CONFERENCE PROCEEDINGS
FULL TEXT PAPERS

edited by
Zofia Rusnak and Beata Zmyślona

Contents

SOME METHOD OF DETECTING THE JUMP CLUSTERING PHENOMENON IN FINANCIAL TIME SERIES

MACIEJ KOSTRZEWSKI

Cracow University of Economics, Department of Econometrics and Operations Research,
27 Rakowicka Street, Cracow, Poland
email: maciej.kostrzewski@uek.krakow.pl

Abstract

News might trigger jump arrivals in financial time series. The "bad" and "good" news seems to have a distinct impact. In the study, the Bayesian double exponential jump-diffusion model and the jump-diffusion model with M jumps are applied to model logarithmic rates of return and detect jumps. The technique proposed in the paper enables the analysis of jump frequency. The methodology, founded upon the idea of latent variables and Markov Chain Monte Carlo methods, is illustrated with two empirical studies. Periods of no jumps alternating with the ones of frequent jumps confirm the existence of jump clustering.

Key words: *double exponential jump-diffusion model, JD(M)J model, Bernoulli jump-diffusion model, Bayesian inference, MCMC methods, latent variables, jump clustering.*

1. Introduction

Jumps in the time series of asset returns are often defined as the values exceeding some arbitrarily chosen thresholds. Obviously, different thresholds lead to a different number of jumps. Thresholds are commonly set symmetrically either around zero or the sample mean and are defined as a multiple of the sample standard deviation. Quite often the empirical distribution of logarithmic rates of return features negative skewness, in the case of which symmetric thresholds seem no longer valid. In what follows, the latent variables are used to identify the data points which feature a jump. Specifically, it is assumed that a jump occurs at the i-th moment if the posterior probability of a jump exceeds an arbitrarily chosen value corresponding to the aforementioned thresholds. However, the problem of asymmetry or symmetry is not a matter here. In the research we are preoccupied with inference about jumps (in particular detecting them and analyzing the frequency of jumps) rather than relating them with, e.g., macroeconomic releases, with the caveat, however, that we make no pretense of settling how good or how bad the models in question perform in terms of either the in- or the out-of-sample fit.

The presented methodology enables one to identify (in probabilistic terms) the moments when a jump has occurred. Having that it is further of a particular interest to examine whether the jump activity is variable over time and whether jumps tend to cluster over time. It is worth noting that even though the models under consideration – similarly as some other common specifications – do not account for any dependence structure in the occurrence of jumps, it is still informative (in the context of detecting jump clusters) to inspect the series of time elapsed between consecutive jumps, for it still can exhibit patterns suggestive of clustering. The analysis is performed within two Bayesian models: the double exponential jump diffusion process (DEJD model) and the jump-diffusion model with M jumps (JD(M)J model) developed by (Kostrzewski, 2012, 2013a, 2013b).

The very term "jump clustering" – quite analogous to the one of "volatility clustering", pervading the GARCH and SV literature – means that jump arrivals (or times between two consecutive jumps) tend to cluster. Jump clusters have been already discussed in the financial econometrics literature – see, e.g., (Maheu and McCurdy, 2004), (Yu, 2004), (Knight and Satchell, 2007), (Lee and Mykland, 2008), (Lee, 2012). The main idea is based on the assumption of a stochastic jump intensity which follows an autoregressive structure or a self-exciting process. Moreover, (Lee, 2012) investigated the jump size dynamics and showed evidence of a short-term jump size clustering.

The contribution of the paper resides in performing an analysis of jump frequency and designing some simple method of detecting the jump clustering phenomenon. The main idea of the methodology is based on three steps:

1. Identify times of jumps by means of any Bayesian jump diffusion model.

2. Inspect the frequency of jumps by plots of the posterior probabilities of jumps, the series of waiting times between successive jumps and the plot of the autocorrelation function of these times.

3. Make a decision, based on the plots, about the jump clustering.

The use of the proposed methodology is particularly justified in the context of settling whether the structure of some common jump-diffusion models should be extended so as to take the jump clustering phenomenon explicitly into account.

2. The models

Consider a standard Wiener process $W = (W_t)_{t \geq 0}$, a Poisson process $N = (N_t)_{t \geq 0}$ with the intensity parameter $\lambda > 0$, and a sequence of independent random variables $Q = (Q_j)_{j \geq 1}$. Let us assume that W, N and Q are mutually independent. Finally, $S = (S_t)_{t \geq 0}$ denotes the price process of some risky asset.

