Senior Lecturer Virgil DAMIAN, PhD
E-mail: virgil.damian@fin.ase.ro
The Bucharest University of Economic Studies
Cosmin – Octavian CEPOI, PhD
E-mail: cosmin.cepoi@fin.ase.ro
The Bucharest University of Economic Studies

VOLATILITY ESTIMATORS WITH HIGH-FREQUENCY DATA FROM BUCHAREST STOCK EXCHANGE

Abstract. *In this paper, we have focused on accurately volatility estimation due to its crucial importance in investment and risk management activities. Based on tick by tick data, provided by Thomson Reuters, we have realized a comparative study among different high-frequency volatility estimators for some of the most important three companies listed on Bucharest Stock Exchange. Our findings emphasize that the presence of jumps or microstructure noises affect the efficiency of realized volatility estimator. So, based on data architecture, we have used adequate estimators jump and noise robust. We concluded that for less liquid markets, the presence of more visible jumps leads to higher intra-day volatilities comparing with more liquid markets.*

Keywords: *stochastic volatility, realized variance, realized kernel, two time scales estimator, jump, stochastic integral representation*.

JEL Classification:C58, C13, G10.

1. Introduction

Although price process of financial instruments is generally observable, the volatility is always an unobservable variable, hence the necessity to be estimated. Over time, parametric models, like ARCH family models, Stochastic Volatility Model, Local Volatility Model *etc.*, have been proposed to deal with this fundamental problem essential for practical applications as pricing financial assets, performance evaluation, risk management *etc*.

Over the last decade, a new paradigm, related to high frequency trading and pricing risk, has been evolved. One popular application of high frequency data represents the estimation of the *Integrated Variation* (IV).A conventional estimator is *Realized Variance* (RV), defined as the sum of sampled squared returns. It is well known in financial econometrics that both jumps and microstructure noise are usually met involved in high frequency time series. Unfortunately, when price series data is contaminated by microstructure noise or jumps, RV estimator is biased. In order to defeat this inconvenient, Barndorff-Nielsen and Shepard(2004), Aït-Sahalia *et al.*(2005), Zhang *et al.* (2005), Boudt and Zhang (2010) *etc.* proposed different kind of high-frequency unbiased variance estimators.

Using three blue chips listed on Bucharest Stock Exchange (BSE), we have tested for the presence of jumps and microstructure noises in price series. Depending on the results obtained, we proposed adequate high-frequency variance estimators. In order to make a back-testing of our results, we provided connections between periods with high-volatility to market announcements .*This approach is a new perspective in volatility estimation based on high frequency data from Bucharest Stock Exchange*.

The paper is organized as follows. Section 2 presents the most important tests for detecting the presence of jumps in price series data. In Section 3 we discuss about the implication of microstructure noises in quality of high-frequency variance estimators computation. Section 4 presents different type of high-frequency estimators for intra-day variance. In Section 5, estimation results are analyzed. Section 6 concludes this research.

2. Tests for the presence of jumps in price series

There are several important tests for detecting the presence of jumps in financial asset prices. For our purpose, we present here two of them: Barndorff-Nielsen and Shephard (2006), respectively Aït- Sahalia and Jacod (2009).

Under the assumption of no arbitrage, price processes must follow a semi-martingale(see, *e.g.*, Delbaen and Schachermayer (1994)) on a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$.The basic mathematical framework underlying jump tests requires that the stochastic log-price process, denoted by $(\eta_t)_{0\leq t<\infty}$, belongs to the class of semi-martingales, represented by an stochastic integral form:

$$\eta_t = \eta_0 + \int_0^t a_u du + \int_0^t \sigma_u dB_u + Z_t, \qquad (2.1)$$

where $(a_u)_{0\leq u\leq t}$ denotes the drift adapted càdlàg process, $(\sigma_u)_{0\leq u\leq t}$ is an adapted càdlàg volatility process associated with the instantaneous conditional mean and volatility of the corresponding return, $(B_u)_{0\leq u\leq t}$ is a standard Brownian motion, and $(Z_t)_{0\leq t<\infty}$ is a jump process defined as $Z_t = \sum_{j=1}^{N_t} k_j$, where k_j are random variables with nonzero values, and $(N_t)_{0\leq t<\infty}$ is a counting process (*e.g.*, a Poisson stochastic process).

Let us consider the time horizon t, divided into M sub-intervals of constant length and let $h = \frac{1}{M}$. Let us define the $i-th$ intraday return on day t (denoted by $r_{t,i}$) as:

$$r_{t,i} := \eta_{t-1+ih} - \eta_{t-1+(i-1)h}, \tag{2.2}$$

for every $i \in \{1,2,\dots,M\}$.

A standard measure of quadratic variation of process (2.1), introduced in Andersen and Bollerslev (1998), is *realized variance* (RV)[1] defined as:

$$RV_{t,M} := \sum_{i=1}^{M} r_{t,i}^2. \tag{2.3}$$

They also proved that the variation of stochastic process $(\eta_t)_{0\le t<\infty}$ can be, in this way, decomposed in two sources of variability, since the limit of $RV_{t,M}(M \to \infty)$, defined in (2.3), can be represented as:

$$\lim_{M\to\infty} RV_{t,M} = \int_0^t \sigma_u^2 du + \sum_{j=1}^{N_t} k_j^2 := IV_t + JV_t, \tag{2.4}$$

where IV and JV denote the *integrated variance* and *jump variation*, respectively. Hence, if it is possible to sample frequently enough, IV_t can be estimated with minimal estimation error.

