ON THE EXISTENCE OF JUMPS IN FINANCIAL TIME SERIES*

Maciej Kostrzewski

Cracow University of Economics
Rakowicka 27, 31-510 Kraków, Poland
and
AGH University of Science and Technology
30 Mickiewicza Av., 30-059 Kraków, Poland

(Received June 11, 2012; revised version received October 15, 2012)

In this research two methods of detecting jumps are presented. One is based on the nonparametric approach, whereas the other — on the JD(M)J model. Bayesian inference is applied to detect jumps in the JD(M)J model. Intraday and daily rates of return are under consideration. The empirical results imply the existence of jumps. The information on existing jumps is exploited in a forecasting experiment focused on Value at Risk predictions.

DOI:10.5506/APhysPolB.43.2001
PACS numbers: 89.65.Gh, 02.50.–r, 05.45.Tp

1. Introduction

In many financial time series, we observe small changes of data over time (the so-called continuous changes) as well as occasional jumps. These time series are often modelled by jump-diffusion processes. The processes are solutions of the stochastic differential equations $dY_t = \mu(t)dt + \sigma(t)dW_t + k(t)dq_t$. The first two elements define the continuous part of the process and form a so-called pure diffusion process. The last element constitutes a pure jump process. In practice, we want to know whether to consider jumps or not. In other words, it is a question whether we should take into account the element $k(t)dq_t$. Some familiar examples of a pure diffusion process are the arithmetic and geometric Brownian motions[1]. The Merton model assumes dynamics of an asset given by the standard jump-diffusion process [2]. The jump-diffusion processes are commonly used in financial econometrics.

* The research was supported by the Polish Ministry of Science and Higher Education. Research Project 2010–2012, No. N N111 429139.

[1] Bayesian estimation of the pure diffusion processes are discussed in [1].

The aim of the paper is to present two methods of jump detection: a parametric and a nonparametric one. We verify the hypothesis of the existence of jumps. It is also interesting to infer about the number of jumps per day.

The nonparametric method of jumps detection is based on the results of Barndorff–Nielsen and Shephard [3–5], utilizing a quadratic variation which is comprised of a sum of two ingredients. One of them corresponds to a pure diffusion process and the other to a pure jump process. In practice employing a realized variance (introduced in [6]) and a realized bipower variation (introduced in [3]) we can separate these two parts. The nonparametric test of detecting jumps is based on the following idea. If a pure jump component is greater than zero, then jumps exist, otherwise not. This approach assumes that a fixed time interval between the following observations should be infinitesimally small. In practice, it means that the test should be limited to high frequency data. The assumption also limits applicability of some other commonly known tests of jumps.

The likelihood function in the Merton model [2] is a product of infinite mixtures of normal distributions. Let us restrict the infinite mixtures to finite ones with the number of components equal $M + 1$. The likelihood function in the Merton model [2] is a product of an infinite number of mixtures of normal distributions. Let us assume that this number equals $M + 1$. If $M = 0$, then we get a pure diffusion process, for $M = 1$ we get the well-known Bernoulli jump-diffusion process. Generally, for $M \geq 0$ the model could be called the jump-diffusion model with M jumps (JD(M)J). In this paper we consider the Bayesian inference for the model. Bayesian comparisons of the JD(M)J models, the full Bayesian significance test and the comparison of some posterior probabilities are used to estimate the number of jumps M. Let us notice that the familiar Bernoulli jump-diffusion process assumes the existence of at most one jump at a fixed term, while the JD(M)J model allows for M jumps.

Using the nonparametric method, we can assess the number of jumps per day. We can compare the number of jumps with the Bayesian inference about M. 5-minute and 10-minute data are used in the nonparametric method, whilst daily data are employed for the parametric model. We compare the assessed numbers of jumps per day which are inferred from the two methods. In this work, we do not distinguish days with and without macroeconomic releases. Such releases could be one of the reasons for jumps [7].

This paper is organized as follows. The nonparametric method of jump detection is outlined in Sec. 2. In Sec. 3, we define the JD(M)J model and present Bayesian framework of the model. Next, in Secs. 4 and 5, we present the results of the empirical research, focusing on testing of jump existence, estimation of the number of jumps and a Value at Risk analysis. The methodology is presented using a series of logarithmic rates of return on KGHM shares.

2. The nonparametric approach

Our basic assumption is that a time series is, or we believe it is, a trajectory of a stochastic process. The values of the time series are positive and a logarithm of the values is given by the process Y. Y is a solution of the stochastic differential equation

$$dY_t = \mu(t)\,dt + \sigma(t)\,dW_t + k(t)\,dq_t\,,$$

where μ, σ and k are some unknown functions, which should guarantee the existence of the unique solution of the above equation. W denotes a Brownian motion and q is a counting process. The component $\mu(t)dt + \sigma(t)dW_t$ defines a pure diffusion process, while $k(t)dq_t$ defines a pure jump process. Therefore Y is called a jump-diffusion process.

A logarithmic rate of return on an interval $[0,t]$ is defined by $X_t = Y_t - Y_0$. Let $[X,X]_t = \int_0^t \sigma^2(s)\,ds + \sum_{0\le s\le t} k^2(s)$ denote a quadratic variation, where the second term is generated by the jumps. We restrict our considerations to a single trading day. Let δ be the time between consecutive observations. Then $1/\delta$ is the number of returns. Finally, $X_j(\delta)$ stands for the jth logarithmic rate of return.

The logic behind the method is to estimate $\sum_{0\le s\le 1} k^2(s)$ by employing a realized variance $RV(\delta) = \sum_{j=1}^{1/\delta} X_j^2(\delta)$ (introduced by [6]) and a realized bipower variation (introduced by [3])

$$BV(\delta) = \mu_1^{-2}(1-2\delta)^{-1}\sum_{j=3}^{1/\delta} |X_j(\delta)||X_{j-2}(\delta)|\,,$$

where $\mu_1 = \sqrt{2/\pi}$.

If $\delta \to 0$, then $RV(\delta) \underset{p}{\to} \int_0^1 \sigma^2(s)ds + \sum_{0\le s\le 1} k^2(s)$ and $BV(\delta) \underset{p}{\to} \int_0^1 \sigma^2(s)ds$. Finally, the sum of squared values of jumps on an interval $[0,1]$ could be estimated by a consistent estimator

$$RV(\delta) - BV(\delta) \underset{p}{\to} \sum_{0\le s\le 1} k^2(s)$$

for an infinitesimal small δ. In practice the sampling frequency δ often corresponds to five- or ten-minutes data.