The logarithm of S is governed by a jump-diffusion process that constitutes the solution of the equation:

$$d(\ln S_t) = \left(\mu - \frac{1}{2}\sigma^2 \right)dt + \sigma dW_t + QdN_t. \tag{1}$$

It might be shown that:

$$\ln\left(\frac{S_{t+\Delta}}{S_t} \right) = \left(\mu - \frac{1}{2}\sigma^2 \right)\Delta + \sigma(W_{t+\Delta} - W_t) + \sum_{i=N_t+1}^{N_{t+\Delta}} Q_i, \quad \Delta > 0. \tag{2}$$

The process is built of two components: the (pure) diffusion part,

$$\left(\mu - \frac{1}{2}\sigma^2 \right)\Delta + \sigma(W_{t+\Delta} - W_t), \tag{3}$$

representing continuous variations in the series, and the (pure) jump component, $\sum_{i=N_t+1}^{N_{t+\Delta}} Q_i$,

reflecting abnormal (extreme) movements in returns. The continuous price behavior between jumps is described by the geometric Brownian motion, W, while the arrival rate of jumps is described by the homogeneous Poisson process, N, and the jump magnitudes – by Q.

The distribution of logarithmic rates of return, $\ln\left(\frac{S_{t+\Delta}}{S_t}\right)$, is an infinite mixture (Hanson, Westman 2002):

$$\sum_{k=0}^{\infty} \exp(-\lambda\Delta) \frac{(\lambda\Delta)^k}{k!} f_k(x), \tag{4}$$

where f_k are some densities. Since the series given by (4) is infinite, the density is intractable. Therefore, consider the following finite approximation of (4):

$$\sum_{k=0}^{\infty} \exp(-\lambda\Delta) \frac{(\lambda\Delta)^k}{k!} f_k \approx \sum_{k=0}^{M} \exp(-\lambda\Delta) \frac{(\lambda\Delta)^k}{k!} f_k. \tag{5}$$

The approximation restricts the number of jumps over any time interval Δ to M. The case of $M = 0$ indicates no jumps over interval Δ.

Further considerations are restricted to the discrete time framework. Time series $(x_1, x_2, ..., x_n)$ is comprised of $x_i = \ln\left(\frac{S_{t_{i+1}}}{S_{t_i}}\right)$ observed at times $t_1, t_2, ...$. Moreover, $\Delta \equiv t_{i+1} - t_i > 0$ is a fixed time interval between following observations. Denote the vector of parameters as θ. If we normalize the approximation given by (5), under $M = 1$ we obtain the conditional data density (given the parameters, θ):

$$p(x|\theta) = \frac{1}{1+\lambda\Delta} f_X(x) + \frac{\lambda\Delta}{1+\lambda\Delta} f_{X+Q_j}(x), \tag{6}$$

where $X := \left(\mu - \frac{1}{2}\sigma^2\right)\Delta + \sigma\Delta W_t$, and $Q_j \sim f_{Q_j}$. The first term on the right-hand side of (6) is referred to as the diffusion component, whereas the second one − as the jump-diffusion component.

In what follows, only two models are considered. In the first one, further referred to as the DEJD model, logarithmic rates of return are assumed to follow the distribution given by (6) while Q_j has a double exponential distribution with density

$$f_{Q_j}(x) = p_D \eta_D \exp(\eta_D x) \mathsf{I}_{(-\infty,0)}(x) + p_U \eta_U \exp(-\eta_U x) \mathsf{I}_{[0,\infty)}(x), \tag{7}$$

where $\eta_U > 0, \eta_D > 0$. The second specification, termed the JD(2)J model, is defined by assuming a normal distribution for Q_j: $f_{Q_j}(x) = \frac{1}{\sqrt{2\pi}\sigma_Q} \exp\left(-\frac{1}{2}\left(\frac{x-\mu_Q}{\sigma_Q}\right)^2\right)$, and setting $M = 2$. The choice of $M = 2$, by means of Bayes rules, is presented in (Kostrzewski, 2013a). Both structures are used to model series of daily logarithmic rates of return, indicating $\Delta = \frac{1}{252}$.

