

Extraction of market expectations from risk-neutral density*¹

Josip Arnerić², Zdravka Aljinović³, Tea Poklepović⁴

Abstract

The purpose of this paper is to investigate which of the proposed parametric models for extracting risk-neutral density; among Black-Scholes Merton, mixture of two log-normals and generalized beta; give the best fit. The model that fits sample data better is used to describe different characteristics (moments) of the ex ante probability distribution. The empirical findings indicate that no matter which parametric model is used, the best fit is always obtained for short maturity horizon, but when comparing models in short-run, the mixture of two log-normals gives statistically significant smaller MSE. According to the pair-wise comparison results, the basic conclusion is that the mixture of two log-normals is superior to the other parametric models and has proven to be very flexible in capturing commonly observed characteristics of the underlying financial assets, such as asymmetries and “fat-tails” in implied probability distribution.

Key words: market expectation, risk-neutral density, mixture of log-normals, Black-Scholes Merton, generalized beta, maturity horizon, DAX index options

JEL classification: C14, C58, G1

* Received: 10-11-2015; accepted: 15-12-2015

¹ This work has been fully supported by Croatian Science Foundation under the project “Volatility measurement, modeling and forecasting” (5199).

² Assistant Professor, University of Zagreb, Faculty of Economics and Business, Trg. J. F. Kennedyja 6, 10000 Zagreb, Croatia. Scientific affiliation: econometric methods and models, financial time series and volatility, VAR models and cointegration, GARCH and MGARCH models, stochastic processes and risk management. Phone: +385 1 238 3361. Fax: +385 1 233 5633. E-mail: jarneric@efzg.hr. Website: <http://www.efzg.unizg.hr/default.aspx?id=15332>. (corresponding author).

³ Full Professor, University of Split, Faculty of Economics, Cvite Fiskovića 5, 21000 Split, Croatia. Scientific affiliation: mathematical finance – financial modelling, mathematics in economics and multiple-criteria decision making. Phone: +385 21 430 644. Fax: +385 21 430 701. E-mail: zdravka.aljinovic@efst.hr. Website: <http://www.efst.unist.hr/content.php?k=fakultet&p=69&osoba=lil>.

⁴ PhD student, University of Split, Faculty of Economics, Cvite Fiskovića 5, 21000 Split, Croatia. Scientific affiliation: statistics and econometrics in business, financial econometrics, time series and neural networks. Phone: +385 21 430 761. Fax: +385 21 430 701. E-mail: tpoklepo@efst.hr. Website: <http://www.efst.unist.hr/content.php?k=fakultet&p=69&osoba=tpoklepo>.

1. Introduction

Extracting information that is embedded in financial asset prices is helpful to both market participants and public authorities, especially in implementing monetary policy, because this information reflects market expectations in the future. In this context futures and option prices can be used for extracting market expectations due to their forward-looking nature. Extracted information is mostly concentrated on implied volatility, i.e. volatility computed from a certain option pricing model. For example one can invert Black-Scholes option pricing model to extract implied volatility using observable option prices in the market. Extracting important but unobservable parameters from option prices is not limited to implied volatility, i.e. in a similar manner as implied volatility is derived from option prices, the entire distribution; so called “implied probability distribution” can also be estimated. Hence, the result is the *ex ante* probability distribution that market participants would have expected if they were risk neutral. This means that the estimated implied probability distribution does not take into account the degree of risk aversion of investors. The risk-neutral probability distribution and the associated risk-neutral density function (RND) can describe different characteristics (moments) of the *ex ante* probability distribution, that includes mean, standard deviation, skewness and kurtosis. Hence, variations over time in the implied moments should provide a good indication of changes in the market’s assessment of future developments in the underlying asset.

All techniques for estimating RND functions from option prices are related to the result of Breeden and Litzenberger (1978) and Banz and Miller (1978). They showed that a set of option prices with the same maturity but with different exercise prices can be used to extract the entire probability distribution of the underlying asset at the maturity of the option.

According to Jondeau et al. (2007), the models for estimating RNDs can be divided into two main categories: structural and non-structural. Structural approaches propose a full description for the stock price dynamics, but they are rarely used due to the large number of unknown parameters that should be estimated. Non-structural approaches yield a description of the RND without completely describing the dynamic of price. The non-structural models can be parametric, semi-parametric and non-parametric. Parametric models propose a direct expression for the RND, without referring to any price dynamic. They include mixture of distributions and a wide collection of distributions allowing for higher moments. Semi-parametric and non-parametric models propose some approximation of the true RND. Most commonly used semi-parametric models include Edgeworth expansions and Hermite polynomials, while non-parametric models include spline methods, tree-based methods, maximum entropy principle and kernel regression. However, they do not try to give explicit form for the RND, and “the data is left to speak for itself”

(Jondeau et al., 2007). Therefore, in this paper parametric models will be explained in detail and used in empirical research. Namely, three parametric models; mixture of two log-normals (MLN), Black-Sholes Merton (BSM) and generalized beta (GB2) respectively, are compared to find which one describes the “best” RND for prediction of *ex ante* moments of DAX index. Since there is no consensus in the literature about the “best” parametric model for RND estimation, the goal of this study is to compare the performances of three most commonly used models. Due to different features of proposed models, we believe that they are extreme sensitive to different maturities of call and put options on underlying asset. Hence, the main hypothesis is that *the mixture of two log-normals gives the most accurate implied probability distribution compared to the other parametric models for extracting market expectations in the short run.*

The remainder of the paper is organized as follows. Section 2 presents the literature review. Section 3 describes methodology and method of analysis. Section 4 presents data and empirical results. Section 5 gives interpretation of results with discussion. Finally, some conclusions and directions for future research are provided in Section 6.