To describe the algorithm of detecting jumps, let us define the tripower quarticity (introduced by [8])

$$TQ(\delta) = \delta^{-1}\mu_{4/3}^{-3}(1-4\delta)^{-1}\sum_{j=5}^{1/\delta} |X_j(\delta)|^{4/3}|X_{j-2}(\delta)|^{4/3}|X_{j-4}(\delta)|^{4/3}\,,$$

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

If $\delta \to 0$, then $TQ(\delta) \underset{p}{\to} \int_0^1 \sigma^4(s)ds$. Let α be the significance level and $\Phi_{1-\alpha}$- the corresponding $(1-\alpha)$ — quantile of the standard normal distribution $N(0,1)$. The following algorithm was designed in [9].

Algorithm 1 *1. The number of jumps equals zero. Define an empty set of indices $SJ = \emptyset$.*

2. Calculate $Z(\delta) = \frac{\delta^{-1/2}(RV(\delta)-BV(\delta))RV^{-1}(\delta)}{\sqrt[2]{\left(\mu_1^{-4}+2\mu_1^{-2}-5\right)\max\{1,TQ(\delta)BV^{-2}(\delta)\}}}$[2].

3. If $Z(\delta) > \Phi_{1-\alpha}$ then the number of jumps increases by one.

(a) Determine k, such that $|X_k(\delta)| = \max_{j\in\{1,\ldots,1/\delta\}\setminus SJ}\{|X_j(\delta)|\}$
Define $SJ := SJ \cup \{k\}$
Replace $RV(\delta) := RV(\delta) - X_k^2(\delta) + \frac{\delta}{1-\delta}\sum_{j=1,j\notin SJ}^{1/\delta} X_j^2(\delta)$.

(b) Return to the step two.

4. If $Z(\delta) \le \Phi_{1-\alpha}$ then exit.

In Sec. 3.2, we apply the algorithm for every single day over a given period of time. It is easy to see that the inferred number of jumps per day depends on the value of α. We assumed that α equals 0.01. The value is one of the recommended magnitudes in the literature [8, 10].

3. The JD(M)J model

In the Merton model [2], the price of a share is governed by the jump-diffusion process S, which is the solution of the equation

$$dS_t = \mu S_t dt + \sigma S_t dW_t + \left(e^Q - 1\right) S_t dN_t\,, \tag{1}$$

where W is a standard Wiener process, N is a Poisson process with an intensity $\lambda > 0$, $(Q_j)_{j\ge 1}$ are independent normal random variables with mean μ_Q and variance σ_Q^2. Moreover, W, N and Q are stochastically independent. Then, logarithm of the price, $Y_t = \ln S_t$, and the process of logarithmic rates of return $Y_{t+\Delta} - Y_t = \ln(\frac{S_{t+\Delta}}{S_t})$, are given by

$$dY_t = \left(\mu - \tfrac{1}{2}\sigma^2\right) dt + \sigma dW_t + Q dN_t\,, \tag{2}$$

$$\ln\left(\frac{S_{t+\Delta}}{S_t}\right) = \left(\mu - \tfrac{1}{2}\sigma^2\right)\Delta + \sigma\left(W_{t+\Delta} - W_t\right) + \sum_{j=N_t+1}^{N_{t+\Delta}} Q_j\,, \quad \Delta > 0\,. \tag{3}$$

[2] If $\delta \to 0$, then $Z(\delta) \underset{d}{\to} N(0,1)$.

The probability density function of $\ln(\frac{S_{t+\Delta}}{S_t})$ is given in [11]

$$p_{\ln\left(\frac{S_{t+\Delta}}{S_t}\right)}(x)=\sum_{k=0}^{\infty}\exp\left(-\lambda\Delta\right)\frac{\left(\lambda\Delta\right)^k}{k!}\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q k,\sigma^2\Delta+\sigma_Q^2 k\right)\ , \tag{4}$$

where $\phi(\cdot;m,s^2)$ is the density of a normal distribution with mean m and variance s^2. The likelihood function is a product of the infinite mixture of the normal densities. Let us consider the approximation

$$p_{\ln\left(\frac{S_{t+\Delta}}{S_t}\right)}(x)=\sum_{k=0}^{\infty}\exp\left(-\lambda\Delta\right)\frac{\left(\lambda\Delta\right)^k}{k!}\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q k,\sigma^2\Delta+\sigma_Q^2 k\right) \tag{5}$$

$$\approx\sum_{k=0}^{M}\exp\left(-\lambda\Delta\right)\frac{\left(\lambda\Delta\right)^k}{k!}\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q k,\sigma^2\Delta+\sigma_Q^2 k\right)\ , \tag{6}$$

where $M \geq 0$ is a fixed constant. In other words, we assume that the random variable of the number of jumps per Δ takes values in the set $\{0,1,...,M\}$. Let us normalize the approximation (6) so as to obtain a probability density function

$$p\left(x\left|\theta,M\right.\right)=\sum_{k=0}^{M}w_k\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q k,\sigma^2\Delta+\sigma_Q^2 k\right)\ , \tag{7}$$

where $w_k = \frac{(\lambda\Delta)^k}{k!}\left[\sum_{j=0}^{M}\frac{(\lambda\Delta)^j}{j!}\right]^{-1}$ for $k=0,\dots,M$. The model considered in this paper is given by (7). We call it a Jump-Diffusion model with M-Jumps, or JD(M)J model, in short. Note that under $M=0$ Eq. (7) defines the arithmetic Brownian motion.

The Bernoulli jump-diffusion model [12], or BJD in short, allows only for two possibilities — either a single or no jumps (at all) over a fixed period of time Δ. Then, the density is given by

$$\begin{aligned} p\left(x\left|\theta,\mathrm{BJM}\right.\right) &= \left(1-\lambda\Delta\right)\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta,\sigma^2\Delta\right) \\ &\quad +\lambda\Delta\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q,\sigma^2\Delta+\sigma_Q^2\right)\ , \end{aligned}$$

where $\lambda\Delta$ is assumed to be close to zero. Note that in the JD(1)J model the density is given by

$$\begin{aligned} p\left(x\left|\theta,M=1\right.\right) &= \frac{1}{1+\lambda\Delta}\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta,\sigma^2\Delta\right) \\ &\quad +\frac{\lambda\Delta}{1+\lambda\Delta}\phi\left(x;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q,\sigma^2\Delta+\sigma_Q^2\right)\ , \end{aligned}$$

and so the two models (BJD and JD(1)J) are equivalent. It is interesting to investigate whether the JD(1)J model (or, equivalently, the Bernoulli jump-diffusion model) is an adequate model for financial time series in comparison with the pure-diffusion model or the models with more than one jump at a fixed time interval. Further considerations are restricted to the JD(M)J models.