3. The Bayesian models

A Bayesian statistical model is defined by the joint density:

$$p(x,\theta) = p(x|\theta)p(\theta), \tag{8}$$

where $x = (x_1, ..., x_n)$ is the observed data, θ is the vector of unknown parameters, $p(x|\theta)$ is the sampling density and $p(\theta)$ is the prior density. The inference rests upon the posterior

density $p(\theta|x)$ of θ given data x (Bernardo and Smith, 2002):

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}. \qquad (9)$$

Given x, $p(x|\theta)$ – as a function of θ – is called the likelihood function, whereas $p(x)$ is the marginal data density.

Under the DEJD specification the process S depends on six unknown parameters $\theta = (\mu, \sigma, \lambda, p_U, \eta_U, \eta_D)$, where $\theta \in \mathsf{R} \times (0, \infty) \times (0, \infty) \times (0,1) \times (0, \infty) \times (0, \infty)$. While analyzing a time series which is (or, rather, is believed to be) a trajectory of a jump-diffusion process, one does not actually know if a given data point has been generated by the pure diffusion or the jump-diffusion component. In other words, one cannot determine which component of the series in (6), i.e. $f_X(x)$ or $f_{X+Q}(x)$ underlies a given observation. To manage the problem let us introduce latent variables $\xi = (\xi_1, ..., \xi_n)$, where $\xi_i \in \{-1, 0, 1\}$. The value $\xi_i = 0$ means no jump at (an interval) $t = i\Delta$. The values $\xi_i = -1$ and $\xi_i = 1$ mean that a jump occurs and its value is either negative or positive, respectively. Moreover, it is convenient to introduce latent variables $J = (J_1, ..., J_n)$ corresponding to the value of a jump, where:

$$p(J_i = j|\theta, \xi_i = -1) = \eta_D \exp(\eta_D j)\mathsf{I}_{(-\infty, 0)}(j),$$
$$p(J_i = j|\theta, \xi_i = 0) = \delta_0(j), \qquad (10)$$
$$p(J_i = j|\theta, \xi_i = 1) = \eta_U \exp(-\eta_U j)\mathsf{I}_{(0, \infty)}(j),$$

and δ_0 is the Dirac delta function.

The Bayesian model is then given by:

$$p(x, \theta, \xi, J) = p(x|\theta, \xi, J)p(\theta, \xi, J). \qquad (11)$$

More information about the Bayesian DEJD model could be find in (Kostrzewski, 2013b).

Under the JD(2)J specification the process S depends on five unknown parameters $\theta = (\mu, \sigma, \lambda, \mu_Q, \sigma_Q^2)$, where $\theta \in \mathsf{R} \times (0, \infty) \times (0, \infty) \times (0,1) \times (0, \infty) \times (0, \infty)$. The idea of the Bayesian JD(M)J specification is similar to the DEJD one. More information about the Bayesian JD(2)J model can be find in (Kostrzewski, 2012, 2013a).

Posterior characteristics of the unknown quantities are calculated via the Markov Chain Monte Carlo (MCMC) methods (Gamerman and Lopes, 2006), combining the Gibbs sampler, the independence and the sequential Metropolis-Hastings algorithms, as well as the acceptance-rejection sampling (Chib and Greenberg, 1995). The theorems stated in (Kostrzewski, 2013a, 2013b) make the algorithm ready to use. The prior structures introduced into the models in question are presented in (Kostrzewski, 2013a) and (Kostrzewski, 2013b) for the JD(M)J model and the DEJD model, respectively.

Formally, the occurrence of a jump is equivalent to $\xi_i \neq 0$. Unfortunately, one does not observe ξ_i, but the posterior probability of a jump, $P(\xi_i \neq 0|x)$, can be evaluated for each day $i = 1, ..., n$. Let us assume that a jump occurs at the i-th period if the probability $P(\xi_i \neq 0|x)$ exceeds an arbitrarily chosen value of 0.5. The resulting series consisting of zeros and ones corresponding to such i's that $P(\xi_i \neq 0|x) \leq 0.5$ and $P(\xi_i \neq 0|x) > 0.5$, respectively, is further

employed in assessing the frequency of jumps as well as studying the jump clustering phenomenon.

4. Examples

This section illustrates the methodology outlined above. Two real-world datasets are under study. First, the series of daily logarithmic rates of return on the ICE ECX futures based on the underlying European Union CO2 emission allowances is fit with the DEJD structure. Secondly, the dataset of logarithmic rates of return on the S&P100 Index is fit with the JD(2)J specification.