2.1. Barndorff-Nielsen and Shepard test (BS test)

Historically, the first statistical test to identify jumps in high-frequency time series was introduced by Barndorff-Nielsen and Shephard (2004, 2006). To estimate integrated variance (IV estimator; see (2.4)) in presence of jumps, Barndorff-Nielsen and Shephard propose the *realized bi-power variation* (denoted by BPV), defined by:

$$BPV_{t,M} := \sum_{i=2}^{M} |r_{t,(i-1)}| \cdot |r_{t,i}|. \tag{2.5}$$

The main idea of using bi-power variation, as an IV estimator, is that the likelihood of observing presence of jumps in two consecutive returns approaches zero sufficiently fast as the sampling frequency increases. Thus, the jumps contribution is eliminated since the product of two arbitrary consecutive returns is influenced by the diffusion part only.

Since the limit of realized volatility, for $M \to \infty$, converges to the sum of integrated variance and jump variation (see (2.4)), it follows that the difference $RV_{t,M} - BPV_{t,M}$ captures the jump part only. Taking into account the joint Central Limit Theorem of RV and BPV, they proposed the following test statistics:

[1] In the literature it is often also called *realized volatility* though in a strict sense, the terminology "volatility" is typically used for σ_t rather than for σ_t^2.

$$BS_{t,M} := \frac{RV_{t,M} - BPV_{t,M}}{\sqrt{\theta_2 \cdot h \cdot QPQ_{t,M}}}, \tag{2.6}$$

which converges to a standard normal distribution with appropriate scaling. In (2.6), $QPQ_{t,M}$denotes the *realized quad-power quarticity*, given by:

$$QPQ_{t,M} := M \sum_{i=4}^{M} |r_{t,(i-3)}| \cdot |r_{t,(i-2)}| \cdot |r_{t,(i-1)}| \cdot |r_{t,i}| \longrightarrow \int_0^t \sigma_u^4 du. \tag{2.7}$$

Also, in (2.6) θ_p is given by:

$$\theta_p := \mu_{2/p}^{-2p} \cdot \left[\mu_{4/p}^{p} + (1-2p) \cdot \mu_{2/p}^{2p} + 2 \cdot \sum_{j=1}^{p-1} \mu_{p/4}^{p-j} \cdot \mu_{2/p}^{2j} \right], \tag{2.8}$$

whereμ_p is a notation for the $p - th$*absolute moment* of a standard normal random variable $U \sim \mathcal{N}(0,1)$, defined by:

$$\mu_p := \mathbb{E}[|U|^p] = \frac{1}{\sqrt{\pi}} 2^{p/2} \Gamma\left(\frac{p+1}{2}\right), \tag{2.9}$$

with$\Gamma(\cdot)$ the *Euler integral of the second kind* (*Gamma function*).

2.2. Aït-Sahalia and Jacod test (AJ test)

This test, based on the research of Aït-Sahalia and Jacod (2009), uses the convergence properties of power variation and its dependence on the time scale on which it is measured.

The main idea is to compare the multi-power variation of returns recorded at equidistant time intervals at an infinitesimal scale (h), denoted by $r_{t,i}$, with those recorded at the slower time scale (kh), denoted by $y_{t,i}$, for every $i \in \{1,2,\dots,M/k\}$.

They found that the asymptotic value (obtained for $M \to \infty$) of the realized power variation($\hat{S}_t(p,k,h)$ introduced below) is invariant with respect to different sampling scales and that their value is 1in case of jumps and a known number in the absence of jumps ($k^{p/2} - 1$, where p is a positive number; see Aït-Sahalia and Jacod (2009) for details and numerical examples).

Technically speaking, the statistic of an *AJ test* is:

$$AJ_{t,M} := \frac{\hat{S}_t(p,k,h) - k^{p/2-1}}{\sqrt{\hat{V}_{t,M}}}, \qquad p > 2, \tag{2.10}$$

where $\hat{S}_t(p,k,h) := \frac{\hat{B}_t(p,kh)}{\hat{B}_t(p,h)}$, and $\hat{B}_t(p,h) := \sum_{i=1}^{M} \left| r_{t,i} \right|^p$ denotes the usual power variation. $\hat{V}_{t,M}$ denotes the *asymptotic variance* of $\hat{S}_t(p,k,h)$ and is defined by:

$$\hat{V}_{t,M} := h \frac{N(p,k) \cdot \hat{A}_t(2p,h)}{\left[\hat{A}_t(p,h)\right]^2}, \tag{2.11}$$

where:

$$N(p,k) := \frac{1}{\mu_p^2} \left[k^{p-2}(k+1)\mu_{2p} + k^{p-2}(k-1)\mu_p^2 - 2k^{p/2-1}\mu_{k,p} \right], \tag{2.12a}$$

$$\hat{A}_t(p,h) := \frac{h^{1-p/2}}{\mu_p} \sum_{i=1}^{M} \left| r_{t,i} \right|^p \chi_{\{|r_{t,i}| \leq \alpha h^{\varpi}\}}, \tag{2.12b}$$

$$\mu_{k,p} := \mathbb{E}\left[|U|^p \left| U + \sqrt{k-1} V \right|^p \right], \tag{2.12c}$$

with p, k, α, ϖ parameters, and $V \sim \mathcal{N}(0,1)$. μ_p was introduced in section 2.1, relation (2.9). $\chi_{\{\cdot\}}$ denotes the *indicator function*.

Consequently, the AJ test identifies the presence of jumps using the ratio of realized power variation sampled from two time scales. The null hypothesis of this test is that no jumps occurred.