Let us assume that a time series $x=(x_1,\ldots,x_n)$ is observed at $(t_1,\ldots,t_n)$, where $x_i=\ln\left(\frac{S_{t_i}}{S_{t_i-\Delta}}\right)$. There are five unknown parameters of the JD(M)J model $(\mu,\sigma^2,\lambda,\mu_Q,\sigma_Q^2)\in\Theta$, where $\Theta=\mathbb{R}\times(0,\infty)\times(0,\infty)\times\mathbb{R}\times(0,\infty)$. Θ is an admissible set of the parameters.

If we analyse the path of a jump-diffusion process we do not know whether the observations, or which of them, have resulted from the pure diffusion or the jump-diffusion process. In other words, we do not know which component of the sum in the density (7), given by

$$\phi\left(\cdot;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q k,\sigma^2\Delta+\sigma_Q^2 k\right)$$

for $k=0,\ldots,M$, is "responsible for" the observation. In order to manage the problem, let us introduce the latent vector variable $Z=(Z_1,\ldots,Z_n)$ such that $Z_i\in\{0,1,\ldots,M\}$ and $P(Z_i=j)=w_j$, where $i\in\{1,\ldots,n\}$ and $j\in\{0,1,\ldots,M\}$. The random variable Z_i is the number of jumps at the ith interval of time. The likelihood function is then given by

$$p\left(x\,|Z,\theta\right)=\prod_{i=1}^{n}\sum_{k=0}^{M}w_k\phi\left(x_i;\left(\mu-\tfrac{1}{2}\sigma^2\right)\Delta+\mu_Q Z_i,\sigma^2\Delta+\sigma_Q^2 Z_i\right)\,.$$

It is convenient to introduce the reparametrization: $L=\lambda\Delta$, $h_\sigma=\frac{1}{\sigma^2}$, $h_Q=\frac{1}{\sigma_Q^2}$. The vector of unknown parameters is denoted by $(Z,\theta)=(Z_1,\ldots,Z_n,\mu,h_\sigma,L,\mu_Q,h_Q)$.

The Bayesian model is defined by $p(x,Z,\theta)=p(x|Z,\theta)p(Z,\theta)$. The prior structure for (Z,θ) is defined by

$$p\left(Z,\theta\right)=p\left(Z\,|\theta\right)p\left(\theta\right)=p\left(\mu\right)p\left(h_\sigma\right)p\left(L\right)p\left(\mu_Q\right)p\left(h_Q\right)\prod_{i=1}^{n}p\left(Z_i\,|\theta\right)\,,$$

where

$$P\left(Z_i=j\,|\theta\right)=w_j\,,\qquad w_k=\frac{(\lambda\Delta)^k}{k!}\left[\sum_{j=0}^{M}\frac{(\lambda\Delta)^j}{j!}\right]^{-1}\quad\text{for } k=0,\ldots,M\,,$$

$$p\left(\mu\right)=\phi\left(\mu;m_\mu,s_\mu^2\right)\,,\qquad p\left(\mu_Q\right)=\phi\left(\mu_Q;m_Q,s_Q^2\right)\,,$$

$$p\left(h_\sigma\right)\propto h_\sigma^{(\nu_\sigma-2)/2}\exp\left(-Ah_\sigma/2\right)\,,\qquad p\left(h_Q\right)\propto h_Q^{(\nu_Q-2)/2}\exp\left(-Bh_Q/2\right)\,,$$

$$p\left(L\right)\propto L^{(\nu_L-2)/2}\exp\left(-L/2\right)\,.$$

The Bayesian inference is hinged on calculation of a posterior distribution. The distribution is often sampled from by means of Markov Chain Monte Carlo methods [13]. In this paper, posterior characteristics of the parameters and latent variables are estimated by the hybrid Metropolis–Hastings algorithm within the Gibbs sampler and the acceptance–rejection sampling method [14].

The hypothesis about existing jumps in the Bayesian framework is equivalent to the question about the magnitude of M. If M is greater than zero it means that jumps should be taken into account. If M is greater then one we reject the Bernoulli jump-diffusion model. In what follows, we compare the JD(M)J models for some values of M. The best model hints at the appropriate (*i.e.* data-supported) value of M.

4. Empirical research

Let us notice that the primary assumption of the nonparametric approach is the stochastic differential equation $dY_t = \mu(t)dt + \sigma(t)dW_t + k(t)dq_t$ for the logarithm of the prices process, whilst for the parametric model $dY_t = \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma dW_t + QdN_t$. The last equation is a special case of the former with $\mu(t) \equiv \mu - \frac{1}{2}\sigma^2$, $\sigma(t) \equiv \sigma$, $k(t) = Q_t$ and $q_t = N_t$. It is interesting to compare the outcomes of the nonparametric and Bayesian inference. We collate the results of testing the hypothesis of the existence of jumps along with the estimated number of jumps over the same period of time.

We apply the nonparametric method separately at each single day for the intraday data. The consecutive intraday data appears every five or ten minutes. In the case of the JD(M)J model, we consider daily quotations (closing prices). The daily data appear successively at $\Delta = \frac{1}{252}$.

4.1. Data

KGHM is a copper producer and one of the largest Polish exporters. The company's share prices contribute to the WIG20 Index. Let us consider the time series of intraday and daily values of logarithmic growth rates of KGHM quotations on the Warsaw Stock Exchange from January 23, 2006 to February 22, 2010. The intraday data are five and ten minute returns. The data are calculated as logarithmic returns of volume-weighted means of prices for every 5-minute and 10-minute periods. We consider continuous trading *i.e.* transactions from 9:00 a.m. to 4:10 p.m. We discard the null intervals (*i.e.* the ones without any transactions). The null five- and ten-minute periods were accordingly 6 and 5 per cent. The data are depicted in Fig. 1. The horizontal lines present a band of two standard deviations in width.

