815 | 0.1240 |
| SHENHUA | 0.3435 | 0.1256 | 0.1443 | 0.1050 |
| WANHUA | 0.2885 | 0.1018 | 0.1629 | 0.1131 |
| **ZIJIN** | **0.1083** | 0.1239 | **0.2627** | 0.1537 |
| LONGI | 0.2917 | 0.1351 | 0.1654 | 0.1260 |
| SANY | 0.2959 | 0.1320 | 0.2089 | 0.1365 |
| CTGDF | 0.3080 | 0.1217 | 0.1514 | 0.1101 |
| GREE | 0.3132 | 0.1135 | 0.1709 | 0.1156 |
| YILI | 0.3198 | 0.1111 | 0.1343 | 0.0874 |
| MOUTAI | 0.3823 | 0.1245 | 0.1915 | 0.1423 |
| HENGRUI | 0.3247 | 0.1305 | 0.1548 | 0.1157 |
| CHNT | 0.2804 | 0.1254 | 0.1617 | 0.1190 |
| PING AN | 0.2672 | 0.1210 | 0.1643 | 0.1171 |
| CMB | 0.3304 | 0.1194 | 0.1747 | 0.1230 |
| LUXSHARE | 0.4305 | 0.1525 | 0.1949 | 0.1237 |
| HIKVISION | 0.3369 | 0.1119 | 0.1567 | 0.1117 |
| ZTE | 0.2670 | 0.1017 | 0.1651 | 0.1244 |
| WINGTECH | 0.3852 | 0.1697 | 0.1796 | 0.1254 |
| CYP | 0.2455 | 0.1048 | 0.1926 | 0.1205 |
| **GDPD** | **0.1248** | 0.1150 | **0.2543** | 0.1432 |

Next, we take the market index SH_INDEX and stock ZIJIN as the examples to further illustrate the results. Figure 1 shows the time series of daily jump contribution estimated by tick data and five-minute data respectively for Shanghai composite index. From the figure, one may see clearly that the solid line (jump contribution results of tick data) is almost over the dotted line (results of five-minute data).

**Figure 1**

**Time Series of Daily Jump Contribution Results of Tick Data and Five-minute Data for SH_INDEX**

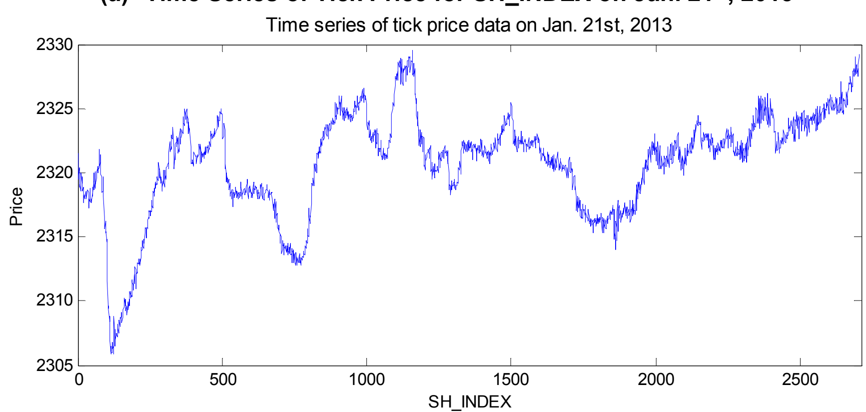



In order to look into the cause, we further present the time series of tick price data and five-minute price data data for SH_INDEX on Jan. 21st, 2013 in Figure 2. On this day, the average jump contribution estimated by tick data is 0.3193, and the average jump contribution estimated by five-minute data is 0.1557. From Figure 2, one may see that comparing the first half of the subfigure (a) with the corresponding part of (b), the obvious several large jumps can be observed at both frequencies, while in the second half of the subfigure (a), there are obviously a lot of fluctuations at tick frequency, but they are not reflected at five-minute frequency in subfigure (b) due to the lower resolution. This may be the situation that leads to the lager jump contribution at higher frequency.