3. Identifying the presence of microstructures noise in price process

Let us denote by $(\tilde{\eta}_t)_{0 \leq t < \infty}$, the log-price process of some asset, for which its analytical representation can be modeled by an Itô process:

$$\tilde{\eta}_t = \tilde{\eta}_0 + \int_0^t \tilde{a}_u du + \int_0^t \tilde{\sigma}_u dB_u, \tag{3.1}$$

where $(\tilde{a}_u)_{0 \leq u \leq t}$, respectively $(\tilde{\sigma}_u)_{0 \leq u \leq t}$ are stochastic processes and both fulfill regularity conditions. If we consider the observed log-price process (3.1), we can infer that the data series consist of the so-called “true” logarithmic price process, but it also contains noise. The consistency of the realized variance estimator builds on the hypothesis that prices can be modeled according to the semi-martingale (3.1) and can be sampled arbitrarily frequently. In practice, the sampling frequency is certainly limited by the transaction frequency or actual quotation. Moreover, transaction prices are subject to market microstructure effects, such as the bid-ask bounce effect.

Hence, instead of the Itô representation (3.1), it is more realistic to assume the observable price process to be given by:

$$s_{t,i} = \tilde{\eta}_{t,i} + \varepsilon_{t,i}, \tag{3.2}$$

where $(\tilde{\eta}_t)_{0 \le t < \infty}$ is the *latent true*, or so-called *efficient*, price that follows the semi-martingale given in (3.1),and $(\varepsilon_t)_{0 \le t < \infty}$ is a zero mean error term interpreted as the *market microstructure noise*. In this case, the observed intraday return is given by:

$$r_{t,i} = \tilde{r}_{t,i} + u_{t,i}, \tag{3.3}$$

where $\tilde{r}_{t,i} := \tilde{\eta}_{t-1+ih} - \tilde{\eta}_{t-1+(i-1)h}$ and the intraday noise increment is $u_{t,i} = \varepsilon_{t-1+ih} - \varepsilon_{t-1+(i-1)h}$.

If ε_t is assumed to be i.i.d.,with $\sigma_\varepsilon^2 := \mathbb{E}[\varepsilon_t^2] := \omega^2$, then the observed high-frequency returns follow an MA(1) process. Moreover, it can be proved that:

$$\mathbb{E}\left[RV_{t,M}\right] = IV_t + 2M\omega^2. \tag{3.4}$$

Hence, as Hansen and Lunde (2006) proved in their research, RV is a biased estimator of integrated variance, with bias term $2M\omega^2$. Obviously, for $M \to \infty$, $RV_{t,M}$ diverges to infinity.Thus, the estimation is dominated by market microstructure noise.

Since it is beyond the aim of this article to provide an in-depth discussion of alternative integrated variance estimators, we only briefly present, in section 4,two of the most popular estimators accounting for the presence of noise used in section 5. It's about the *realized kernel* estimator, introduced by Barndorff-Nielsen *et al.* (2008) and about the *jump robust two scale realized variance*, introduced by Zhang *et al.* (2010). Other usual estimators are the *maximum likelihood* estimator proposed by Aït-Sahalia*et al.* (2005) and the *pre-averaging* estimator suggested by Jacod*et al.* (2009).

4. High-frequency data estimators

The object of interest in this kind of study is to estimate the continuous part of the quadratic variation, or the integrated variance (IV), defined in (2.4).

4.1. Estimation techniques without microstructure noise

Obviously, if in the considered series there are *no jumps* and *no microstructure noise*, then the most popular estimator is *realized variance* (see (2.3)).

Andersen*et al.* (2012) have proposed a new set of estimators for IV *in the presence of jumps*, but when *no microstructure noise* occurred. They are based on the minimum (named $MinRV_{t,M}$) and median (named $MedRV_{t,M}$), respectively of a number of consecutive absolute intraday returns, as follows:

$$MinRV_{t,M} = \frac{\pi}{\pi - 2} \cdot \frac{M}{M-1} \cdot \sum_{i=1}^{M-1} \left[\min\left\{\left|r_{t,i}\right|, \left|r_{t,(i+1)}\right|\right\}\right]^2, \tag{4.1a}$$

$$MedRV_{t,M} = \frac{\pi}{6 - 4\sqrt{3} + \pi} \cdot \frac{M}{M-2} \cdot \sum_{i=2}^{M-1} \left[\mathrm{med}\{|r_{t,(i-1)}|, |r_{t,i}|, |r_{t,(i+1)}|\}\right]^2. \qquad (4.1b)$$

The robustness of the $MinRV_{t,M}$ and $MedRV_{t,M}$ estimators stems from the fact that returns influenced by a large jump in price series are eliminated by the min or med functions. That is, if a significant jump occurs within one of the two return terms for the (4.1a) estimator, the *min* function squares the adjacent diffusive return.

4.2. Estimation techniques with noise induced

For the case when *no jumps* occurred, but *microstructure noise* can be *identified*, Barndorff-Nielsen *et al.* (2008) proposed a class of consistent kernel based estimators, *realized kernels*. This method extends the realized volatility literature, which has been shown –empirically– to significantly improve the ability to forecast volatility. The realized kernel estimator is defined by:

$$KRV_{t,M} = \gamma_{t,0}(h) + \sum_{i=1}^{H} k\left(\frac{i-1}{H}\right)\{\gamma_{t,i}(h) - \gamma_{t,-i}(h)\}, \qquad (4.2a)$$

$$\gamma_{t,i}(h) = \sum_{j=1}^{M} \left(s_{t-1+jh} - s_{t-1+(j-1)h}\right) \cdot \left(s_{t-1+(j-i)h} - s_{t-1+(j-i-1)h}\right), \qquad (4.2b)$$

where $\gamma_{t,i}(h)$ denotes the $i-th$ *realized auto-covariance function*, H denotes a parameter that controls the *bandwidth*, and $i \in \{-H, \dots, -1, 0, 1, \dots, H\}$. In formulae (4.2a), $k(x)$ denotes the *kernel function*, *i.e.*, a deterministic weight function depending on a bandwidth H, with $x \in [0,1]$. If $k(0) = 1$, $k(1) = 0$ and $H = cM^{2/3}$, the resulting estimator is asymptotically mixed Gaussian and converges at rate $M^{1/6}$. Here, the constant c can be optimally chosen as a function of the kernel and the integrated quarticity, such that the asymptotic variance of the estimator is minimized.