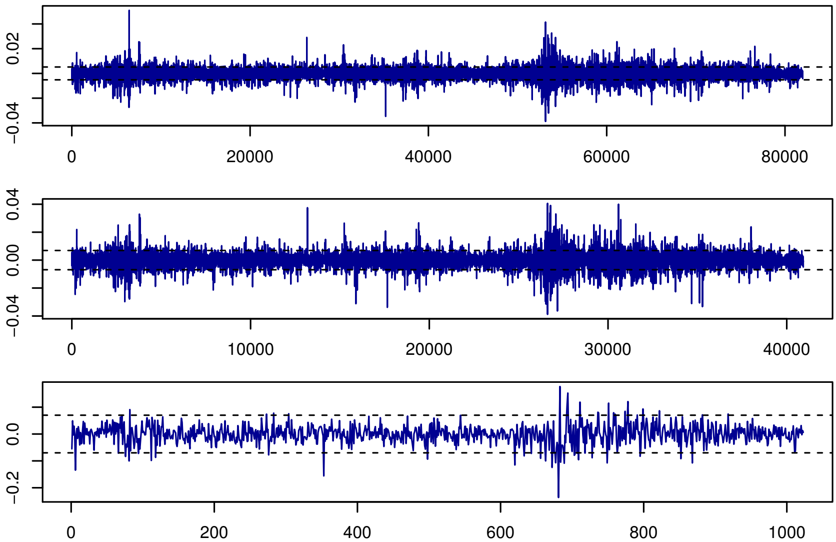

Fig. 1. The time series of the 5-minute (top), 10-minute (middle) and daily (bottom) returns of KGHM.

The nonparametric approach assumes an infinitesimally small period of time between the following observations, *i.e.* $\delta \to 0$. On the other hand, in the case of too small values of δ market mictrostructure noise could affect the results [15–17]. That is why 5-minute intraday data are recommended for liquid shares, whereas 10-minute or less-frequent data for less liquid assets. The KGHM asset is not high liquid. However, in order to compare, we present results for both sets of intraday data. Basic descriptive statistics of the modelled series are reported in Table I.

TABLE I

Descriptive statistics of the data.

KGHM	Mean	Median	Min	Max	Std.Dev.	$\frac{\text{Std.Dev.}}{\text{Median}}$	n
5-minutes	−0.000027	0	−0.03867	0.05101	0.0026	∞	81988
10-minutes	−0.000054	−0.000011	−0.0388	0.04063	0.0034	324	40926
daily	0.00016	0.00115	−0.2362	0.17693	0.0350	30	1023

The 5-minute data are the most volatile, as indicated by the ratios of standard deviation to median.

4.2. Empirical results — jump detection

At first, we present the results of the nonparametric method. We can assess the number of jumps for each day. Figure 2 depicts the numbers of jumps detected by Algorithm 1. The dotted lines repesent one, two and three jumps per day. In the case of 5-minute returns, the maximum number of jumps equals eight, whilst for 10-minute returns this number equals three. It follows that 5-minute returns are "less smooth" than the 10-minute returns, which is consistent with our expectations.

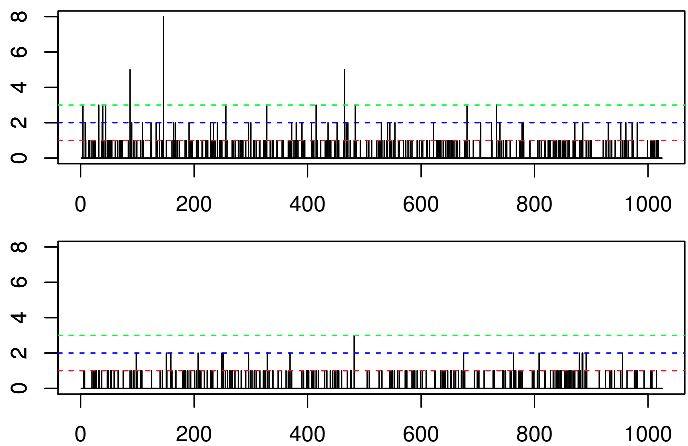

Fig. 2. Numbers of jumps for 5-minute (top) and 10-minute (bottom) returns.

Table II presents the proportion of days with zero, one, two or more jumps per day. On more than 70 per cent of the days no jumps have occured. The results emphasize a crucial role of the pure diffusion part. However, on more than 20 per cent of days there has been at least one jump. The results validate the existence of jumps. The frequency of days with two or more jumps is evidently greater than zero. It seems that the Bernoulli jump-diffusion model may not be an appropriate tool in comparison with models allowing for more than one jump over a fixed period of time.

TABLE II

Frequency of days with a fixed number of jumps.

No. of jumps	0	1	2	3	4	5	6	7	8
5-minutes	0.71	0.24	0.04	0.01	0	0.002	0	0	0.001
10-minutes	0.78	0.199	0.0166	0.001	0	0	0	0	0

Further, we report on the results of the parametric approach. The prior assumptions were defined in Sec. 3. The hyperparameters are specified as follows: $A = B = 1$, $\nu_L = 6\Delta$ (10Δ or 800Δ), $m_\mu = 0.1$ (or 0.01), $s_\mu^2 = 1$ (or 10^7), $\nu_\sigma = 5$ (0.25, 0.1 or 0.5), $m_Q = 0.1$ (or 0.01), $s_Q^2 = 1$ (or 10^7), $\nu_Q = 5$ (0.25, 0.1 or 0.5). Different values of the hyperparameters yield very similar posterior results. The outcomes of the MCMC methods do not depend on starting points. The numerical algorithms applied in the research require monitoring of convergence of the Marov chains to their limiting stationary distributions[3]. To this end we resort to visual inspection of plots displaying means, standard deviations and CUMSUM statistics [18].

Table III depicts posterior means and standard deviations of the parameters. It is easy to see that the pure diffusion process along with the pure jump process play an essential role. The inclusion of the jump component into the model changes the estimation outcomes of the pure diffusion parameters. The result justifies a need for jump-diffusion models. The lack of apparent differences in the estimation of the posterior means and standard deviations for $M = 1$ and $M = 10$ suggests that both models lead to very similar conclusions.

TABLE III

Posterior means and standard deviations for KGHM.