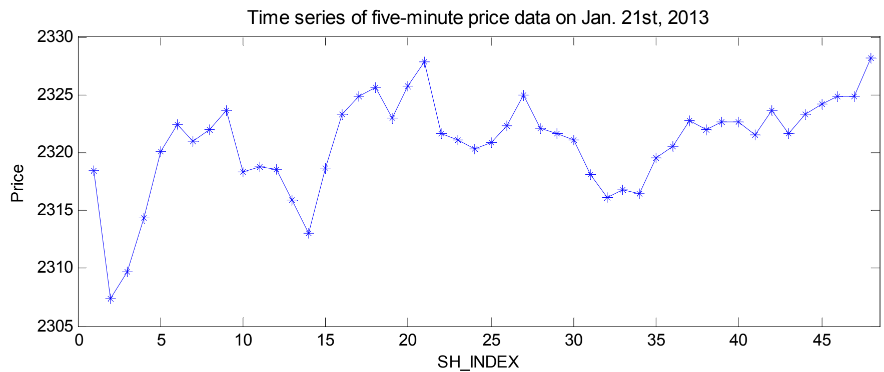
**Figure 2**

**Time Series of Tick Price and Five-minute Price for SH_INDEX on Jan. 21st, 2013**

**(a) Time Series of Tick Price for SH_INDEX on Jan. 21st, 2013**



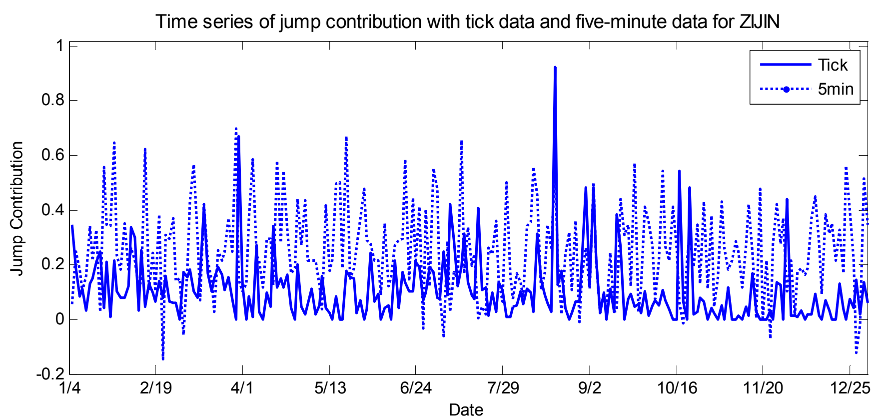
**(b) Time Series of Five-minute Price for SH_INDEX on Jan. 21st, 2013**




Now, we turn to analyze the exception - stock ZIJIN. Figure 3 shows the time series of jump contribution estimated by tick data and five-minute data for stock ZIJIN. From the figure, one may see that the most jump contributions estimated by tick data (solid line) are smaller than the results of five-minute data (dotted line).

**Figure 3**

**Time Series of Daily Jump Contribution Results of Tick Data and Five-minute Data for Stock ZIJIN**

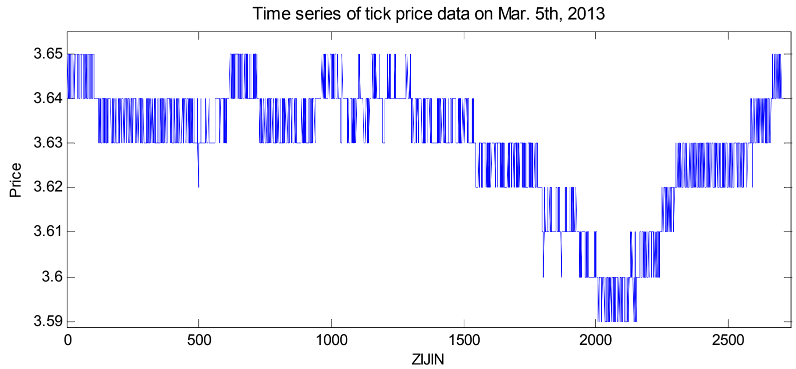



We further present the time series of tick price data and five-minute price data for stock ZIJIN on Mar. 5th, 2013 in Figure 4. On this day, the average jump contribution estimated by tick data is 0.1026, and the average jump contribution estimated by five-minute data is 0.5670. The figure shows that in subfigure (a), the price at tick frequency almost moves back and forth over a small interval, but the price sampling at five minutes exhibits some relatively large steps, which might be caused by several small movements in the same direction. We find that this situation appears frequently for stock ZIJIN in 2013, which may be the reason for the results that the average jump contribution at tick frequency is smaller than the ones at five-minute frequency.