2. Literature review

Most commonly used approach for extraction of market expectations rely on Breeden and Litzenberger (1978) results and require the existence of a continuum of European options with the same time to maturity on underlying asset, spanning exercise prices from zero to infinity, and with these assumptions and in the absence of market frictions, it is possible to imply the underlying risk-neutral distribution (RND). Breeden and Litzenberger (1978) first showed how the second partial derivative of the call-pricing function with respect to the exercise price is directly proportional to the RND function. However, this method is based upon the assumption that there exist traded options for many strikes, which is not the case in practice since options contracts are only traded at discretely spaced time points. Because of that, some approximation for the second derivative is necessary and more than one implied distribution can be included depending on the approximation chosen. This has led to development of various techniques, methods and approaches that put more structure into the option prices. Hence, various structural and non-structural models have been developed and studied in the literature. Due to the large number of unknown parameters included, structural models are rarely used for RND estimation. Non-structural models can be divided into three categories: parametric, semiparametric and nonparametric models. Through the literature inspection, there are plenty of papers regarding semiparametric and nonparametric approaches for estimating RND (Jackwerth and Rubinstein, 1996; Andersen and Wagner, 2002; Tavin, 2011; Smith, 2012; Breeden and Litzenberger, 2014; Datta et al, 2014; Malz, 2014), some of them comparing parametric and nonparametric approaches

(Bahra, 1997; Aparicio and Hodges, 1998; Figlewski, 2009; Mizrach, 2010) and only few are dealing only with parametric approaches (Malz, 1997; Bookstaber and McDonald, 1987; Aguilar and Hördahl, 1999; Söderlind, 2000; Gemmill and Saffekos, 2000; Vähämaa, 2005; Cheng, 2010; Markose and Alentom, 2011; Grith and Krätschmer, 2011; Khrapov, 2014). Since the semi- and non-parametric approaches do not give an explicit form for the RND and there is not made any assumption about the RND, in this paper focus is put on parametric approaches, i.e. three most commonly used parametric models for RND estimation: mixture of two log-normals (MLN), Black-Sholes Merton (BSM) and generalized beta (GB2) model. The most commonly studied distribution for option pricing is the mixture of log-normal (MLN) densities, since it can be seen as an extension of the BSM model that uses only one log-normal density. Since the implied variance is not constant at different strike prices, the assumption of a single distribution is too crude, while the mixture of two or more distributions with different means and variances provide considerable flexibility in the distribution. The framework used in MLN follows Ritchey (1990), who notes that a wide variety of shapes may be approximated with a mixture of log-normal distributions. In this paper only two log-normals are used for RND estimation, requiring five parameters to estimate. Moreover, generalized distribution is also used since it is extremely flexible requiring only four parameters to estimate.

Bahra (1997) gives an extensive overview of previous literature on estimation of RNDs. Using both parametric and nonparametric approaches, and considering both advantages and disadvantages of each approach, he estimates RNDs for LIFFE index options for 2 dates in March 1996 with time to expiration in June 1996 (103 and 105 days to expiration), for three-month sterling interest rate for 2 dates in May 1995 with expiration date in September 1995 (133 and 135 days to expiration) and for Philadelphia Stock Exchange (PHLX) currency options. Although there is an enormous contribution of this paper in synthesising various models for RND estimation, they are not compared to each other to provide evidence of superiority of a chosen method.

Aguilar and Hördahl (1999) present examples of estimated implied distributions for interest rates (options on three-month Eurodollar interest rate on 3 dates in 1996 and 1997 with time to maturity of 77 days) and equities (Swedish OMX index on 4 dates in 1998 with time to maturity of 60 days) using a mixture of two log-normal distributions, and for currencies (Swedish krona – SEK/USD and SEK/DEM on 2 dates in 1998 and 1999 with one month to expiration) using Malz (1997) method. They conclude that implied probability distributions provide an indication of the market's assessment concerning the uncertainty of future events, which vary substantially over time. The authors give an overview of parametric methods used to estimate RND for interest rates and equities (mixture of two log-normals) and for currencies (Malz method) and show their good potential. However, it does not compare different parametric approaches.

Söderlind (2000) study how market expectations about future UK monetary policy changed in period from June to November 1992 around events of ERM crisis. It extracts RNDs from daily option prices on the mark-pound exchange rate, 10-year sterling yield to maturity, 3-month sterling interest rate and 3-month mark interest rate, using MLN model and only arguing that it has to be better than BSM model.

Gemmill and Saffekos (2000) use option prices to construct implied (risk neutral) distributions to prove whether they are useful in forecasting market movements and/or revealing the investor sentiment. They estimate the implied distribution as a mixture of two log-normal distributions and test its one-day ahead forecasting performance, using data on LIFFE's FTSE-100 index options from 1987 to 1997. They found that the two log-normal distributions method performs much better than Black-Scholes approach at fitting the observed option prices, by comparing their RMSE (root mean squared errors), but it is only marginally better at predicting out-of-sample prices. Moreover, they do not compare the performances of the methods regarding different maturity horizons.

Jondeau et al. (2007) give much newer overview of various methods for extracting RNDs using both structural and non-structural methods, and parametric, semiparametric and nonparametric approaches. They also compare several techniques to estimate RNDs using data on CAC40 options. The first one is the numerical approximation of RND based on Breeden and Litzemberger (1978) approach, then with Generalized Beta distribution, then using a mixture of two log-normal distributions and finally using Edgeworth expansion around the log-normal distribution. All of them are compared to BSM. The mixture of two log-normals provides the best fit measured by MSE (mean squared errors) and ARE (absolute relative errors). Although it uses different approaches and compares them, it does not provide any information about maturity of options and other information important for estimation purposes.

Given a cross-section of daily option prices, Khrapov (2014) finds the risk-neutral densities, or equivalently the set of parameters, that minimize the option pricing errors using a mixture of log-normal and a generalized beta distribution. This estimation is repeated every Wednesday from 10th January 1996 to 25th July 2012 for VIX options with maturities up to one year to collect the time series of density parameters. It turns out that these parameters are highly predictable. The results are compared using RMSE and IVRMSE for different maturities. It can be concluded that mixture of log-normals always outperforms the BSM model, whereas for generalized beta distribution nothing is concluded.

The abovementioned papers provide an extensive literature review, detailed description of the methods chosen and clear start for further investigation. Some of the papers only give a theoretical proof of advantages of one over the other method; most of them do not even compare the models or do it improperly, even sometimes

giving an advantage to less parsimonious models. Moreover, different maturities are taken into account in some of the papers, but nothing is concluded regarding that. Since previous researches only conclude that none of the method is clearly superior to the others, different approach is needed. Therefore, in this paper different parametric approaches are explained, analysed and compared for extracting implied RND. Moreover, they are compared regarding different maturities in order to conclude which method has the best forecasting ability in different time horizons.

3. Methodology and method of analysis

It should be noted once again that in this paper we do not use structural models as they are rarely used for estimating RND's due to the large number of unknown parameters. Hence non-structural parametric models are used to infer the RND from option prices: Black-Scholes-Merton (BSM), mixture of two log-normals (MLN) and generalized beta model (GB2).

