

*Piotr Płuciennik*

*Adam Mickiewicz University, National Bank of Poland*

## Forecasting Financial Processes by Using Diffusion Models

**A b s t r a c t.** Time series forecasting is one of the most important issues in the financial econometrics. In the face of growing interest in models with continuous time, as well as rapid development of methods of their estimation, we try to use the diffusion models to modeling and forecasting time series from various financial markets. We use Monte-Carlo-based method, introduced by Cziraky and Kucherenko (2008). Received forecasts are confronted with those determined with the commonly applied parametrical time series models.

**K e y w o r d s:** diffusion models, ex-post forecasts, Monte-Carlo simulation, the GARCH model, the ARIMA model, unit-root.

### 1. Introduction

Models with continuous time and its particular case – diffusion models are exceptionally important class of models. On the developed financial markets there are available quotations containing full information about transaction prices, so called tick-by-tick data. It provides natural motivation to applying diffusion models or jump diffusion models to examination of financial instruments price series. Diffusion models were initially used to short-term rate modeling (Merton, 1973; Vašiček, 1977; Cox, at al. 1985). They gained in importance in the early 70s, when Black and Scholes (1973) introduced European call and put option pricing model in which the underlier price was modeled with simple diffusion model called Geometric Brownian Motion. In the following years many modification of Black-Scholes models were introduced. Merton (1974) supposed additionally that the risk-free rate is modeled by diffusion model, and Heston (1993) assumed that the volatility process is described by mean reverting diffusion models. Forms of Black-Scholes formula were introduced for american options, term options and even volatility index options, which had just appeared on the market.

Other application areas of models with continuous time are issues concerning modeling and forecasting interest rates (term structure) and valuation of very complicated derivatives based on debt instruments. This subject was brought up among others by Jagannathan et al. (2004), which applied multidimensional CIR models to caps and swaptions pricing, Tamba (2006), which used Hull-White diffusion model to Bermudian swaption pricing, and Mannolini, Mari, Renò (2008), which priced caps and floors by extended CIR models.

The aforementioned derivatives are of outstanding significance. They played the crucial role in risk management and in aggressive investment strategies. The first hedging strategies were proposed by Black and Scholes (1973). Nowadays there are strategies which allow to hedge the positions in swaption (Javaheri et al., 2004; Howison et al., 2004) and VIX Options (Psychoyios, Skiadopoulos 2006; Sepp 2008; Broadie, Jain 2008).

In the following article we use diffusion models to forecast the logarithmic levels of DAX, CAC40, NASDAQ and WIG20 indexes. The parameter estimates were obtained by modern Phillips and Yu (2009) method and more classical, introduced by Hansen (1982), Generalized Method of Moments (GMM) with covariance matrix, estimated by using Bartlett kernels, as a weight matrix (Newey, West, 1987). We determine the forecasts by using Monte-Carlo methods and compare its quality with the forecasts which we obtained by using popular parametrical time series models.

## 2. Models

We use popular diffusion models to describe logarithmic prices of financial instruments. The diffusion models were originally used to describe the evolution of short-term rates. Their significant feature is the mean reverting property.

The most simple diffusion model – Vašíček (1977) model – assume that the price process is modeled by the following stochastic differential equation

$$dX_t = \kappa(\mu - X_t)dt + \sigma dB_t,$$

with initial condition  $X_{t_0} = x_0$ . Parameters  $\kappa$ ,  $\mu$  and  $\sigma$  are strictly positive. Parameter  $\mu$  can be interpreted as a long term mean level,  $\kappa$  as a speed of reversion, and  $\sigma$  as a instantaneous volatility.

Cox, Ingersoll and Ross (1985) introduced model called CIR, which is an extension of Vašíček model. The  $X_t$  evolution is described by the formula

$$dX_t = \kappa(\mu - X_t)dt + \sigma \sqrt{X_t} dB_t,$$

with initial condition  $X_{t_0} = x_0$ . Parameters  $\kappa$ ,  $\mu$  and  $\sigma$  are strictly positive, and have the same interpretation as in Vašíček model. The square root in the

diffusion function allows to avoid the possibility of nonpositive values of  $X_t$ , provided that the condition  $2\kappa\mu > \sigma^2$  is met.

Chan at al. (1992) introduced model called CKLS. Authors assume that  $X_t$  evolution is described by the following diffusion models

$$dX_t = \kappa(\mu - X_t)dt + \sigma X_t^\beta dB_t,$$

with initial condition  $X_{t_0} = x_0$ . Similarly as in Vašíček and CIR model,  $\kappa$ ,  $\mu$  and  $\sigma$  are strictly positive. An additional  $\beta$  parameter is called the elasticity of variance parameter, and  $\beta \in [0,1]$ . By simply placing the appropriate restrictions on the four parameters  $\kappa$ ,  $\mu$ ,  $\sigma$  and  $\beta$  we can obtain 7 other diffusion models – among others the Vašíček and the CIR model. (see Chan at al., 1992).