In the following, we propose to briefly present a more elaborated procedure used for estimating IV, *under* specific *jump* and *noise assumptions*. The main idea is referring to *averaging* and *sub-sampling*. The sub-sampling methodology, originally proposed by Zhang *et al.* (2005), is based on the idea of averaging over various realized variances assembled by sampling sparsely over high-frequency sub-samples.

For example, let us suppose that the intraday observations are assigned to K sub-samples. Using a systematic allocation, 5min. returns can be sampled at the time moments 10:40, 10:45, 10:50,…, respectively at the time moments 10:41, 10:46, 10:51 and so on.

After partitioning the whole sample into Ksub-samples, let us denote by $\overline{RV}_{t,K}$the *average$K$sub-sampled realized variance* as:

$$\overline{RV}_{t,K} := \frac{1}{K} \sum_{i=1}^{M-K+1} \left[s_{t-1+(i+K)h} - s_{t-1+ih} \right]. \tag{4.3}$$

The $\overline{RV}_{t,K}$ estimator is still biased, but this property depends now on the average size of the sub-samples rather than on the entire volume of observations. Realized variance, constructed from all observations, as usual denoted by$\widetilde{RV}_{t,M}$ and defined similar as in relation (2.3) for returns introduced by (3.3), is used for bias correction, yielding the *two time scales estimator*:

$$TSRV_{t,(M,K)} := \left(1 - \frac{\bar{M}}{M}\right)^{-1} \left[\overline{RV}_{t,K} - \frac{\bar{M}}{M}\widetilde{RV}_{t,M}\right], \ \text{with}\bar{M} := \frac{M-K+1}{K}. \tag{4.4}$$

Taking the difference $\overline{RV}_{t,K} - \frac{\bar{M}}{M}\widetilde{RV}_{t,M}$cancels the effect of the microstructure noise.

Zhang *et al*. (2005) proved that if no jumps occur, $TSRV$is a consistent estimator for the daily IV, when spot prices are contaminated by noise. Under the hypothesis subject to jumps occur, $TSRV$ estimates the IV plus the sum of squared intraday jumps.

Boudt and Zang (2010) proposed a jump robust version of two time scales estimator, named *jump robust$TSRV$*. It excludes from the estimator computation the returns that exceed a threshold of their distribution established under the hypothesis of no jumps. Under the semi-martingale model for log-prices (3.2), it follows:

$$\begin{aligned} s_{t-1+(i+K)h} - s_{t-1+ih} \\ &= \int_{t-1+ih}^{t-1+(i+K)h} \tilde{a}_u du + \int_{t-1+ih}^{t-1+(i+K)h} \tilde{\sigma}_u dB_u \\ &+ \varepsilon_{t-1+(i+K)h} - \varepsilon_{t-1+ih}. \end{aligned} \tag{4.5}$$

If $\tilde{a}_u = 0, \mathbb{P} - a.s.$, then:

$$\zeta_i := \frac{s_{t-1+(i+K)h} - s_{t-1+ih}}{\left\{\int_{t-1+ih}^{t-1+(i+K)h} \tilde{\sigma}_u^2 \, du + 2\omega^2\right\}^{1/2}} \longrightarrow \mathcal{N}(0,1). \tag{4.6}$$

Having in mind this, let us define the following indicator function:

$$\mathbb{I}_K(i,\xi) = \begin{cases} 1, & \text{if } \zeta_i \le \xi \\ 0, & \text{otherwise} \end{cases}. \tag{4.7}$$

Applying the truncation in both components of $TSRV$ (see (4.4)), we obtain:

$$\overline{\mathcal{RV}}_{t,K} := \frac{c_\xi}{K} \sum_{i=1}^{M-K-1} \left[s_{t-1+(i+K)h} - s_{t-1+ih}\right]^2 \cdot \mathbb{I}_K(i,\xi), \tag{4.8a}$$

$$\widetilde{\mathcal{RV}}_{t,M} := c_\xi \sum_{i=1}^{M} \left[s_{t-1+(i+1)h} - s_{t-1+ih}\right]^2 \cdot \mathbb{I}_1(i,\xi), \tag{4.8b}$$

where the constant $c_\xi^{-1} := F_{\chi_3^2}(\xi)$ adjusts the bias due to the thresholding, and $F_{\chi_N^2}(\cdot)$is the cumulative distribution function of chi-squared distribution with N degrees of freedom.One way to improve the stability of the estimator proposed bellow (see (4.10)) is to adjust it for the sub-sample of data we truncated in. In this order, let us introducing:

$$\overline{\mathcal{RV}}_{t,K}^* := \frac{c_\xi^*}{K} \cdot \frac{\sum_{i=1}^{M-K-1}\left[s_{t-1+(i+K)h} - s_{t-1+ih}\right]^2 \cdot \mathbb{I}_K(i,\xi)}{\frac{1}{M-K-1}\sum_{i=1}^{M-K-1} \mathbb{I}_K(i,\xi)}, \tag{4.9a}$$

$$\widetilde{\mathcal{RV}}_{t,M}^* := c_\xi^* \cdot \frac{\sum_{i=1}^{M}\left[s_{t-1+(i+1)h} - s_{t-1+ih}\right]^2 \cdot \mathbb{I}_1(i,\xi)}{\frac{1}{M}\sum_{i=1}^{M} \mathbb{I}_1(i,\xi)}, \tag{4.9b}$$

where $c_\xi^* = F_{\chi_1^2}(\xi) / F_{\chi_3^2}(\xi)$. For example, if $\xi = 9$, then $c_\xi^* = 1.027$.