The option pricing model based on the log-normal assumption with constant variance across exercise prices and across maturities has been developed by Black and Scholes (1973) and Merton (1973). Log-normal assumption originates from a geometric Brownian motion (GBM), which means that the price of the underlying asset is log-normally distributed variable, and its returns are normally distributed. Based on these assumptions the price of a European call option can be written as:

$$c = S_T \phi(d_1) - X e^{-r(T-t)} \phi(d_2) \quad (1)$$

Based on put-call parity, if the price of call option is known, the price of a European put option can be written as:

$$p = X e^{-r(T-t)} - S_T + c \quad (2)$$

where

$$d_1 = \frac{\ln\left(\frac{S_T}{X}\right) + \left(r + \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}} \quad (3)$$

$$d_2 = d_1 - \sigma\sqrt{(T-t)}$$

and where c is the price of European call option, p is the price of European put option, S is the price of the underlying asset, X is the strike price or the exercise price, r is the risk-free interest rate, $(T-t)$ is the time to expiration date, σ is the standard deviation of returns, i.e. volatility of the underlying asset, and $\phi(\cdot)$ denotes

cumulative distribution function for the standard normal distribution. Therefore, the price of the call option is a function of five variables, whose values are observable, except for the volatility.

To be able to price an option it is required to estimate the future value of volatility. It can be extracted from the Black-Scholes formula if the price of the option on market is observed, and it is known as implied volatility. In the same way as the implied volatility can be found from the option prices, the whole distribution of the underlying asset can also be estimated from option prices, and it is known as implied probability distribution. From implied probability distribution the moments that describe different characteristics of the distribution can be computed. These moments include mean (expectation), standard deviation (variance), skewness and kurtosis.

If a call option is $c = e^{-r(T-t)} \int_0^{\infty} \max(S_T - X, 0) q(S_T) dS_T$, based on Breeden and Litzenberger (1987), the risk neutral density is given by differentiating this equation twice, with respect to exercise price X :

$$q(S_T) = e^{r(T-t)} \frac{\partial^2 c}{\partial X^2} \quad (4)$$

Using simultaneously observed call and put option prices with the same maturity but with different exercise prices, the parameter vector $\theta \in \Theta$ can be estimated by minimising the sum of squared deviations between the observed (c_j^* , p_j^*) and theoretical prices (c_j , p_j) as:

$$\min_{\theta \in \Theta} \sum_{j=1}^m (c_j - c_j^*)^2 + \sum_{j=1}^n (p_j - p_j^*)^2 \quad (5)$$

The log-normal assumption does not hold in practice, i.e. implied volatilities of options on different exercise prices are not the same, not are volatility smiles for different maturities. Therefore, it is required to use more general option pricing model than BSM. One alternative is to use the mixture of log-normal (MLN) densities, since it can be seen as an extension of the BSM model that uses only one log-normal density. The prices of European call and put options at time t can be written as:

$$c = e^{-r(T-t)} \int_X^{\infty} q(S_T) (S_T - X) dS_T, \quad (6)$$

$$p = e^{-r(T-t)} \int_0^X q(S_T) (X - S_T) dS_T, \quad (7)$$

where $e^{-r(T-t)}$ is the discount factor based on risk-free interest rate r and $q(S_T)$ denotes the risk-neutral density function for the price of the underlying asset on the expiration date T . From the previous equations it can be seen that using observed option prices, we should be able to extract the market's estimate of the probability distribution $q(S_T)$. It should be noted that, in theory, any density functional form $q(S_T)$ can be used in equations (6) and (7), however this paper assumes that a mixture of two log-normal distributions is suitable to describe the underlying distribution for S_T :

$$q(S_T) = \theta L(\alpha_1, \beta_1, S_T) + (1 - \theta) L(\alpha_2, \beta_2, S_T), \quad (8)$$

where

$$L(\alpha_i, \beta_i, S_T) = \frac{1}{S_T \beta_i \sqrt{2\pi}} e^{\left(-\frac{(\ln S_T - \alpha_i)^2}{2\beta_i^2} \right)}, i = 1, 2, \quad (9)$$

and where θ is the weighting parameter that determines the relative influence of two log-normal distributions on the terminal distribution. Parameters α_i and β_i indicate location and dispersion for each log-normal distribution, which determine the mean and variance of the distributions according to:

$$\mu_i = e^{\alpha_i + \frac{\beta_i^2}{2}} \quad (10)$$

and

$$\sigma_i^2 = e^{2\alpha_i + \beta_i^2} (e^{\beta_i^2} - 1). \quad (11)$$

Inserting equations (8) and (9) into equations (6) and (7), the values of the call and put options can be expressed as:

$$c = e^{-r(T-t)} \int_X^\infty \left[\frac{\theta}{S_T \beta_1 \sqrt{2\pi}} e^{\left(-\frac{(\ln S_T - \alpha_1)^2}{2\beta_1^2} \right)} + \frac{1 - \theta}{S_T \beta_2 \sqrt{2\pi}} e^{\left(-\frac{(\ln S_T - \alpha_2)^2}{2\beta_2^2} \right)} \right] (S_T - X) dS_T, \quad (12)$$

$$p = e^{-r(T-t)} \int_X^\infty \left[\frac{\theta}{S_T \beta_1 \sqrt{2\pi}} e^{\left(-\frac{(\ln S_T - \alpha_1)^2}{2\beta_1^2} \right)} + \frac{1 - \theta}{S_T \beta_2 \sqrt{2\pi}} e^{\left(-\frac{(\ln S_T - \alpha_2)^2}{2\beta_2^2} \right)} \right] (X - S_T) dS_T. \quad (13)$$

Using at least five simultaneously observed call and put option prices with the same maturity but with different exercise prices, the five parameters ($\theta, \alpha_1, \beta_1, \alpha_2, \beta_2$) can be estimated by minimising the sum of squared deviations between the observed (c_p^*, p_j^*) and theoretical prices (c_p, p_j) as:

$$\min_{\theta, \alpha_1, \beta_1, \alpha_2, \beta_2} \sum_{j=1}^m (c_j - c_j^*)^2 + \sum_{j=1}^n (p_j - p_j^*)^2 \quad (14)$$

Moreover, according to Bahra (1997), additional information can be exploited by including the forward price as an additional variable in the minimisation problem, where the forward price is the mean of the risk-neutral distribution, given by:

$$F(t, T) = \theta e^{\alpha_1 + \frac{\beta_1^2}{2}} + (1 - \theta) e^{\alpha_2 + \frac{\beta_2^2}{2}} \quad (15)$$