All relevant computations were performed in R (R Core Team, 2013). The numerical algorithms applied in the research require monitoring the convergence of the generated Markov chain to its limiting stationary distribution. Convergence of all MCMC samplers exploited in the research is confirmed by the visual inspection of ergodic means, standard deviations and CUMSUM statistics plots. The results are robust to the choice of the starting point for the MCMC procedure.

4.1 Analysis for the CO2 under DEJD model

The European Climate Exchange (ECX) manages Carbon Financial Instruments traded on the ICE Futures Europe electronic platform. The ICE ECX futures are based on the underlying European Union CO2 emission allowances (EUA) and traded on the platform. They are standardized products. The contract volume amounts to 1,000 CO2 EU Allowances. Each EU Allowance entitles one to emit one tonne of carbon dioxide equivalent gas. Future prices are quoted in Euro per metric tonne. Trading model is based on continuous trading.

Further analysis is restricted to closing prices of future contracts expiring on 16 December, 2013, over the period 3 January, 2011 through 1 October, 2013. The number of observations equals $n = 709$. Let us note that the period under study covers Phase II (2008-2012) and Phase III (2013-2020) of the European Union Emissions Trading Scheme (EU ETS). The EU ETS enables market participants to use the Phase II permits during Phase III.

Table 1 contains some basic posterior characteristics of the DEJD model's parameters. The presented results are based on $200,000$ MCMC draws, preceded by $30,000$ burn-in cycles.

Table 1 Posterior means ($E(\cdot \,|\, x)$) and standard deviations ($D(\cdot \,|\, x)$) for the ICE ECX futures

θ	$E(\cdot \,|\, x)$	$D(\cdot \,|\, x)$
μ	0.138	0.282
σ	0.360	0.032
λ	119.164	46.923
p_U	0.412	0.079
η_D	22.170	3.660
η_U	20.365	4.051

Source: Own elaboration.

Figure 1 displays the marginal posteriors of parameters in the DEJD model, along with the prior densities. The plots reveal a considerable contribution of the data to the shape of marginal posteriors (except for μ).
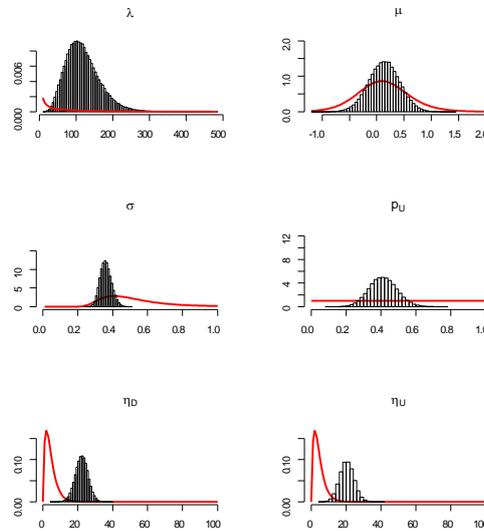
Figure 1 Marginal posterior (bars) and prior densities (solid line) of parameters in the DEJD model estimated for the ICE ECX futures.
Source: Own elaboration.

Figure 2 depicts the logarithmic rates of return (with a band of $\bar{x}_n - 2 \cdot \sigma_n$ and $\bar{x}_n + 2 \cdot \sigma_n$ where $\bar{x}_n$ and σ_n denote the sample mean and standard deviation, respectively), and the values of probabilities $P(\xi_i = -1|x)$ and $P(\xi_i = 1|x)$ against the number of successive days, provided that $P(\xi_i = -1|x) > 0.5$ and $P(\xi_i = 1|x) > 0.5$. Note that higher posterior probabilities of a jump go along with higher volatility of the time series.
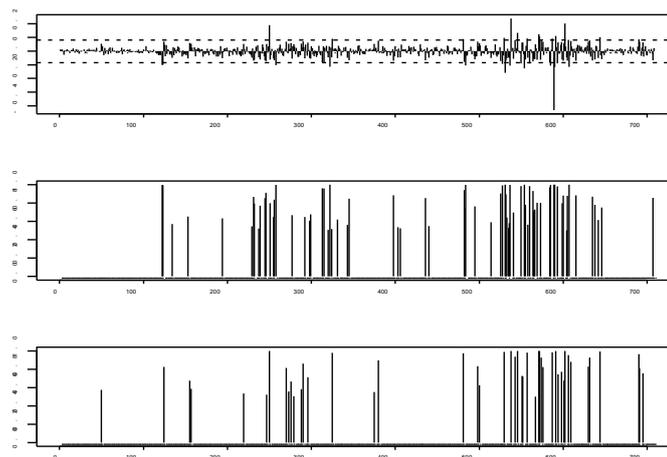
Figure 2 The modeled time series of daily log-returns (top), the series of
$\{P(\xi_i = -1|x) \mathrm{I}_{\{i \,:\, P(\xi_i = -1|x) > 0.5\}}(i); i = 1, 2, ..., n\}$ (middle), and the series of
$\{P(\xi_i = 1|x) \mathrm{I}_{\{i \,:\, P(\xi_i = 1|x) > 0.5\}}(i); i = 1, 2, ..., n\}$ (bottom).
Source: Own elaboration.