Model	JD(0)J		JD(1)J		JD(10)J	
θ	$E(\cdot\|x)$	$D(\cdot\|x)$	$E(\cdot\|x)$	$D(\cdot\|x)$	$E(\cdot\|x)$	$D(\cdot\|x)$
λ	—	—	3.815	(1.84)	3.767	(1.78)
μ	0.19	(0.266)	0.254	(0.251)	0.255	(0.25)
μ_Q	—	—	−0.025	(0.086)	−0.025	(0.085)
σ^2	0.31	(0.014)	0.254	(0.016)	0.254	(0.015)
σ_Q^2	—	—	0.078	(0.045)	0.08	(0.046)

Let us look at the values of the posterior means of σ^2: 0.31, 0.254 and 0.254 for the JD(0)J, JD(1)J and JD(10)J model, respectively. The highest value is for JD(0)J. It has a reasonable interpretation — the process without jumps needs a higher value of the volatility parameter σ^2. The JD(M)J models for $M > 0$ "absorb some volatility" by the jump components. Histograms of the posterior distributions and densities of their prior counterparts distributions are collected in Fig. 3. The posteriors are far distinguishable from their prior counterparts, indicating a strong data contribution to the inference.

[3] The numerical calculations were performed in R and Maple 14, using the program created by the author.

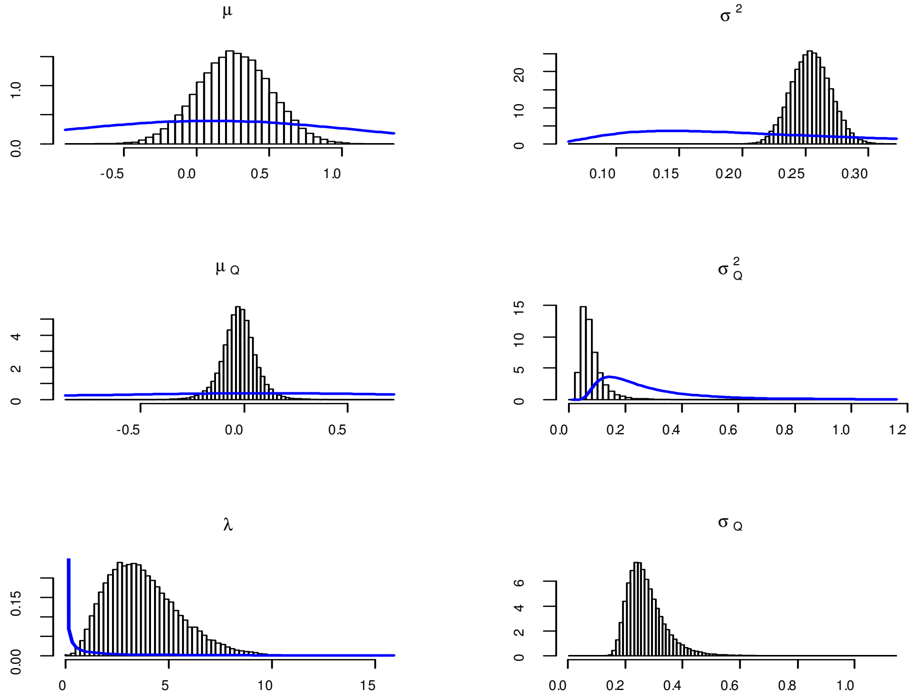

Fig. 3. Histograms of the parameters' posteriors along with the prior densities (solid lines).

We now focus on the choice of an appropriate value of M. Three ways of addressing the problem are possible.

The first method is based on comparison of the competitive models $\mathrm{JD}(M_1)\mathrm{J}$ and $\mathrm{JD}(M_2)\mathrm{J}$. The model with the highest posterior probability is perceived as the best one. We have to compare

$$p\left(\mathrm{JD}(M_1)\mathrm{J}\right|x) = \frac{p\left(\mathrm{JD}(M_1)\mathrm{J}\right)p\left(x\right|\mathrm{JD}(M_1)\mathrm{J})}{\sum_{k=1}^{2} p\left(\mathrm{JD}(M_k)\mathrm{J}\right)p\left(x\right|\mathrm{JD}(M_k)\mathrm{J})}$$

and $p(\mathrm{JD}(M_2)\mathrm{J}|x) = 1 - p(\mathrm{JD}(M_1)\mathrm{J}|x)$. Newton–Raftery estimator may be employed to calculate the marginal data density $p(x|\mathrm{JD}(M)\mathrm{J})$ [19]. The estimator is simulation-consistent, although, in general, it does not have a finite variance. Plots of its values feature often jumps, and evince instability. Therefore, it is advisable to run long Markov chains. In the empirical research an improved version of the estimator is employed. The estimator was formed not from complete likelihood values, but from marginal likelihoods obtained by integrating out variables $Z_1, \ldots, Z_n$. It is consistent with the recommendation of [20].

Let us consider four models: JD(0)J, JD(1)J, JD(2)J, and JD(10)J. It is assumed that the prior probabilities of each model are equal. The Newton–Raftery estimators are employed to calculate the posterior probabilities of the models, based on which we obtain $\log_{10}\left(\frac{p(\mathrm{JD}(1)\mathrm{J}|x)}{p(\mathrm{JD}(0)\mathrm{J}|x)}\right) \approx 9.95$ and $\log_{10}\left(\frac{p(\mathrm{JD}(10)\mathrm{J}|x)}{p(\mathrm{JD}(1)\mathrm{J}|x)}\right) \approx 0.345$. It proves the advantage of the jump-diffusion process over the pure diffusion one. It clearly argues for the existence of jumps. It seems that the JD(10)J specification is slightly favoured against the JD(1)J model. Yet the result should be taken with caution on account of instability of the Newton–Raftery estimator. The above results are based on 250,000 draws of the Markov chain and 10,000 burn-in passes. A longer Markov chain (of 3,700,000 draws) was generated to increase the credibility of the Newton–Raftery estimator. Figure 4 presents values of $\frac{p(\mathrm{JD}(2)\mathrm{J}|x)}{p(\mathrm{JD}(1)\mathrm{J}|x)}$ against the number of the MCMC passes. The logarithm of Bayes factor $\log_{10}\left(\frac{p(\mathrm{JD}(2)\mathrm{J}|x)}{p(\mathrm{JD}(1)\mathrm{J}|x)}\right) \approx 0.7$. It slightly favours the JD(2)J model over the JD(1)J one. However, taking into account the instability of the Newton–Raftery estimator, we ought to be circumspect in drawing any conclusion if a logarithm of Bayes factor is as low as 0.345 or 0.7.