This yields with:

$$\min_{\theta, \alpha_1, \beta_1, \alpha_2, \beta_2} \sum_{j=1}^m (c_j - c_j^*)^2 + \sum_{j=1}^n (p_j - p_j^*)^2 + \left(\theta e^{\alpha_1 + \frac{\beta_1^2}{2}} + (1 - \theta) e^{\alpha_2 + \frac{\beta_2^2}{2}} - e^{-r(T-t)} S \right) \quad (16)$$

subject to

$$\begin{aligned} \beta_1, \beta_2 &> 0 \\ 0 &\leq \theta \leq 1 \end{aligned} \quad (17)$$

Since evaluating equations (12) and (13) is computationally demanding due to the definite integral incorporated in these equations, the minimisation problem in (16) can be solved using closed-form solutions to equations (12) and (13) given as:

$$c = e^{-r(T-t)} \left\{ \theta \left[e^{\alpha_1 + \frac{\beta_1^2}{2}} N(d_1) - XN(d_2) \right] + (1 - \theta) \left[e^{\alpha_2 + \frac{\beta_2^2}{2}} N(d_3) - XN(d_4) \right] \right\}, \quad (18)$$

$$\begin{aligned} p = e^{-r(T-t)} &\left\{ \theta \left[-e^{\alpha_1 + \frac{\beta_1^2}{2}} N(-d_1) - XN(-d_2) \right] + \right. \\ &\left. + (1 - \theta) \left[-e^{\alpha_2 + \frac{\beta_2^2}{2}} N(-d_3) - XN(-d_4) \right] \right\}, \quad (19) \end{aligned}$$

where

$$\begin{aligned} d_1 &= \frac{-\ln X + \alpha_1 + \beta_1^2}{\beta_1}, d_2 = d_1 - \beta_1 \\ d_3 &= \frac{-\ln X + \alpha_2 + \beta_2^2}{\beta_2}, d_4 = d_3 - \beta_2 \end{aligned} \quad (20)$$

The closed-form solution uses the cumulative normal distribution rather than the log-normal density function which eases the numerical computation.

Based on Liu et al (2004) the standard deviation, skewness and kurtosis of a mixture can be derived from:

$$E[S_T^n] = \theta \alpha_1^n e^{\frac{1}{2}(n^2-n)\beta_1^2 T} + (1-\theta) \alpha_2^n e^{\frac{1}{2}(n^2-n)\beta_2^2 T} \quad (21)$$

Bookstaber and McDonald (1987) proposed the generalized beta distribution of the second kind (GB2) to model asset returns. GB2 has four positive parameters $\theta = (a, b, p, q)$ that should be estimated, that permit general combinations of the mean, variance, skewness and kurtosis of a positive random variable, thus enabling the shape of the density to be flexible (Liu, et. al, 2004). The density of the GB2 is given by:

$$q(S_T) = \frac{a}{b^{ap} B(p, q)} s^{ap-1} \left[1 + \left(\frac{s}{b} \right)^a \right]^{-(p+q)}, \text{ for } s > 0, \quad (22)$$

with $B(p, q) = \Gamma(p)\Gamma(q)/\Gamma(p, q)$.

The density is risk-neutral when

$$S_T = e^{-r(T-t)} \frac{bB\left(p + \frac{1}{a}, q - \frac{1}{a}\right)}{B(p, q)}, \quad (23)$$

and its moments are

$$E[S_T^n] = \frac{b^n B\left(p + \frac{n}{a}, q - \frac{n}{a}\right)}{B(p, q)} \text{ for } n < aq. \quad (24)$$

Therefore, GB2 is a function of four parameters: a , b , p and q . These parameters work interactively in determining the shape of the distribution. The b is a scale parameter, which stretches out the distribution. For large values of a parameter, the scale parameter has a direct relation to the mean of the distribution, i.e. the parameter b will approximate the mean. A change in the value of b will have an effect on the height of the density. The parameters a , p and q are shape parameters. The a parameter determines the “speed” with which the tails of the density function approach to the X-axis. A higher value of parameter a implies a quicker approach to the axis. The product aq has a direct effect on the fatness of the distribution. Also, parameters p and q will dictate the skewness of the distribution.

Generalized distribution is extremely flexible, and it includes a number of well known distributions as special limiting cases. Distributions with large, even infinite higher moments can be specified by the choice of parameters. This flexibility

allows a direct representation of different degrees of fat tails in the distribution (Bookstaber and McDonald, 1987).

4. Data and empirical analysis

In empirical researches, there are often used prices collected at the end of the trading day or the average of the bid and ask price. Moreover, it is common to use the out-of-the money options, because of their greater liquidity. The same is done in this research. Data sets includes averages of the last bid and the last ask options prices, both for calls and puts on the same strikes, in period from July 2014 to July 2015. However, the selection of dates to include is not set by chance. After the observation of DAX index movements, and before the research is conducted, it was important to select dates at which the estimation would be performed. Dates around peaks and bottoms are selected. Moreover, each selected date corresponds to the third Friday of each month, i.e. expiration day of option prices. In the observed period from July 2014 to July 2015, DAX index was on average 10412.94 points, reaching its minimum value of 8571.95 in October 2014, and its maximum value of 12374.73 in April 2015. Time series of DAX index prices is presented by Figure 1. Moreover, the RND were estimated for one and two months in advance, i.e. time to expiration is both 28 (and 35) and 63 (or 56) days depending on number of days in a month. Therefore, RNDs are estimated for 11 months, for 2 different maturity horizons and with 3 different parametric models, yielding with 66 different RNDs. Three models will be compared to conclude which method has the best forecasting accuracy in different time horizons. The research is conducted in “R” software using “RND” package.

Figure 1: Closing prices of DAX index from July 2014 to July 2015



Source: <http://finance.yahoo.com/q/hp?s=%5EGDAXI+Historical+Prices>
[25th July 2015]

Since all three approaches are parametric, i.e. mixture of two log-normal distributions (MLN), Black-Scholes-Merton model (BSM) and generalized beta distribution (GB2), estimated parameters of each model are presented in Tables 1-3 for three chosen expiration dates, 17th October 2014, 16th January 2015, and 19th June 2015 respectively.