Taking all above into account, the *jump robust two time scales estimator* (abbreviated by $JRTSRV$) is defined by:

$$JRTSRV_{t,(M,K)} := \left(1 - \frac{\overline{M}}{M}\right)^{-1} \left[\overline{\mathcal{RV}}_{t,K}^* - \frac{\overline{M}}{M}\widetilde{\mathcal{RV}}_{t,M}^*\right], \tag{4.10}$$

with$\overline{M}$ introduced by (4.4).

In order to compute the indicator function $\mathbb{I}_K(i,\xi)$, we need to estimate $\int_{t-1+ih}^{t-1+(i+K)h} \tilde{\sigma}_u^2 \, du$ and the variance of noise. For the case when no jumps occur, Zhang *et al.* (2005) proved that:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{2n} \cdot \left[\widetilde{RV}_{t,M} - TSRV\right] \xrightarrow{\mathbb{P}} \sigma_\varepsilon^2. \tag{4.11}$$

This result also remains valid under the hypothesis of jumps, given that the terms $\widetilde{RV}_{t,M}$ and $TSRV$ are equally influenced by jumps. For the estimation of the integral term, for relatively small values of K (e.g., $K \leq 300$for $M = 23\,400$), we can use the approximation $\int_{t-1+ih}^{t-1+(i+K)h} \tilde{\sigma}_u^2 \, du \approx \frac{Kh}{t} \cdot \widehat{\int_0^t \tilde{\sigma}_u^2 du}$. For estimating the quantity$\int_0^t \tilde{\sigma}_u^2 du$, an iterative approach can be used.

5. Data and results

In order to perform a comparison among the most known high-frequency estimators, we have selected three blue chips listed on Bucharest Stock Exchange (BSE):Banca Română pentru Dezvoltare (BRD), Fondul Proprietatea (FP) and OMV Petrom (SNP) during September 19, 2012 and April 30, 2013. The data has been provided by *Thompson Reuters Eikon* platform, consisting of prices recorded tick by tick. Descriptive statistics about trades and time interval between trades are presented in Table1.

Table 1: Statistics of trades

Stocks	*Average Time Interval (minutes)*	*Average Trades per day*	*MeanBid*	*MeanAsk*
BRD	**10.6**	**36**	**8.3100**	**8.3600**
FP	**4.38**	**92**	**0.5882**	**0.5892**
SNP	**6.93**	**57**	**0.4345**	**0.4361**

According to Table 1, FP has been the most traded stock, having an average of 92 trades *per* day, followed by SNP with an average of 57 trades *per* day, and BRD with an average of 36 trades *per* day. The ratio $spread/mid\ qoute$was 0.60% for BRD, 0.17% for FP and 0.37% for SNP. Therefore, we may conclude that the $spread/mid\ qoute$ ratio, as an appraisal for trading costs, is strongly positively related to its liquidity.

5.1. Jump detection in price series

In order to select the best high-frequency estimator for intra-day price series, we need to study, first of all, the presence of jumps. In this respect, we perform two tests (AJ-test and BS-test, presented in section 2) for the presence of jumps, each day in the sample, at different time intervals. Once all estimations have been compiled, we have selected the days for which the null hypothesis (of no jumps) was accepted and we have computed the proportion of these days in the total sample. The results are presented in Table 2.

Table2: Jump tests at different frequencies (percentage of days)

H(0): *No jumps*		5 minute aggregation	10 minutes aggregation	15 minutes aggregation	20 minutes aggregation
		Accept H(0) at 5% level			
BRD	AJ Test	**15.38%**	**11.54%**	**8.46%**	**6.92%**
	BS Test	**33.33%**	**44.19%**	**50.39%**	**63.56%**
FP	AJ Test	**2.27%**	**1.52%**	**4.55%**	**0.00%**
	BS Test	**67.18%**	**74.05%**	**77.10%**	**74.81%**
SNP	AJ Test	**10.69%**	**4.58%**	**3.82%**	**5.34%**
	BS Test	**50.39%**	**61.24%**	**70.54%**	**66.67%**

According to Table 2, some major differences between this two jump tests may be observed. AJ – test returns an average percentage of 10.58% for BRD, 2.09% for FP and 6.11% for SNP. In the case of BS – test, the average results are different: no jumps occurred in 47.87% of days for BRD price series, 73.29% for FP and 62.21% for SNP.

An interesting aspect has been relieved from the results: AJ – test returns an average proportion of days with no jumps with a lower value if we have a liquid market, such as SNP’s and a bigger one in a less liquid market case, such as BRD’s. On the other side, BS – test returns an average proportion of days with no jumps with a higher value if we have a liquid market, such as SNP’s and a lower one in a less liquid market case, such as BRD’s. So, we may conclude that the absence of jumps undervalues the risk in the case of a liquid market if we use AJ-test and overvalues it in the same market if we use BS-test. In the next sub-sections, we are going to use the results returned by BS – test, considering them reliable on account of our data set architecture. Our first choice would be probably AJ – test if the ticks were considerably more frequently recorded. We accept null hypothesis for values of over 50%, rejecting it otherwise. According to our previous results we cannot reject the null hypothesis for FP and SNP, but we do reject it for BRD.