### 3. Determination of One-Day *Ex-Post* Forecast from Diffusion Models

Denote the  $l$ -step forecast of  $X_{t+l}$  as  $\hat{X}_t(l)$ . Assuming that the minimum squared error is the loss function, the forecast  $\hat{X}_t(l)$  is the random variable chosen such that

$$\mathbb{E}[X_{t+l} - \hat{X}_t(l)]^2 \leq \min_g \mathbb{E}[X_t - g(X_t, \dots, X_1)]^2,$$

where  $g(X_t, \dots, X_1)$  is measurable function towards  $\sigma$ -algebra generated by the information available up to time  $t$  inclusive. We can show that

$$\hat{X}_t(h) = \mathbb{E}(X_{t+h} | \mathcal{F}_t).$$

Therefore, if we assume that process is described by diffusion models and length of one step equals  $h$ , then

$$\begin{aligned} \hat{X}_t(h) &= \mathbb{E} \left[ X_0 + \int_0^{t+h\delta} \mu(X_s, s, \hat{\theta}_1) ds + \int_0^{t+h\delta} \sigma(X_s, s, \hat{\theta}_2) dB(s) \middle| \mathcal{F}_t \right] \\ &= \mathbb{E} \left[ X_t + \int_t^{t+h\delta} \mu(X_s, s, \hat{\theta}_1) ds + \int_t^{t+h\delta} \sigma(X_s, s, \hat{\theta}_2) dB(s) \right] \\ &= X_t + \int_t^{t+h\delta} \mu(X_s, s, \hat{\theta}_1) ds + \mathbb{E} \int_t^{t+h\delta} \sigma(X_s, s, \hat{\theta}_2) dB(s), \end{aligned}$$

where  $\hat{\theta}_1$  and  $\hat{\theta}_2$  are parameter estimates vectors of drift and diffusion respectively, obtained on the quotations up to time  $s$ . As we know,

the quotations are available only in discrete intervals. Consequently, we approximate the forecast  $\hat{X}_t(l)$  by the Euler scheme of the form

$$\hat{X}_t^{(\delta)}(h) = X_t^{(\delta)} + \sum_{k=0}^{h-1} \mu(X_{t+k|n}^{(\delta)}, t, \hat{\theta}_1) \delta, + \sum_{k=0}^{h-1} \sigma(X_{t+k|n}^{(\delta)}, t, \hat{\theta}_1) \sqrt{\delta} \varepsilon_{t+k}^{(\delta)}. \quad (1)$$

Moreover, for all  $t$  holds  $X_t^{(\delta)} = X_{[t/\delta]}$ . Cziraky i Kucherenko (2008) obtain estimates of  $\hat{X}_t^{(\delta)}(l)$  by repeating the above recursion using  $N$  independent realizations of innovations vectors  $(\varepsilon_{n+1}^{(\delta)}, \dots, \varepsilon_{n+k}^{(\delta)})$ , and for any realization they determine the trajectory (1) with initial condition  $X_t^{(\delta)}$ . The MC estimator of  $\hat{X}_t(l)$  is then given by the average of the last elements of every trajectory.

One-step ex-post forecasts are obtained by fitting the model using data up to time  $T$ , and then computing the usual fitted equation and residuals for periods  $T+1$  to  $T+F$ , with additional assumption that the quotation which precede the forecast is known.

#### 4. The Data

We take into account daily levels of German DAX, French CAC40, American NASDAQ and Polish WIG20 indexes from the period 2. January 2001 to 29. December 2006. In both series we observe the logarithmic trend. During the mentioned period the trend is growing. Therefore, we decide to model logarithmic levels of the indexes. The significance of the trend is then marginalized. The descriptive statistics are given in the following Table 1.

Table 1. Descriptive statistics of the logarithmic levels of indexes CAC40, DAX, NASDAQ and WIG20 from the period 2. January 2001 to 29. December 2006

| Time series | obs. number | mean    | std. dev. | skewness | kurtosis | min.   | max.   |
|-------------|-------------|---------|-----------|----------|----------|--------|--------|
| CAC40       | 1534        | 8.3138  | 0.19922   | -0.1759  | 2.0605   | 7.7845 | 8.6993 |
| DAX         | 1528        | 8.40231 | 0.24511   | -0.3745  | 2.3597   | 7.6976 | 8.824  |
| NASDAQ      | 1554        | 7.55651 | 0.1818    | -0.8033  | 3.0256   | 7.0158 | 7.9583 |
| WIG20       | 1546        | 7.4555  | 0.34958   | 0.4318   | 2.0328   | 6.8979 | 8.1745 |

#### 5. Empirical Research

In the following section we present results of one day ex-post forecasts quality testing of examined time series, which we obtain from diffusion model. The model is estimated by using the modern two-stage Phillips and Yu (2009) method and the Generalized Method of Moments. We obtain the parameter estimation of diffusion models by using our own procedures in Matlab (Phillips-Yu method) and by using the Matlab libraries by Cliff (2003) (GMM). The values of parameters estimations are given in Table 2.