Table 1: RND's parameters estimates for 1 month and 2 months maturity horizons with *ex ante* moments extracted at expiration date 17th October 2014

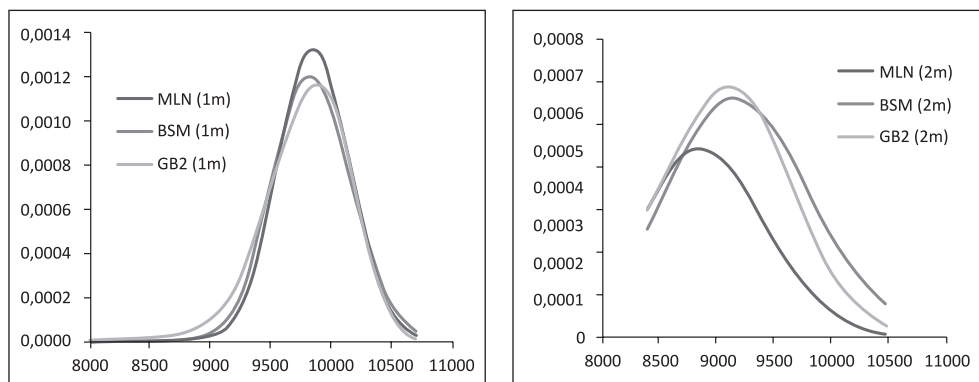
Model (T-t)	BSM		MLN		GB2	
	1m	2m	1m	2m	1m	2m
α	9.194294	9.12535				
β	0.033633	0.06805				
θ			0.993439	0.315686		
α_1			9.195662	9.195227		
α_2			8.790679	9.091337		
β_1			0.030309	1.81E-08		
β_2			0.194057	0.061025		
a					19.13343	20.00000
b					12026.84	9548.811
p					2.514844	0.991215
q					104.2054	2.385819
μ	9834.54	9267.05	9851.40	9048.34	9800.27	9180.06
σ	437.99	519.50	431.80	570.80	438.17	517.37
α_3	1.39	-0.61	1.39	-1.14	1.15	-0.96
α_4	2.41	2.11	2.34	1.89	2.21	2.14
MSE	768.20	6631.00	192.00	3472.00	894.40	7256.00
DM test	-10.1561***		-3.6758***		-3.7529***	
BSM			-3.8055***		-0.623	
MLN					-4.4169***	

Note: *, **, *** represent significance level of 1%, 5% and 10% at which the null hypothesis of Diebold-Mariano test (DM) is rejected.

Source: Authors' calculations

All tables, in addition to the estimated parameters of RND's for 1 month and 2 months maturity horizons, include also *ex ante* moments extracted from each RND, i.e. mean μ , standard deviation σ , skewness α_3 and kurtosis α_4 of DAX index prices at chosen expiration date.

Figure 2: Estimated RNDs for 1 month and 2 months maturity horizons on 17th October 2014



Source: Authors' calculations

In each Table 1-3 six RNDs have been estimated based on two expiration dates and using three parametric models. The same is done for 8 more expiration dates, but not presented here due to the lack of space (results are available upon request).

Table 2: RND's parameters estimates for 1 month and 2 months maturity horizons with *ex ante* moments extracted at expiration date 16th January 2015

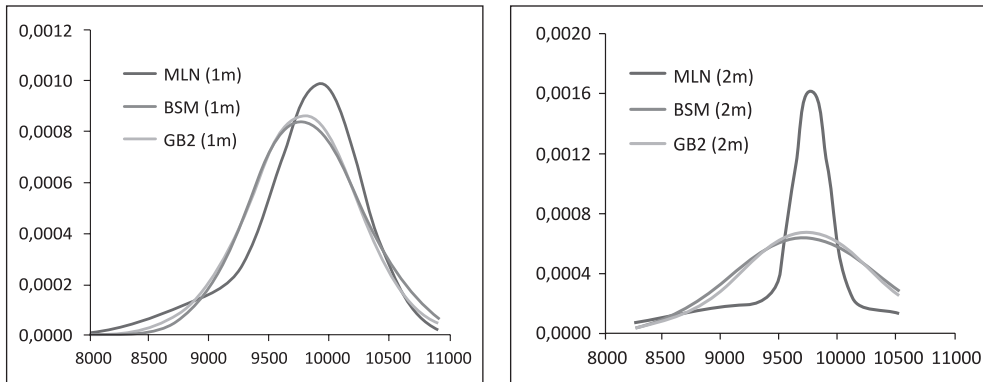
Model (T-t)	BSM		MLN		GB2	
	1m	2m	1m	2m	1m	2m
α	9.189182	9.18157				
β	0.048679	0.064523				
θ			0.201139	0.506042		
α_1			9.123738	9.173626		
α_2			9.20441	9.186189		
β_1			0.058149	0.099549		
β_2			0.034469	0.014332		
a					20	20
b					10032.04	9847.41
p					2.032562	1.429859
q					3.189957	1.772992
μ	9770.54	9523.59	9784.38	9598.07	9745.99	9536.00
σ	443.01	462.05	480.90	416.97	455.62	459.46
α_3	-0.11	-0.27	-0.86	-1.24	-0.31	-0.37
α_4	2.68	2.64	3.90	4.38	3.06	2.78
MSE	30.720	989.50	58.100	757.300	148.8000	979.400
DM test	-1.7245**		-1.4516*		-1.7408**	
BSM			-5.6849***		-8.7384***	
MLN					-5.77***	

Note: *, **, *** represent significance level of 1%, 5% and 10% at which the null hypothesis of Diebold-Mariano test (DM) is rejected.

Source: Authors' calculations

However, in Tables 1-3, and Figures 2-4 one can notice that there are differences between extracted three risk-neutral densities. Therefore, the purpose of the present paper is to investigate which model and at which maturity horizon fits the future distribution of DAX index most accurately.

Figure 3: Estimated RNDs for 1 month and 2 months maturity horizons on 16th January 2015



Source: Authors' calculations

As a comparison criterion, mean square error (MSE) is used. MSE was calculated as the mean square difference between observed call and put option prices and expected (theoretical) call and put option prices obtained from three parametric models for the same strikes. Furthermore, Diebold-Mariano test (DM) is used to test which model has lower MSE. Null hypothesis of DM test is that two forecasts have the same accuracy. Since the forecasts values in the sample are the same as fitted values then DM test in this paper is used to test which model has lower mean square error. Model that has significantly lower MSE is considered as most appropriate.