5.2. Microstructure noise

Microstructure noise is a deviation from fundamental value of an asset that is induced by the characteristics of the market under considerations like bid-ask bounce, the discreteness of price change, latency, and traders with different degree of information. If microstructure noises are presented in some data series of prices some estimation techniques of realized variance, based on high-frequency data, would be unstable. One way to identify microstructure noises is to aggregate the data into bigger time intervals. If the microstructure noises are present, results would be quite different.

In Figures 1–3 are presented the realized variance measures for all three stocks, at ten, fifty and twenty minute aggregation, respectively. Big differences among these plots are visible in the case of BRD and partially for SNP.

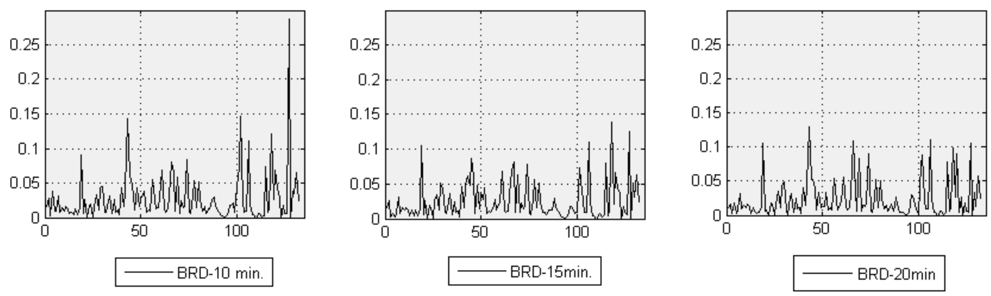

Figure 1. Daily realized variance at different frequencies – BRD (percentage)

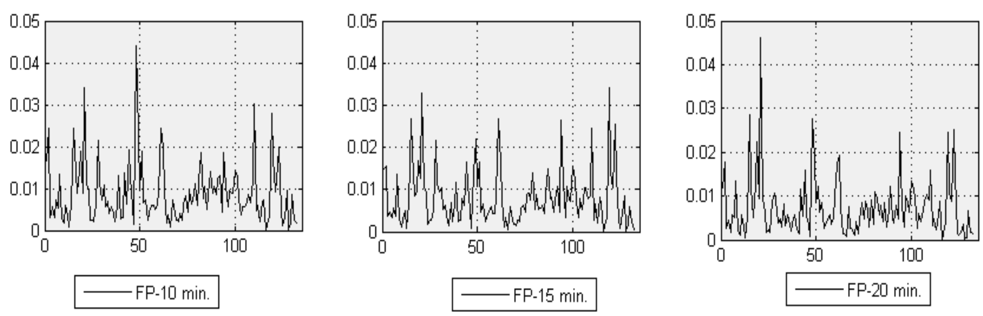

Figure 2. Daily realized variance at different frequencies – FP (percentage)

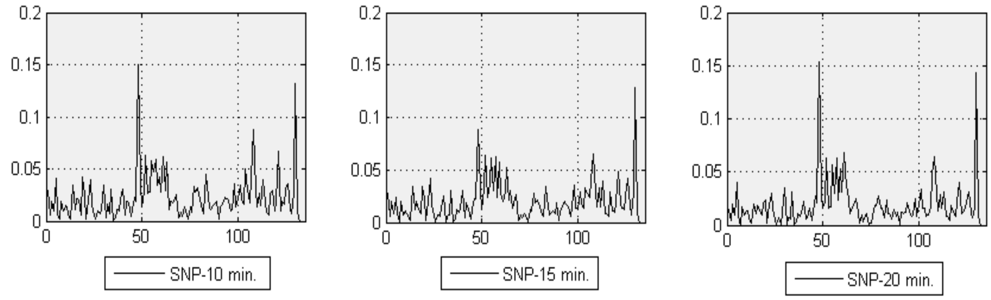

Figure 3. Daily realized variance at different frequencies – SNP (percentage)

In Figure 4 we plot average realized variance over different time intervals. In the case of BRD, the nature of data didn't allow a higher level of aggregation, so we select five levels of aggregation. For FP and SNP we choose ten levels of aggregation.

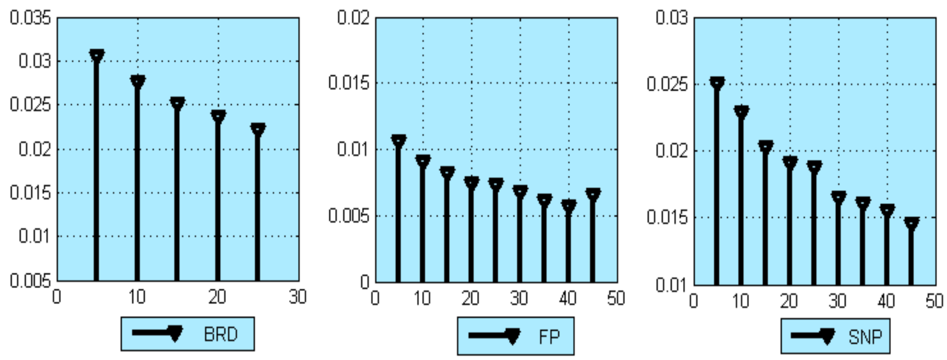

Figure 4.Variance signature plot for realized variance (percentage)

The presence of microstructure noises is very visible for BRD, since it cannot be identified a stabilization level, so if one must estimate a model with intra-day data of BRD, the presence of microstructure noise shall be considered. For FP, the absence of microstructures noise is obvious. We may see a stabilization of realized variance for thirty minutes aggregation interval, although the differences in values are very small. Thus, we may eliminate the hypothesis of microstructure noises for FP. We shall not have the same approach for SNP, even for a realized variance stabilization ascertained by thirty minutes aggregation interval. The differences among these estimators are prominent, almost as significant as BRD's realized variance value.