Figure 3 presents the posterior probabilities of a jump, i.e. $P(\xi_i \neq 0 | x)$ (only the ones that exceed 0.5), the values of the time elapsed between consecutive jumps (with the number of a jump represented by the horizontal axis) and the corresponding (empirical) autocorrelation function (ACF) calculated for the series of times between consecutive jumps.

One can easily observe periods of no jumps alternating with the ones of frequent jumps. The values of the ACF for the first lag equals 0.2 and appears statistically significant (at the significance level of 0.05). The results may be perceived as indicative of the jump clustering phenomenon.
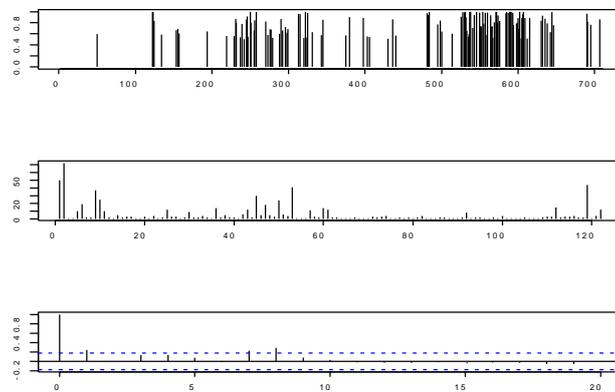
Figure 3 The time series of the posterior probabilities of a jump (only the ones exceeding 0.5; top), the series of times between successive jumps (middle), and the autocorrelation function of the times between consecutive jumps (bottom).

Source: Own elaboration.

4.2 Analysis for the S&P100 under JD(2)J model

To illustrate the methodology advanced in the paper above, we also analyze a series of daily logarithmic rates of return on the S&P 100 Index over the period from March 5, 1984 through July 8, 1997. The series has already been employed by (Honore, 1998), who fits it with the Bernoulli jump-diffusion model by means of the maximum likelihood method, as well as by (Kostrzewski, 2013a). Quotations on the S&P100 Index have been downloaded from http://www.econstats.com.

Figure 4 displays the posterior probabilities of a jump (only the ones that exceed 0.5), the series of waiting times between consecutive jumps and the ACF for the latter series. Similar remarks to the ones formulated in the previous case can be made here. Clearly, periods of no jumps alternate with the ones of frequent jumps. That the jumps noticeably tend to cluster over time is also indicated by the first and the second lag autocorrelation coefficients, which are equal around 0.4 and statistically significant. More details on the estimation results are available from the author upon request.
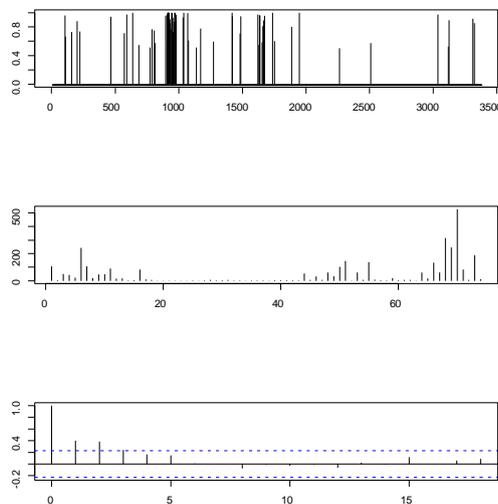
Figure 4. The time series of the posterior probabilities of a jump (only the ones exceeding 0.5; top), the series of times between successive jumps (middle), and the autocorrelation function of the times between consecutive jumps (bottom).