Table 3: RND's parameters estimates for 1 month and 2 months maturity horizons with *ex ante* moments extracted at expiration date 19th June 2015

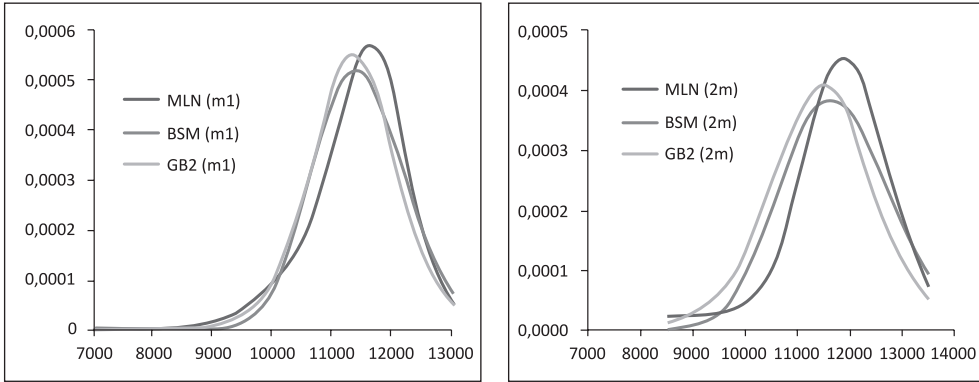
Model (T-t)	BSM		MLN		GB2	
	1m	2m	1m	2m	1m	2m
α	9.346663	9.369684				
β	0.067334	0.088918				
θ			0.303163	0.109155		
α_1			9.288703	9.188557		
α_2			9.370498	9.387833		
β_1			0.077069	0.142239		
β_2			0.048859	0.067263		
a					20	20
b					11540.48	11714.69
p					1.306704	0.803193
q					1.66167	1.061787
μ	11422.54	11550.45	11446.08	11667.88	11322.75	11351.40
σ	705.05	849.26	752.90	874.94	720.21	901.08
α_3	-0.11	-0.28	-0.60	-0.92	-0.27	-0.38
α_4	2.63	2.60	3.25	4.10	3.06	2.86
MSE	731300.0	1558	24.1800	148.80	570.000	3006.000
DM test	-9.7392***		-9.3281***		-10.8407***	
BSM			-9.769***		-9.7635***	
MLN					-13.5996***	

Note: *, **, *** represent significance level of 1%, 5% and 10% at which the null hypothesis of Diebold-Mariano test (DM) is rejected.

Source: Authors' calculations

Firstly, the testing procedure is computed to compare the performance of RND's extracted between 1 month and 2 months maturity horizons within each model. Secondly, three proposed models are used for pairwise comparison (MLN vs. BSM, MLN vs. GB2 and BSM vs. GB2) only in short-run as previously has been concluded that the short-run forecasts yield better results with smaller MSE.

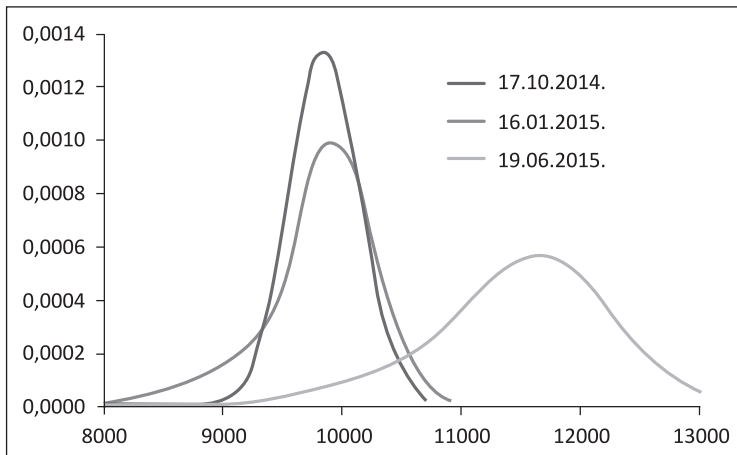
Figure 4: Estimated RNDs for 1 month and 2 months maturity horizons on 19th June 2015



Source: Authors' calculations

Changes in the implied moments, extracted from RND's, between two successive time points should provide valuable information of changes in the market's assessment of future developments in the underlying asset. According to the MLN model that has most accurate predictive ability within one month maturity horizon, three RND's are compared to describe these changes (Figure 5).

Figure 5: RNDs at 3 expiration dates based on MLN model (maturity horizon 1m)



Source: Authors' calculations

For ease interpretation of changes in the market expectations of DAX index an implied (*ex ante*) moments at three chosen expiration dates are also presented in Table 4.

Table 4: Implied (*ex ante*) moments extracted from estimated RND's at three chosen expiration dates (MLN model with maturity horizon of one month)

Implied moments	17/10/2014	16/01/2015	19/06/2015
μ	9851.40	9784.38	11446.08
σ	431.80	480.90	752.90
α_3	1.39	-0.86	-0.60
α_4	2.34	3.90	3.25

Source: Authors' calculations

Changes in the implied moments in observed sample period reveal that the value of DAX index is at the higher level with higher implied volatility and that implied (risk-neutral) distribution is more negatively skewed on the left tail then on the right tail even both tails are heavier compared to the implied distribution at previous expiration dates.

5. Interpretation of results and discussion

The results indicate that the short-run forecasts yield better results with smaller MSE, i.e. within each estimated model (BSM, MLN and GB2) the null hypothesis of DM test can be rejected in favour of one-sided alternative. The null hypothesis is that the two models have the same forecast accuracy (Diebold and Mariano, 1995).

The results show superiority of MLN model in short run against BSM and GB2, i.e. the short-run MLN forecasts have lower MSE than short-run forecasts of both BSM and GB2 models. The same results were obtained and confirmed for 8 more expiration dates, but not presented here due to the lack of space. Moreover, BSM model shows rather divergent results, yielding in some cases with lower and in some cases with extremely higher MSE.

Namely, on third Friday in September the value of DAX index was still recovering from its' plunge in August. Therefore, the market sentiment was optimistic on 17th October 2014 yielding with higher expected value of DAX index one month ahead, lower standard deviation as an indication of smaller uncertainty and positive skewness which indicates that market perceives the probability of positive outcomes to be higher than the probability of negative outcomes.

For expiration date 16.1.2015. results indicate similar value of the implied mean but higher implied standard deviation (volatility) with implied asymmetry concentrated on the left tail of distribution ($\alpha_3 < 0$). As implied kurtosis is increasing a distribution has heavier tails (Figure 5). It is important to note that the kurtosis is a measure of fat tails of the distribution, i.e. leptokurtic distribution has fat-tails ($\alpha_4 > 3$) regarding the curvature of distribution.