5.3. Volatility estimators

In this sub-section we are going to compute estimation for different high-frequency estimators based on equities prices from BSE. From now, we accepted the hypothesis of jumps for BRD and reject it for SNP and FP. In the case of microstructure noises, we found that only price series that is not affected substantially by microstructure noises effect is only FP share.

We selected 5 minutes and 10 minutes interval aggregation, mainly to gain the information contented in price series. If the prices were recorded at millisecond, an optimal aggregation interval would be, for example, 1 minute. In our case, considering a small number of trades per day, a higher level aggregation for volatility estimators is possible to undervalue the measure of risk.

Table 3.Variance estimators for BRD, FP and SNP using high-frequency data

Stocks	***Microstructure noise: No***			
	Jumps: No		***Jumps: Yes***	
	Estimator: $\boldsymbol{RV}$		Estimator: $\boldsymbol{MinRV}$	
	5 min. aggregation	10 min. aggregation	5 min. aggregation	10 min. aggregation
BRD	**0.01434%**	**0.01517%**	**0.01149%**	**0.01406%**
FP	**0.00784%**	**0.00708%**	**0.00691%**	**0.00649%**
SNP	**0.01145%**	**0.01541%**	**0.00908%**	**0.01308%**
Stocks	***Microstructure noise: Yes***			
	Jumps: No		***Jumps: Yes***	
	Estimator: $\boldsymbol{KRV}$		Estimator: $\boldsymbol{JRTSRV}$	
	5 min. aggregation	10 min. aggregation	5 min. aggregation	10 min. aggregation
BRD	**0.02339%**	**0.02474%**	**0.0171%**	**0.0191%**
FP	**0.00786%**	**0.00773%**	**0.0061%**	**0.0060%**
SNP	**0.01920%**	**0.01585%**	**0.0154%**	**0.0178%**

In Table 3, a comparison between high-frequency estimates is presented. Some conclusions can be drawn:

(a) In the case of BRD, presence of both noise and jumps must be considered. In this case, estimated variance is 0.0171% for five minutes aggregation and 0.0191% for ten minutes aggregation. If we eliminate jump restriction, we overestimate the risk. If we eliminate the microstructure noises restriction, risk is underestimated. If we eliminate both, risk is underestimated also.

(b) In the case of FP we have no noises or jumps. In this case estimated variance is 0.00784% for five minutes aggregation and 0.00708% for ten minutes aggregation.

Considering presence of microstructure noises for FP price series wouldn't be a big mistake, having in mind the variance signature plot. In his case the results are the same, with small differences. If we are considering jumps for the price series of FP, the risk is underestimated.

(c) In the case of SNP we have microstructures noises but there is no jump. In this case estimated variance is 0.01920% for five minutes aggregation and 0.01585% for ten minutes aggregation. If we accept the hypothesis of jumps, at five minutes aggregation, the risk is undervalued and at ten minutes aggregation, the risk is undervalued. In the case of SNP, having these contradictions based on interval aggregation, we must accept the fact that microstructures noises have a greater impact that we observed and a higher level of aggregation is recommended.

In Table 4 we present the measure of risk in annualized terms, for the recommended estimators corresponding to five minutes aggregation data and ten minutes aggregation data.

Table 4. Annualized risk measures for BRD, FP and SNP

$Stocks$	$JRTSRV$		RV		KRV	
	5 min. aggregation	10 min. aggregation	5 min. aggregation	10 min. aggregation	5 min. aggregation	10 min. aggregation
BRD	**24.28%**	**24.97%**	–	–	–	–
FP	–	–	**14.06%**	**13.36%**	–	–
SNP	–	–	–	–	**22.00%**	**19.99%**

In general, even if there are no significant difference between average risk at five minutes aggregation and ten minutes aggregation, considering that a trade for any equity is taking place once at ten minutes, we recommend this aggregation time when making assumptions about (daily) volatility.

5.4. Results interpretation

BRD has manifested more daily volatility ($1.3856\% = \sqrt{0.0191\%}$) and an underperformance related to peers (*e.g.*, Erste Bank) and market (BET-C and BET-FI) due to many reasons:***(i)***the European financial sector has suffered most in 2012 due to the European sovereign crisis increasing CDS swaps and loans portfolios deterioration;***(ii)***the Romanian financial sector has suffered most since 2009 because it was the moment when most of the underperforming loans had started to be provisioned, causing a loss for BRD of over 330 million RON by the end of 2012. The loss was in line with the general trend of the Romanian banking sector;***(iii)***BRD situation was even harsher due to predatory lending. By the end of October 2012,competent Romanian authorities have made public the results of a fraud, causing significant losses for the bank, the day of November 16, 2012 being the most volatile day (26%*per* year);(***iv)***moreover, BRD has not beneficiated of favorable price support from investors and insiders, SIF Muntenia being a major insider selling many stocks. Without favorable support from insiders or investment funds (like FP and SNP had), BRD stock prices had a less favorable trend than market, sector and peers.

FP has had the best evolution of all companies analyzed, manifesting the least daily volatility ($0.8414\% = \sqrt{0.00708\%}$) and over-performance the market (BET-C, BET-FI), due to some factors briefly explained below: ***(i)*** a very strong price support from insiders and issuer. One of Elliot's funds (Manchester Securities Corp.)has been a strong bidder during November 2012 and April 2013. Moreover, starting with April 10, 2013, FP has started the second program of share repurchases, being a strong bidder as well. Their bidding support has smoothed significantly the FP's price during the period analyzed, making it higher even in periods of overall market downside trend (April 2013);***(ii)*** another reason for believing the major catalyst role of Elliot's fund and FP in smoothening the price evolution between September 2012 and middle December 2012 was the instance when Elliot's fund presence in the market was minimal. Moreover, January effect for FP was less strong than the rest of the market (first period of January), the evolution of the price being stronger only in March (contrary to the market) and April 2013 regardless of consistent positive financial results for the period analyzed. January 4th, 2013 has been the day with the highest high-frequency annualized volatility estimated (30% *per* year). This value maybe related to the effect of the first trading day of the year.