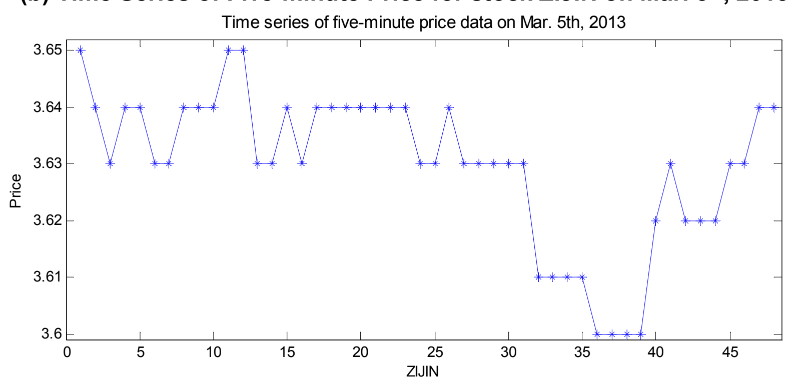
**Figure 4**

**Time Series of Tick Price and Five-minute Price for Stock ZIJIN on Mar. 5th, 2013**

**(a) Time Series of Tick Price for stock ZIJIN on Mar. 5th, 2013**




**(b) Time Series of Five-minute Price for stock ZIJIN on Mar. 5th, 2013**




## *4.2 Jump Contribution when Jump Occurred*

In this subsection, we focus on the jump contribution to the total price variation on the days when the jumps occurred. Based on the noise-robust estimators of QV and IV, the method of testing whether the jumps happened can be easily constructed. Following the method proposed by Christensen *et al.* (2014), the test statistic is given by

$$Z_{PTV} = \frac{n^{1/4}(\ln PRV - \ln PTV)}{\sqrt{\Sigma_{11} + \Sigma_{22} - 2\Sigma_{12}} \,/\, \int_0^1 \sigma_s^2 ds} \xrightarrow{d} N(0,1)\,, \tag{16}$$

where the covariance $\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$ can be estimated by block subsample method; see Christensen *et al.* (2014) for details. We apply this testing method to the tick data and apply the testing method based on BV proposed in Huang and Tauchen (2005) (the statistic denoted by $Z_{BV}$) without consideration of microstructure noise to five-minute data. For each valid trading day, we can first obtain the testing results of whether there are jumps. Then, we can measure the jump contribution on the days when jumps occurred. According to the jump testing method, we define the jump contribution on the days when jumps occurred as

$$JC_t^{PTV} = \left(1 - \frac{PTV}{PRV}\right) I_{\{|Z_{PTV}| > z_{\alpha/2}\}}, \text{ and } JC_t^{BV} = \left(1 - \frac{BV}{RV}\right) I_{\{|Z_{BV}| > z_{\alpha/2}\}}. \quad (17)$$

Table 7 gives the mean values and standard deviations of daily $JC_t^{PTV}$ for tick data and $JC_t^{BV}$ for five-minute data, and the number of the days when jumps occurred (the columns of Counts) tested by $Z_{PTV}$ and $Z_{BV}$, respectively, as well.

From Table 7, one may see that for most stocks (18 stocks among 21 stocks), the number of days when jump occurred tested at tick frequency is larger than the number of days tested at five-minute frequency, which further verifies the results obtained from Table 6. The exceptions marked in bold in the table also contain stock ZIJIN and GDPD, which is consistent with the results in Table 6. Besides, on the trading days when the jumps occurred, there is no clear pattern of mean values of jump contribution between the two estimation methods, since the mean values not only depend on the number of days when jumps occurred but also the variation levels of small jumps and big jumps.