Source: Own elaboration.

5. Final remarks

The presented methodology is capable of detecting jumps and the jump clustering phenomenon. The empirical examples show that jump may indeed tend to cluster and that the phenomenon itself can actually be diagnosed by means of models whose structure does not incorporate potential autocorrelation of the waiting time between jumps *per se*. Jump activity may vary over time so that the assumption of a constant jump intensity might appear no longer valid.

It appears that the methodology may provide some indications as to whether or not employ jump-diffusion models for a time series. In that context, one of the main conclusions (and a very broad one) is that financial time series models which allow jumps as well as jump clustering, indeed, might be empirically adequate. Therefore, the focus of future research should be placed upon, for instance, specifications with stochastic jump intensity employing either Hawkes processes (Hawkes, 1971) or non-homogeneous Poisson processes, which would enhance the model structure so as to account for jump clustering explicitly. Moreover, further studies could concentrate on jump clustering, yet under stochastic volatility framework.

References
1. BERNARDO J. M., SMITH A. F. M. 2002. Bayesian Theory. East Sussex: Wiley Series in Probability and Statistics. 2002. ISBN 978-0471494645.

2. CHIB S., GREENBERG E. 1995. Understanding the Metropolis-Hastings Algorithm. In The American Statistician, 1995, vol. 49, iss. 4, pp. 327-335.
3. GAMERMAN, D., LOPES H. F. 2006. Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference. Boca Raton: Chapman & Hall/CRC. 2006. ISBN 978-1584885870.
4. HANSON, F.B., WESTMAN J.J., 2002. Stochastic Analysis of Jump-diffusions for Financial Log-return Processes. In Pasik-Duncan B. (ed.) Stochastic Theory and Control. Proceedings of a Workshop held in Lawrence, Kansas. Lectures Notes in Control and Information Sciences, Springer-Verlag, New York. 2002. ISBN 978-3-540-43777-2.
5. HAWKES, A.G. 1971. Spectra of Some Self-Exciting and Mutually Exciting Point Processes. In Biometrika, 1971, vol. 58, iss. 1, pp. 83-90.
6. HONORE, P. 1998. Pitfalls in Estimating Jump-Diffusion Models (January 24, 1998). [cit. 08-04-2014] http://dx.doi.org/10.2139/ssrn.61998.
7. KNIGHT, J. L., SATCHELL, S. 2007. GARCH Processes − Some Exact Results, some Difficulties and a Suggested Remedy. In KNIGHT, J. L., SATCHELL, S. (eds.) Forecasting Volatility in the Financial Markets. Amsterdam: Elsevier. 2007. ISBN 075066942.
8. KOSTRZEWSKI, M. 2012. On the Existence of Jumps in Financial Time Series. In Acta Physica Polonica B, 2012, vol. 43, iss. 10, pp. 2001-2019.
9. KOSTRZEWSKI, M. 2013a. Bayesian Inference for the Jump-Diffusion Model with M Jumps. In Communications in Statistics − Theory and Methods, 2013. DOI: 10.1080/03610926.2012.755202.
10. KOSTRZEWSKI, M. 2013b. Bayesian DEJD model and detection of asymmetric jumps (December 22, 2013) [cit. 09-04-2014], http://arxiv.org/abs/1404.2050.
11. LEE, S.S. 2012. Jumps and Information Flow in Financial Markets. In Review of Financial Studies, 2012, vol. 25, iss. 2, pp. 439-479.
12. LEE, S. S., MYKLAND P.A. 2008. Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics. In The Review of Financial Studies, 2008, vol. 21, iss. 6, pp. 2535-2563.
13. MAHEU, J. M., MCCURDY T. H. 2004. News Arrival, Jump Dynamics, and Volatility Components for Individual Stock Returns. In The Journal of Finance, 2004, vol. 59, iss. 2, pp. 755-793.
14. R CORE TEAM. R: a language and environment for statistical computing. The R Foundation for Statistical Computing, 2013, http://www.r-project.org/.
15. YU, J. 2004. Empirical Characteristic Function Estimation and Its Applications. In Econometric Reviews, 2004, vol. 23, iss. 2, pp. 93-123.