The same changes in the implied moments can be noticed at expiration date 19th June.2015, which means that the value of DAX index is at the higher level with higher implied volatility. Distribution is also more negatively skewed on the left tail then on the right tail even both tails are heavier compared to the implied distribution at expiration date 16th January 2015. This reveals that the market participants perceive great uncertainty with the development of the DAX index during the life of the options with higher probability of negative outcomes, i.e. the market sentiment was pessimistic.

6. Conclusion

Empirical results confirm the hypothesis that the mixture of two log-normals gives the most accurate implied probability distribution compared to the other parametric models for extracting market expectations in the short run. The results reveal that no matter which parametric model is used, all of them give better short-term forecasts. In pairwise comparison for short-term prediction, mixture of two log-normals approach is superior to the others according to the DM test. Moreover, MLN model has proven to be very flexible, which means that it is possible to obtain a wide variety of different implied distributions, and therefore it can capture commonly observed characteristics of financial assets, such as asymmetries and “fat-tails” in implied probability distribution. In this way, departures from the assumptions of the Black-Scholes model can be taken into account in a relatively simple way. The results also show how RND function responds to the arrival of new information, and how market risk preferences vary over time. Results show that the risk neutral density for the DAX index is far from the lognormal density assumed by the Black-Scholes model, i.e. it is strongly negatively skewed and fat-tailed. We also found that when the underlying index moves, the RND not only moves along with the index, but it also changes shape. The present paper contributes to the existing literature in a several ways. First, it reveals at which maturity horizon the prediction of RND is most accurate given by the “best” parametric model. It should be noted that parametric models are neglected in the literature hitherto compared to the nonparametric ones. Second, it introduces the Diebold-Mariano test of performances comparison of the RND estimators. Third, it covers the most recent period which have not been in the center of research interest. Fourth, investors are able to apply the chosen RND estimator and obtain the “most accurate” implied probability distribution for extracting market expectations. The limitation of this research is that results are carried out only for DAX index, i.e. developed stock market. Namely, the same research cannot be compute for other markets due to infrequent option trading and illiquid option contracts. Therefore options prices can be less informative about market expectations; especially at emerging markets where such data are not available. Moreover, option trading does not exist in some countries. Further,

the value of the true variance, skewness and kurtosis are unknown to us (cannot be observed) and the implied higher order moments can only be compared with their realized counterparts that requires intraday observations (high frequency data). Therefore, direction for the future research is to compare implied higher moments with realized higher moments which can be used as best proxies of unknown distribution characteristics of the underlying asset.

Extracting such information is of great interest both to market participants and policy-makers such as central banks. Market participants are interested in forecasting future volatility, since it is an important variable in portfolio management. To a central bank, it is important to be able to interpret market participants' expectations concerning future monetary policy. A central bank's ability to understand market expectations is of great importance to compare the central bank's view of the economy with the market's view, but also because market expectations can be self-fulfilling to some extent.

References

- Aguilar, J., Hördahl, P. (1999) "Option prices and market expectations", *Sveriges Riksbank Quarterly Review*, Vol. 1, pp. 43–70., http://www.riksbank.se/Upload/Dokument_riksbank/Kat_publicerat/Artiklar_PV/qr99_1_artikel3.pdf (4.12.2015).
- Andersen, A.B., Wagener, T. (2002) "Extracting risk neutral probability densities fitting implied volatility smiles: some methodological points and an application to the 3M Euribor futures option prices", ECB Working Paper No. 198., <http://ssrn.com/abstract=359060> (4/12/2015).
- Aparicio, S.D., Hodges, S. (1998) "Implied Risk-Neutral Distribution: A comparison of Estimation Methods", <http://www2.warwick.ac.uk/fac/soc/wbs/subjects/finance/research/wpaperseries/1998/98-95.pdf> (5/10/2015).
- Bahra, B. (1997) "Implied risk-neutral probability density functions from option prices: theory and application", Bank of England, Working Paper, No. 66, doi: 10.2139/ssrn.77429.
- Banz, R.W., Miller, M. (1978) "Prices for State-contingent Claims: Some Estimates and Applications", *The Journal of Business*, Vol. 51, No. 4, pp. 653–672, doi: 10.1086/296026.
- Black, F., Scholes, M. S. (1973) "The Pricing of Options and Corporate Liabilities", *Journal of Political Economy*, Vol. 81, pp. 637–659, doi: 10.1086/260062.
- Bookstaber, R.M., McDonald, J.B. (1987) "A General Distribution for Describing Security Price Returns", *The Journal of Business*, Vol. 60, No. 3, pp. 401–424, doi: 10.1086/296404.

- Breeden, D.T., Litzenberger, R.H. (1978) "Prices of State-Contingent Claims Implicit in Option Prices", *The Journal of Business*, Vol. 51, No. 4, pp. 621–651, doi: 10.1086/296025.
- Breeden, D.T., Litzenberger, R.H. (2014) "Central Bank Policy Impacts on the Distribution of Future Interest Rates" (February 20, 2014) http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2642363 (30.10.2015), doi: 10.2139/ssrn.2642363.
- Datta, D.D., Londono, J.M. and Ross, L.J. (2014) "Generating Options-Implied Probability Densities to Understand Oil Market Events", Board of Governors of the Federal Reserve System, International Finance Discussion Paper, Number 1122, October, <http://www.federalreserve.gov/econresdata/ifdp/2014/files/ifdp1122.pdf> (4.15.2015).
- Diebold, F. X., Mariano, R. S. (1995) "Comparing Predictive Accuracy", *Journal of Business and Economic Statistics*, No. 13, pp. 253–263, doi: 10.1080/07350015.1995.10524599.
- Figlewski, S. (2009) "Estimating the Implied Risk Neutral Density for the U.S. Market Portfolio", Volatility and Time Series Econometrics: Essay in Honor of Robert F. Engle, Tim Bollerslev, Jeffrey R. Russell, Mark Watson, Oxford, UK: Oxford University Press, 2008; NYU Working Paper No. FIN-08-004, doi: 10.1093/acprof:oso/9780199549498.003.0015.
- Gemmill, G., Saflekos, A. (2000) "How useful are implied distributions? Evidence from stock index options", *The Journal of Derivatives*, Vol. 7, No. 3, pp. 83–91, doi: 10.3905/jod.2000.319123.
- Jackwerth, J.C., Rubinstein, M. (1996) "Recovering Probability Distribution from Option Prices", *The Journal of Finance*, Vol. 51, No. 5, pp. 1611–1631, doi: 10.2307/2329531.
- Jondeau, E., Poon, S and Rockinger, M. (2007) "Financial modelling under Non-Gaussian Distributions", Springer-Verlag, London, doi: 10.1007/978-1-84628-696-4.
- Khrapov, S. (2014) "Option pricing via Risk-neutral Density Forecasting", 2nd International Workshop on "Financial Markets and Nonlinear Dynamics" (FMND), June 4-5, 2015, Paris (France) https://editorialexpress.com/cgi-bin/conference/download.cgi?db_name=SNDE2015&paper_id=110 (16/9/2015).
- Liu, X., Shackleton, M.B., Taylor, S.J. and Xu, X. (2004) "Closed-form transformations from risk-neutral to real-world distributions", Essex Finance Centre Discussion Paper No. 04–05, doi: 10.2139/ssrn.424420.
- Malz, A.M. (1997) "Option-implied Probability Distributions and Currency Excess Returns", Federal Reserve Bank of New York Staff Reports, No. 32, doi: 10.2139/ssrn.943500.
- Malz, A.M. (2014) "A Simple and Reliable Way to Compute Option-Based Risk-Neutral Distributions", Federal Reserve Bank of New York Staff Reports, No. 677, doi: 10.2139/ssrn.2449692.