To sum up, for the Chinese stock market, no matter from the analysis of the total sample or the subsample where jumps occurred, we do not find the phenomenon occurring in the US market that the jump contribution is uniformly overestimated by the data at sparse frequency (Christensen *et al*., 2014). On the contrary, for the most stocks, the average jump contribution and the number of the jump days obtained by using data at higher frequency is larger than the results at lower frequency. This phenomenon may be attributed to the frequent small fluctuations of the asset price in the intraday short time period, which may be captured as small jump components in the short-time period, but may be smoothed out over a longer-time period. As a result, the jump contribution is larger at the higher resolution than the lower resolution. Therefore, we think that the levels of jump contribution estimated by correctly applying suitable methods to the data with different frequencies can reflect the contribution of different types of jumps under different resolutions. The final difference of jump contribution between the higher frequency and lower frequency depends on the intraday price dynamics of each asset, and thus there may be no uniform results. Furthermore, jumps contribute more to the total risk in the Chinese stock market. Take the Shanghai composit index for example; the average daily jump contribution is 0.2345 for tick data and 0.2616 for five-minute data when jumps occurred (see the resuts of SH_INDEX in Table 7). Therefore, the jump component is an important risk factor in the Chinese market.

**Table 7**

**Results of Jump Contribution for Tick Data and Five-minute Data when Jumps Occurred**

| Symbol | Tick data | | | Five-minute data | | |
|---|---|---|---|---|---|---|
| | Mean | Std. | Counts | Mean | Std. | Counts |
| SH_INDEX | 0.2345 | 0.1271 | 104 | 0.2616 | 0.0839 | 74 |
| SINOPEC | 0.2112 | 0.0984 | 126 | 0.2726 | 0.1050 | 99 |
| SHENHUA | 0.3096 | 0.0973 | 137 | 0.2451 | 0.0890 | 67 |
| WANHUA | 0.2727 | 0.0784 | 139 | 0.2671 | 0.0986 | 72 |
| **ZIJIN** | 0.1259 | 0.0855 | **96** | 0.3328 | 0.1370 | **147** |

| Symbol | Tick data | | | Five-minute data | | |
|---|---|---|---|---|---|---|
| | Mean | Std. | Counts | Mean | Std. | Counts |
| LONGI | 0.2577 | 0.0814 | 118 | 0.2602 | 0.1052 | 81 |
| SANY | 0.2618 | 0.1015 | 117 | 0.3094 | 0.1217 | 89 |
| CTGDF | 0.2745 | 0.0858 | 132 | 0.2558 | 0.0860 | 68 |
| GREE | 0.2946 | 0.1015 | 170 | 0.2738 | 0.0891 | 75 |
| YILI | 0.2975 | 0.0931 | 129 | 0.2277 | 0.0682 | 51 |
| MOUTAI | 0.3404 | 0.0919 | 117 | 0.3044 | 0.1251 | 82 |
| HENGRUI | 0.2947 | 0.1098 | 146 | 0.2582 | 0.0938 | 63 |
| CHNT | 0.2608 | 0.1013 | 156 | 0.2706 | 0.1024 | 75 |
| PING AN | 0.2661 | 0.0915 | 117 | 0.2661 | 0.1090 | 69 |
| CMB | 0.3160 | 0.1076 | 125 | 0.2801 | 0.0926 | 81 |
| LUXSHARE | 0.3798 | 0.1390 | 123 | 0.2894 | 0.0969 | 97 |
| HIKVISION | 0.3207 | 0.1012 | 157 | 0.2631 | 0.0874 | 67 |
| ZTE | 0.2631 | 0.0885 | 143 | 0.2933 | 0.1013 | 66 |
| **WINGTECH** | 0.3225 | 0.1269 | **91** | 0.2843 | 0.1037 | **107** |
| CYP | 0.2412 | 0.1030 | 139 | 0.2832 | 0.0918 | 94 |
| **GDPD** | 0.1378 | 0.0926 | **110** | 0.3289 | 0.1156 | **142** |

## 5. Conclusions

In this paper, we measure the contribution of jumps to total price variation with the ultra high frequency data under the effect of microstructure noises. By combining the pre-averaging method and threshold technique to deal with noises and jumps respectively, we construct the pre-averaging threshold estimator for the jump variation contribution. Moreover, we compare the finite sample properties of the pre-averaging threshold estimator and the pre-averaging bi-power variation estimator proposed by Christensen *et al.* (2014) by Monte Carlo simulation. The results show the advantages of our estimator for the Lévy jump.