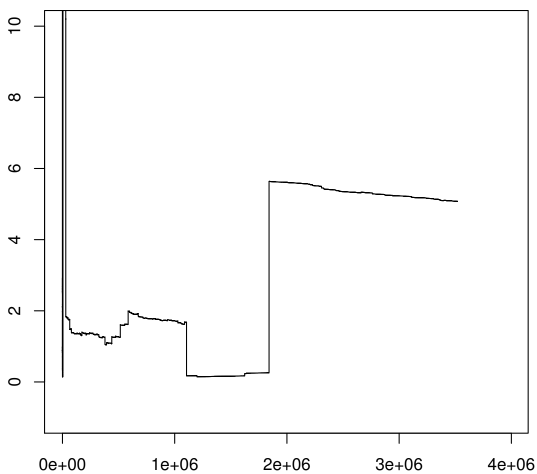

Fig. 4. Values of $p(\mathrm{JD}(2)\mathrm{J}|x)/p(\mathrm{JD}(1)\mathrm{J}|x)$.

The second method is based on an application of the Full Bayesian Significance Test (FBS Test) [21, 22]. In order to verify the hypothesis about existing jumps, we could test the hypothesis H_0: $Z_1 = 0, \ldots, Z_n = 0$ using the Full Bayesian Significance Test. Let us consider the model given by

$$p\left(x, Z_1, \ldots, Z_n, \mu, h_\sigma, L, \mu_Q, h_Q, \theta\right) \, ,$$

where $Z_i \in \{0,1\}$. It corresponds to JD(1)J model. If the probability

$$Ev(H_0) = 1 - P\left((\theta, Z):\ p(\theta, Z\,|x) > \max_{(\theta,Z)\in\theta\times(0,\ldots,0)} p(\theta, Z\,|x)\,|x\right)$$

is large, then the test favours the H_0 and the lack of jumps.

The approximation of $Ev(H_0)$ is based on 800,000 MCMC cycles then

$$\max_{(\theta,Z)\in\theta\times(0,\ldots,0)} p(\theta, Z\,|x) \approx \exp(1966.053) \quad \text{and} \quad Ev(H_0) \approx 0.15\,.$$

The FBS test favours JD(1)J over JD(0)J and we conclude that jumps exist. Figure 5 presents values of $Ev(H_0)$ against the number of the MCMC passes used to approximate $Ev(H_0)$. The estimator of $Ev(H_0)$ quickly converges to 0.15.

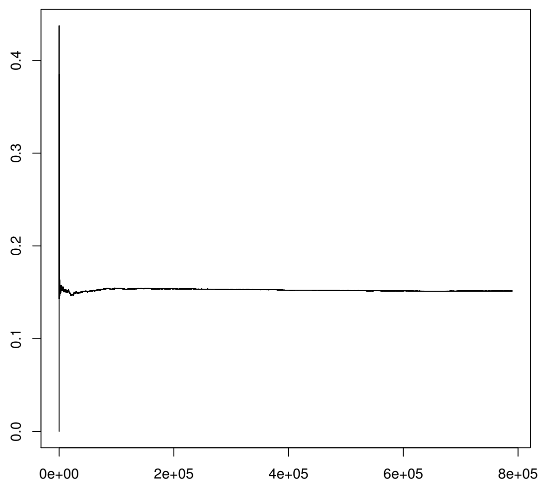

Fig. 5. Values of $Ev(H_0)$ against the number of the MCMC passes.

The third method is based on testing hypothesis: $Z_1 \le j, \ldots, Z_n \le j$ by comparing the posterior probabilities[4] $P(Z_1 \le j, \ldots, Z_n \le j|x; \mathrm{JD}(M)\mathrm{J})$ for $j = 0, \ldots, M$. In that case, we consider a model given by $p(x, Z_1, \ldots, Z_n, \mu, h_\sigma, L, \mu_Q, h_Q, \theta)$, where $Z_i \in \{0, \ldots, M\}$, which corresponds to the JD(M)J specification.

Let us now consider the JD(10)J model and compare posterior probabilities: $P(Z_1 = 0, \ldots, Z_n = 0|x; \mathrm{JD}(10)\mathrm{J}) \approx 0$, $P(Z_1 \le 1, \ldots, Z_n \le 1|x; \mathrm{JD}(10)\mathrm{J}) \approx 0.9762192$, $P(Z_1 \le 2, \ldots, Z_n \le 2|x; \mathrm{JD}(10)\mathrm{J}) \approx 0.9999071$. The presented approximations are calculated with 1,000,000 draws from the

[4] More details about testing hypothesis by comparing posterior probabilities can be found in [23].

posterior distribution. Figure 6 presents values of $P(Z_1 \le 1, \ldots, Z_n \le 1|x;$ JD(10)J) against the number of the MCMC passes. The estimator of the probability quickly converges to 0.9762192. The investigation proves the existence of jumps. Moreover, the method indicates that $M = 2$ is suitable for the modelled series.

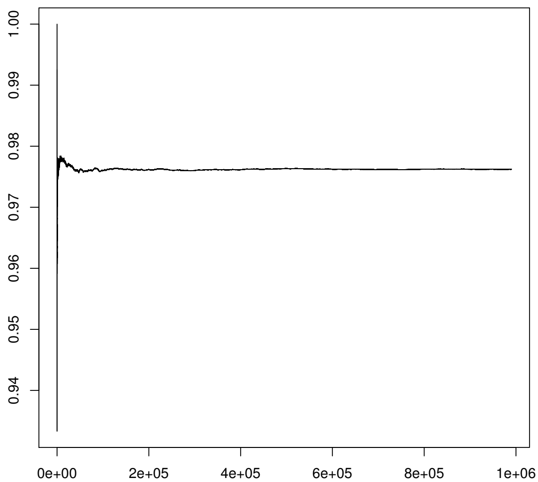

Fig. 6. Values of $P(Z_1 \le 1, \ldots, Z_n \le 1|x; \mathrm{JD(10)J})$ against the number of the MCMC passes.