- Markose, S., Alentom, A. (2011) “The Generalized Extreme Value (GEV) Distribution, Implied Tail Indeks and Option Pricing”, *The Journal of Derivatives*, Vol. 18, No. 3, pp. 35–60, doi: 10.3905/jod.2011.18.3.035.
- Merton, R. C. (1973) “Theory of Rational Option Pricing”, *The Bell Journal of Economics and Management Science*, Vol. 4, No. 1, pp. 141–183, doi: 10.2307/3003143.
- Mizrach, B. (2010) “Estimating Implied Probabilities from Option Prices and the Underlying”, *Handbook of Quantitative Finance and Risk Management*, pp. 515–529, doi: 10.1007/978-0-387-77117-5_35.
- Ritchey, R.J. (1990) “Call option valuation for discrete normal mixtures”, *Journal of Financial Research*, Vol. 13, No. 4, pp. 285–296, doi: 10.1111/j.1475-6803.1990.tb00633.x.
- Smith, T. (2012) “Option-implied probability distributions for future inflation”, *Bank of England Quarterly Bulletin 2012 Q3*, pp. 224-233, <http://ssrn.com/abstract=2150092> (5/12/2015).
- Söderlind, P. (2000) “Market expectations in the UK before and after the ERM crisis”, *Economica*, No. 67, pp. 1–18, doi: 10.1111/1468-0335.00192.
- Tavin, B. (2011) “Implied distribution as a function of the volatility smile”, *Bankers Markets and Investors*, No. 119, pp. 31–42, doi: 10.2139/ssrn.1738965.
- Vähämaa, S. (2005) “Option-implied asymmetries in bond market expectations around monetary policy actions of the ECB”, *Journal of Economics and Business*, No. 57, pp. 23–38, doi: 10.1016/j.jeconbus.2004.07.003.

Procjena tržišnih očekivanja na temelju funkcije gustoće neutralne na rizik¹

Josip Arnerić², Zdravka Aljinović³, Tea Poklepović⁴

Sažetak

Cilj ovog rada je istražiti koji se od predloženih parametarskih modela za procjenu funkcije gustoće neutralne na rizik, Black-Scholes Merton model, kombinacija dvije log-normalne distribucije ili generalizirana beta distribucija, najbolje prilagođava stvarnim podacima. Najbolji model od navedenih dalje se koristi kako bi se opisale karakteristike (momenti) *ex ante* distribucije vjerojatnosti. Rezultati empirijskog istraživanja sugeriraju da bez obzira koji parametarski model koristili, najbolja prilagodba podacima je uvijek s kraćim rokom dospjeća. Uspoređujući modele u kratkom roku, model koji kombinira dvije log-normalne distribucije ima statistički značajno najnižu srednju kvadratnu pogrešku. Na temelju usporedbe rezultata u parovima glavni je zaključak da je kombinacija dvije log-normalne distribucije superiorna u odnosu na ostale parametarske modele te se pokazala fleksibilnom u opisivanju opaženih karakteristika promatrane financijske imovine, kao što su asimetrija i zadebljani „repovi“ implicirane distribucije vjerojatnosti.

Cljučne riječi: tržišna očekivanja, funkcija gustoće neutralna na rizik, kombinacija log-normalnih distribucija, Black-Scholes Mertonov model, generalizirana beta distribucija, vrijeme do dospjeća, opcije na DAX indeks

JEL klasifikacija: C14, C58, G1

¹ Ovaj rad je financirala Hrvatska zaklada za znanost projektom “Mjerenje volatilnosti, modeliranje i prognoziranje” (5199).

² Docent, Sveučilište u Zagrebu, Ekonomski fakultet, Trg. J. F. Kennedyja 6, 10000 Zagreb, Hrvatska. Znanstveni interes: ekonometrijske metode i modeli, financijski vremenski nizovi i volatilnost, VAR modeli i kointegracija, GARCH i MGARCH modeli, stohastički procesi i upravljanje rizicima. Tel: +385 1 238 3361. Fax: +385 1 233 5633. E-mail: jarneric@efzg.hr. Web stranica: <http://www.efzg.unizg.hr/default.aspx?id=15332>. (kontakt osoba).

³ Redoviti profesor, Sveučilište u Splitu, Ekonomski fakultet, Cvite Fiskovića 5, 21000 Split, Hrvatska. Znanstveni interes: matematičke financije – financijsko modeliranje, matematička ekonomija te višekriterijsko odlučivanje. Tel: +385 21 430 644. Fax: +385 21 430 701. E-mail: zdravka.aljinovic@efst.hr. Web stranica: <http://www.efst.unist.hr/content.php?k=fakultet&p=69&osoba=lil>.

⁴ Doktorand, Sveučilište u Splitu, Ekonomski fakultet, Cvite Fiskovića 5, 21000 Split, Hrvatska. Znanstveni interes: statističke i ekonometrijske metode u poslovanju, financijska ekonometrija, vremenski nizovi i neuronske mreže. Tel: +385 21 430 761. Fax: +385 21 430 701. E-mail: tpoklepo@efst.hr. Web stranica: <http://www.efst.unist.hr/content.php?k=fakultet&p=69&osoba=tpoklepo>.