SNP has manifested less daily volatility ($1.2589\% = \sqrt{0.01585\%}$) than BRD and mixed performance related to market and peers (BET-C, respectively OMV) due to many reasons: ***(i)*** upward (positive) trend for BRENT and WTI oil prices during the period analyzed has been a strong support for positive and consistent financial results; ***(ii)*** strong insiders price support due to committed stock acquisitions of Templeton Frontier Markets Fund. During November 2012 and April 2013. By the end of 2014, OMV Petrom has been the largest holding (4.61% percent of the 1 billion US total assets fund). This represents probably the main reason why SNP has strong support and over performance for most of the time related to BET-C and OMV; ***(iii)*** SNP had three main downtrend periods underperforming BET-C: first half of November (Q3 results underestimates), one month period starting the half of January 2013 (2012 financial results under estimates although positive and higher turnover and profit YOY) and another volatile period starting with 21 March. This last very volatile (50% *per* year volatility estimate in March 27, 2013)period may be the result of 10% lower dividend proposal comparing to the year before (0.0280 RON for 2012 comparing to 0.0310 RON for 2011), corroborated with leaving the office of a member of the Board in charge with Exploration and Production. The dividend proposal has been a surprise since the main financial indicators were all positive and higher comparing to the year before (turnover +18% YOY and profit +3% YOY).

6. Conclusions

We have reviewed different ways to estimate the integrated variation based on high frequency data and used them for estimating the daily volatility of three representative stocks traded on BSE, respectively: BRD, FP and SNP. Above all, the presence and significant influence of both jumps and microstructure noise has been also empirically studied. We proposed adequate high-frequency variance estimators.

In the period analyzed, SNP and FP, except for BRD, had an overall positive evolution, with a culminating favorable (upside) trend between middle December 2012 to middle January 2013, confirming the theory of behavioral finance – *January effect bias*. Except for FP, both SNP and BRD have confirmed the overall trend of the BSE market (BET-C and BET-FI index evolution), BRD underperforming the market and SNP over performing it although the general trend being confirmed for the most period. For most of the time we confirm the inverse relationship between stock market evolution and EURRON exchange rate.

Acknowledgement

We wish to express our heartfelt appreciation and deep sense of gratitude to Professor Bogdan Negrea, PhD(The Bucharest University of Economic Studies)and to Professor VasilePreda, PhD (University of Bucharest) for their constantly encouraging support and for being a great source of inspiration and knowledge. Also, we thank to Ciprian Turcan for sharing his experience in preparation of this manuscript.

REFERENCES

[1] **Aït- Sahalia, Y.; Jacod, J. (2009),*Testing for Jumps in a Discretely Observed Process*.***The Annals of Statistics*, 37(1), 184-222;

[2] **Aït-Sahalia Y.; Mykland P.; Zhang, L. (2005), *How often to Sample a Continuous-time Process in the Presence of Market Microstructure Noise*.***The Review of Financial Studies*, 18, 351-416;

[3] **Andersen, T.G.; Bollerslev, T. (1998), *Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts*.** *International Economic Review*, 39(4), 885-905;

[4]**Andersen, T.G.; Dobrev, D.; Schaumburg, E. (2012),** ***Jump-robust Volatility Estimation Using Nearest Neighbor Truncation***. *Journal of Econometrics*, 169(1), 75-93;
[5]**Barndorff-Nielsen, O. E.; Shephard, N. (2004),** ***Power and Bi-power Variation with Stochastic Volatility and Jumps***(with discussion). *Journal of Financial Econometrics*, 2, 1-48;
[6] **Barndorff-Nielsen, O. E.; Shephard, N. (2006),** ***Econometrics of Testing for Jumps in Financial Economics Using Bi-power Variation***.*Journal of Financial Econometrics*, 4(1), 1-30;
[7] **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(6), 1481-1536;
[8] **Bauwens, L.;Hafner, C.; Laurent, S. (2012),*****Volatility Models and their Applications***. New Jersey: *John Wiley & Sons*, Inc.;
[9]**Boudt, K.; Zhang, J. (2013),** ***Jump Robust two Time Scale Covariance Estimation and Realized Volatility Budgets***. *Quantitative Finance*, DOI: 10.1080/14697688.2012.741692;
[10] **Delbaen, F.; Schachermayer, W. (1994),** ***A General Version of the Fundamental Theorem of Asset Pricing***. *Mathematische Annalen*, 300, 463-520;
[11] **Hansen, P.R.; Lunde, A. (2006),** ***Realized Variance and Market Microstructure Noise (with discussion)*.***Journal of Business and Economic Statistics*, 24, 127-218;
[12] **Hautsch, N. (2012),*****Econometrics of Financial High-Frequency Data***. Berlin: *Springer-Verlag*;
[13] **Jacod J.; Li, Y.; Mykland P.; Podolskij M.; Vetter M. (2009),** ***Microstructure Noise in the Continuous Case: The Pre-averaging Approach. Stochastic Processes and their Applications***, 119(7), 2249-2276;
[14]**Pigorsch, C.; Pigorsch, U.; Popov, I. (2012),** ***Volatility Estimation Based on High Frequency Data.*** In J.C. Duan, J.E. Gentle, W.K. Härdle (Eds.), *Handbook of Computational Finance*, *Springer*, 335-369;
[15] **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.