Furthermore, we collate values of posterior probabilities of jumps for every single day. Figure 7 depicts the time series of the daily returns, values of the probabilities $P(Z_i = 1|x, \mathrm{JD(10)J})$ and $P(Z_i > 1|x, \mathrm{JD(10)J})$ against the number ($i = 1, \ldots, n$) of the successive days. Note that the higher posterior probabilities of jumps go with higher volatility of the time series. Clearly, periods of no jumps alternate with the ones of frequent jumps. It suggests existence of jump clustering (it is based on the same idea as volatility clustering). Preliminary empirical researches (not reported in this paper) confirm that the presented methodology is able to detect jump clustering. This issue is left for a further research. The posterior probabilities of more than one jumps ($Z_i > 1$) for every single day are lower than 0.007. Assume for a while that a jump occurs when the posterior probability of a jump exceeds 0.5. Then there are 5 downward jumps and 4 upward jumps. Some additional analysis (peripheral to the present paper and, therefore, not reported here) confirms asymmetry of the jumps. It corresponds to a skew in the empirical data distribution. We note an intuitive relation — higher posterior probabilities of jumps go along with higher absolute values of the data. Jumps' dynamics is explained by the second and higher-numbered components of the mixture. So, jumps are low probable for the pure-diffusion component, *i.e.* normal distribution (observations are in tails of the distribution). In the case of heavy tails distributions (a study on this topic might be found in [24] such jumps (or extreme values) might be typical values.

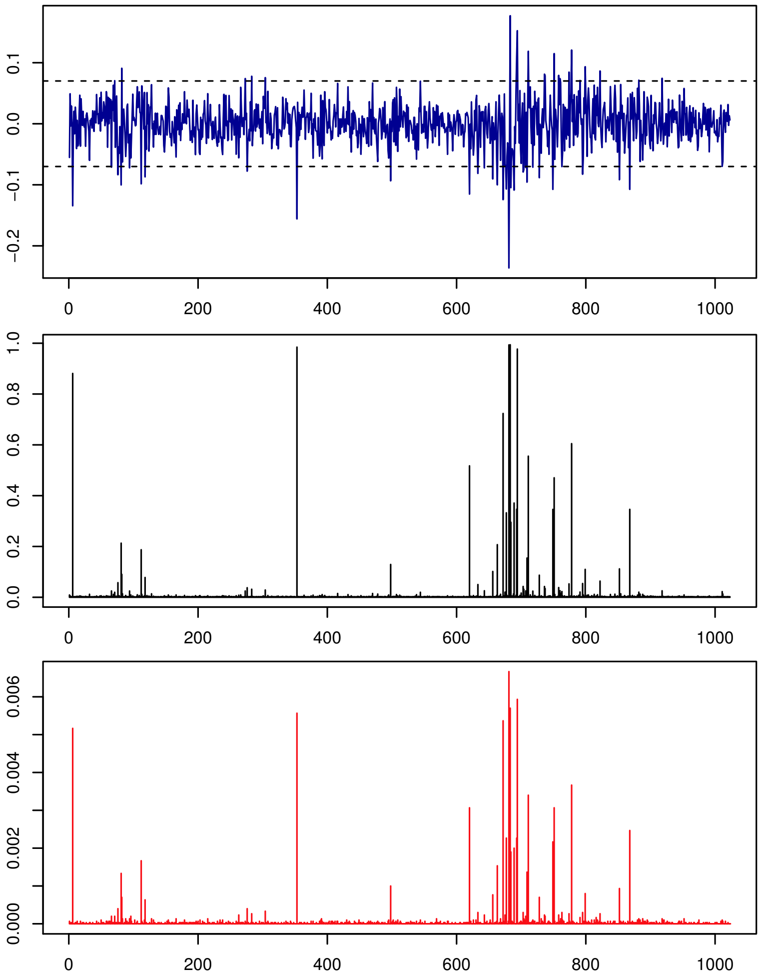

Fig. 7. The time series of the daily returns (top), $i \to P(Z_i = 1|x, \mathrm{JD(10)J})$ (middle) and $i \to P(Z_i > 1|x, \mathrm{JD(10)J})$ (bottom).

To sum up, the two approaches (the nonparametric and the parametric one) detect jumps. Moreover, it is empirically proved that models with more than one jump per day applied to daily data are able to improve fitting.

Additionally, we considered daily quotations of S&P100 Index (March 5, 1984 to July 8, 1997 and simulated data. We detected jumps in S&P100 Index ($E(\lambda|x) \approx 9$). Further details of this research can be found in [25].

4.3. Empirical results — VaR

One of the conclusions of the previous section is that jumps should be taken into account while analyzing the KGHM time series. Employing the parametric model JD(M)J and the Bayesian methods, we assess Value at Risk (VaR) predictions. We consider a long position and a tolerance level α. Then the one-period ahead Value at Risk $\mathrm{VaR}^l_{s:t}(\alpha, t+1)$ at time $t+1$, for $s < t$, is defined as a minus α-quantile of the predictive distribution $p(x_{t+1}|x_s, \ldots, x_t; \mathrm{JD}(M)\mathrm{J})$

$$P\left(x_{t+1} \leq -\mathrm{VaR}^l_{s:t}\left(\alpha, t+1\right)|x\right) = \alpha\,.$$

The predictive distribution reflects uncertainty about the future rates of return (given the data and the Bayesian model specification) while taking into account the parameter uncertainty.

Kupiec test [26] is used to verify the accuracy of the VaR predictions. It is a backtesting procedure which counts VaR exceedances, *i.e.* the number of days the observed returns are lower than the predicted $\mathrm{VaR}^l_{s:t}$. The null hypothesis (H0) reads: the number of VaR breaks equals α. In order to backtest, we calculate 800 one-day ahead forecasts of $\mathrm{VaR}^l_{s:t}$. The forecasts are calculated on the basis of 200 preceding observations. The predictive distributions [27] $p(x_{t+1}|x_{t-200}, \ldots, x_t, M)$ were used to calculate the forecast of $\mathrm{VaR}^l_{t-200:t}(\alpha, t+1)$ according to the formula

$$\int\limits_{-\infty}^{-\mathrm{VaR}^l_{t-200:t}(\alpha,t+1)} p\left(x_{t+1}\left|x_{t-200}, \ldots, x_t; \mathrm{JD}\left(M\right)\mathrm{J}\right.\right) = \alpha\,.$$

Table IV presents the results of the test and values of a conditional expected shortfall ES [28] defined as

$$\mathrm{ES}\left(\alpha, n+1\right) = E\left(-x_{n+1}\middle|x_{n+1} < -\mathrm{VaR}^l_{1:n}\left(\alpha, n+1\right); x, \mathrm{JD}\left(M\right)\mathrm{J}\right)\,,$$

where n is the sample size.

VaR tells us nothing about potential size of the loss, while the conditional expected shortfall does. VaR and ES are essential tools to assess riskiness of trading activities.