Finally, we apply the pre-averaging threshold estimator and bi-power variation estimator for comparison to measure the jump contribution of Chinese stock market using the intraday data of Shanghai Stock Exchange composite index and twenty individual stocks. The empirical results show that for most assets, the average jump contribution to the daily price variation and the number of the days when jumps occurred obtained by using tick frequency data are larger than the results by using five-minute frequency data. This implies the exsitence of the frequent small jump components in the Chinese stock market. The results obtained here are different from the results for US market found in Christensen *et al.* (2014). We believe that the jump contribution estimated by correctly applying suitable methods to the data at different frequencies can factually reflect the contribution of different types of jumps under different resolutions. But, the final difference of jump contribution between the higher frequency and lower frequency depends on the intraday price dynamics of each asset. Moreover, the jump component contributes more to the total price variation and is a very important risk factor in the Chinese stock market.

## References

Aït-Sahalia, Y. and Jacod, J., 2009. Testing for jumps in a discretely observed process. *Annals of Statistics*, 37(1), pp.184-222.

Andersen, T.G., Bollerslev, T. and Diebold, F.X., 2007. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility. *Review of Economics and Statistics*, 89(4), pp.701-720.

Bandi, F.M. and Russell, J.R., 2006. Separating microstructure noise from volatility. *Journal of Financial Economics*, 79(3), pp.655-692.

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

Barndorff-Nielsen, O.E., Shephard, N. and Winkel, M., 2006. Limit theorems for multipower variation in the presence of jumps. *Stochastic Processes and their Applications*, 116(5), pp.796-806.

Barndorff-Nielsen O.E. and Shephard, N., 2002. Estimating quadratic variation using realized variance. *Journal of Applied Econometrics*, 17(5), pp.457-477.

Barndorff-Nielsen, O.E., Hansen, P.R., Lunde, A. and Shephard, N., 2008. Designing realised kernels to measure ex-post variation of equity prices in the presence of noise. *Econometrica*, 76(6), pp.1481-1536.

Boudt, K., Croux, C. and Laurent, S., 2011. Robust estimation of intraweek periodicity in volatility and jump detection. *Journal of Empirical Finance*, 18(2), pp.353-367.

Christensen, K., Oomen, R. and Podolskij, M., 2014. Fact or friction: Jumps at ultra-high frequency. *Journal of Financial Economics*, 114(3), pp.576-599.

Huang, X. and Tauchen, G., 2005. The relative contribution of jumps to total price variance. *Journal of Financial Econometrics*, 3(4), pp.456-499.

Jacod, J., Li, Y., Mykland, P.A. and Podolskij, M., 2009. Microstructure noise in the continuous case: The pre-averaging approach. *Stochastic Processes and their Applications*, 119(7), pp.2249-2276.

Jing, B.Y., Liu, Z. and Kong, X.B., 2014. On the estimation of integrated volatility with jumps and microstructure noise. *Journal of Business and Economic Statistics*, 32(3), pp.457-467.

Lee, S.S. and Mykland, P.A., 2012. Jumps in equilibrium prices and market microstructure noise. *Journal of Econometrics*, 168(2), pp.396-406.

Lee, S.S. and Hannig, J., 2010. Detecting jumps from Lévy jump diffusion processes. *Journal of Financial Economics*, 96(2), pp.271-290.

Mancini, C., 2009. Non parametric threshold estimation for models with stochastic diffusion coefficient and jumps. *Scandinavian Journal of Statistics*, 36(2), pp.270-296.

Oomen, R.C.A., 2006. Comment on 2005 JBES invited address “Realized variance and market microstructure noise” by Peter R. Hansen and Asger Lunde, *Journal of Business and Economic Statistics*, 24(2), pp.195-202.

Podolskij, M. and Vetter, M., 2009. Estimation of volatility functionals in the simultaneous presence of microstructure noise and jumps. *Bernoulli*, 15(3), pp.634-658.

Tauchen, G.E. and Zhou, H., 2011. Realized jumps on financial markets and predicting credit spreads. *Journal of Econometrics*, 160(1), pp.102-118.

Zhang, L., Mykland, P.A. and 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(472), pp.1394-1411.

Zhang, L., 2006. Efficient estimation of stochastic volatility using noisy observations: A multi-scale approach. *Bernoulli*, 12(6), pp.1019-1043.