Let us notice small p-value for the JD(0)J model and $\alpha = 0.01$. In the other cases the test favours unambiguously the null hypothesis. The p-values for $\alpha \in \{0.01, 0.1\}$ and the JD(M)J models for $M \in \{1, 2\}$ are higher than for the pure diffusion specification (JD(0)J). It seems that the JD(M)J models for $M \in \{1, 2\}$ are favoured against JD(0)J. The frequency of VaR exceedances are similar to α. The inclusion of jumps did not improve

TABLE IV

Results of the Kupiec test and values of ES.

α	VaR breaks	Freq. of VaR breaks	p-value	ES$(\alpha, n+1)$
JD(0)J				
0.01	14	0.017	0.054	0.094
0.05	37	0.046	0.622	0.073
0.1	71	0.089	0.280	0.062
JD(1)J				
0.01	12	0.015	0.186	0.210
0.05	44	0.055	0.520	0.093
0.1	75	0.094	0.550	0.070
JD(2)J				
0.01	12	0.015	0.186	0.206
0.05	44	0.055	0.520	0.093
0.1	76	0.095	0.630	0.070

significantly the results of the Kupiec test. The values of the conditional expected shortfalls are higher for $M \geq 1$. In other words, in the case of the models with jumps, average losses upon VaR exceedance are higher. The results for $M = 1$ and $M = 2$ are similar. The outcomes imply that in the case of this particular risk analysis the models with $M > 1$ are unnecessary.

4.4. Summary

In this paper two methods of detecting jumps are presented. One of them is based on the nonparametric approach. The other one — based on the JD(M)J model — builds upon some generalization of the Bernoulli jump-diffusion model. In this paper, the Bayesian framework for estimation and prediction within the model is adopted. The methods are applied to analyse logarithmic returns on the KGHM stock market shares. The nonparametric method is applied to the intraday data, while daily data are modelled with the JD(M)J model. Both methods provide clear evidence for jumps. There are days with more than one jump detected. The posterior probabilities of more than one jump per a single day are low, but (in some cases) higher than zero. Therefore, it appears that both the pure diffusion process and the Bernoulli jump-diffusion process may not be good enough for financial time series analysis. It follows that one should apply the JD(M)J models rather than the familiar Bernoulli jump-diffusion model. We draw these conclusions from different methods and time series (the intraday and daily returns). The outcomes of the Value at Risk predictions imply that in the case of this particular risk analysis the models with $M > 1$ are unnecessary.

The nonparametric approach assumes that a fixed time interval between the following observations should be infinitesimally small. In practice it means that the test should be limited to high frequency data. The Bayesian approach is free from this defect. Moreover, Bayesian statistics equipped with the Markov Chain Monte Carlo methods gives us an easy way of estimating and forecasting.

Useful comments and remarks by anonymous referee are highly appreciated. The author would also like to thank Łukasz Kwiatkowski for language verification of the manuscript.

REFERENCES

[1] M. Kostrzewski, *Bayesian Analysis of Financial Time Series Modelled by Diffusion Processes*, AGH University of Science and Technology Press, 2006 (in Polish).

[2] R.C. Merton, *J. Financial Economics* **3**, 125 (1976).

[3] O.E. Barndorff-Nielsen, N. Shephard, *J. Financial Econometrics* **2**, 1 (2004).

[4] O.E. Barndorff-Nielsen, N. Shephard, *J. Financial Econometrics* **4**, 1 (2006).

[5] O.E. Barndorff-Nielsen, N. Shephard, *J. Econometrics* **131**, 217 (2006).

[6] T.G. Andersen, T. Bollerslev, F. Diebold, P. Labys, *J. Am. Stat. Assoc.* **96**, 42 (2001).

[7] B. Będowska-Sójka, *Dynamic Econometric Models* **11**, 99 (2011).

[8] T.G. Andersen, T. Bollerslev, F.X. Diebold, *Rev. Econ. Stat.* **89**, 701 (2007).

[9] T.G. Andersen,T. Bollerslev, P. Frederiksen, M.O. Nielsen, *Continuous-Time Models, Realized Volatilities, and Testable Distributional Implications for Daily Stock Returns*, Queens Economics Department Working Paper No. 1173, 2008.

[10] T. Ane, C. Metais, *Int. J. Bus. Econ.* **9**, 1 (2010).

[11] F.B. Hanson, J.J. Westman, *Stochastic Theory and Control* **280**, 169 (2002).

[12] C.A. Ball, W.N. Torous, *J. Financial Quant. Anal.* **1**, 53 (1983).

[13] D. Gamerman, H.F. Lopes, *Markov Chain Monte Carlo. Stochastic Simulation for Bayesian Inference*, Chapman & Hall/CRC, 2006.

[14] S. Chib, E. Greenberg, *The American Statistician* **49**, 327 (1995).

[15] M. Doman, *Dynamic Econometric Models* **10**, 5 (2010).

[16] M. Doman, *Mikrostruktura giełd papierów wartościowych*, Poznań University of Economics Press, 2011 (in Polish).

[17] P. Kliber, *Dynamic Econometric Models* **11**, 171 (2011).

[18] B. Yu, P. Mykland, *Statistics and Computing* **8**, 275 (1998).

[19] M.A. Newton, A.E. Raftery, *J. Royal Statistical Society B* **56**, 3 (1994).

[20] A.E. Raftery, M.A. Newton, J.M. Satagopan, P.N. Krivitsky, *Bayesian Statistics* **8**, 1 (2007).

[21] C.A. de Braganca Pereira, J.M. Stern, *Entropy* **1**, 99 (1999).

[22] L. Rifo, S. Torres, *Commun. Statistics — Theory and Methods* **38**, 1262 (2009).

[23] J. Marzec, *Bayesowskie modele zmiennych jakościowych i ograniczonych w badaniach niespłacalności kredytów*, Cracow University of Economics Press, 2008 (in Polish).

[24] J. Kwapień, S. Drożdż, *Phys. Rep.* **515**, 115 (2012).

[25] M. Kostrzewski, `http://home.agh.edu.pl/ kostrzew/BayesianInferenceJDMJ.pdf`

[26] P. Kupiec, *J. Derivatives* **3**, 73 (1995).

[27] J.M. Bernardo, A.F.M. Smith, *Bayesian Theory*, Wiley Series in Probability and Statistics, 2002.

[28] O. Scaillet, *Mathematical Finance* **14**, 115 (2004).