Table 2. Parameter estimations of diffusion models obtained by using the Phillips and Yu method and the Generalized Method of Moments for logarithmic levels of CAC40, DAX, NASDAQ and WIG20 indexes

| Estimation method: |           | Phillips and Yu |          |           |           |
|--------------------|-----------|-----------------|----------|-----------|-----------|
| model              | parameter | CAC40           | DAX      | NASDAQ    | WIG20     |
| Vašíček            | $\kappa$  | -0.10636        | 0.29613  | 0.0000258 | 0.0012536 |
|                    | $\mu$     | 8.1772          | 8.4126   | 7.5829    | 7.4681    |
|                    | $\sigma$  | 0.16926         | 0.21296  | 0.12692   | 0.16780   |
| CIR                | $\kappa$  | -0.10702        | 0.29529  | 0.081467  | 0.00776   |
|                    | $\mu$     | 8.1667          | 8.4223   | 7.5632    | 7.3712    |
|                    | $\sigma$  | 0.058477        | 0.072836 | 0.045667  | 0.052088  |
| CKLS               | $\kappa$  | 0.41333         | 0.51308  | 0.081748  | -0.85398  |
|                    | $\mu$     | 8.3644          | 8.4174   | 7.5632    | 6.557     |
|                    | $\sigma$  | 5.5273          | 5.3016   | 0.72314   | 0.4757    |
|                    | $\beta$   | -4.8234         | -4.2147  | -0.85398  | -0.51843  |
| Estimation method: |           | GMM             |          |           |           |
| model              | parameter | CAC40           | DAX      | NASDAQ    | WIG20     |
| Vašíček            | $\kappa$  | 0.034582        | 0.042929 | 0.91407   | -0.12749  |
|                    | $\mu$     | 8.3644          | 8.4174   | 7.5811    | 6.5984    |
|                    | $\sigma$  | 0.11064         | 0.12291  | 0.26122   | 0.23064   |
| CIR                | $\kappa$  | 0.03683         | 0.044431 | 0.91274   | -0.13399  |
|                    | $\mu$     | 8.3589          | 8.4169   | 7.5827    | 6.6412    |
|                    | $\sigma$  | 0.03826         | 0.042108 | 0.094545  | 0.084197  |
| CKLS               | $\kappa$  | 0.033606        | 0.041253 | 0.97411   | -0.11478  |
|                    | $\mu$     | 8.3669          | 8.4177   | 7.5632    | 6.5059    |
|                    | $\sigma$  | 0.17815         | 0.43424  | 4.5291    | 1.7382    |
|                    | $\beta$   | -0.22647        | -0.59029 | -3.7814   | -1.007    |

Starting values:  $\kappa = 0$ ,  $\mu$  equals the mean of the sample,  $\sigma$  determines the starting value for  $\sigma$  by using Yoshida (1992) estimator for Vašíček and CIR model. As the starting value for  $\sigma$  in CKLS model we take earlier obtained estimation from CIR model. As the  $\beta$ , we take 0.5.

For examined time series we determine 100 ex-post forecasts, by using 10000 Monte-Carlo simulations, and to assess the quality of the forecasts we use common error measures. Small values of error measures are indicative of good quality of the forecasts, and consequently of good model fitting. The good quality also implies using diffusion models as the alternative for parametric models of time series.

Forecast errors were compared with errors obtained from popular time series models forecasts. The grade of time series models had been selected by using the Schwarz Information Criterion. The occurrence of unit root was verified by using Dickey-Fuller (Said, Dickey, 1984) and Phillips-Perron (1988) tests. In the case of failure to reject the  $H_0$  hypothesis we modeled the conditional mean by ARIMA( $p, 1, q$ ) model. Independently we used two innovation distributions – the Student and the Generalized Error Distribution developed by Nelson (1991). Moreover, we verified the existence of the ARCH effect by using Engle (1982) and McLeod-Li (1983) tests. The latter consists in applying Ljung-Box (1978) test to squared residuals of linear model.

As we can observe in Tables 3-6, the forecast errors for WIG20 index are bigger than for other indexes. Polish financial market is still a raising market, and WIG20 volatility is bigger than volatility of indexes traded in mature markets.

The type of applied model do not have big influence on the forecast quality. For indexes CAC40, DAX and WIG20 the forecast was a little bit better when we modeled the logarithmic prices by using diffusion models, but for NASDAQ index the forecast errors were smaller for parametric time series models. From among diffusion models, the best forecasts we obtained using CIR models. Moreover, we can notice that for all examined time series, Phillips and Yu method of parameters estimation leads to smaller forecast errors than GMM method.

Table 3. Values of the forecast errors obtained by using diffusion models. Logarithmic levels of CAC40 index

| Error                         | Phillips and Yu method    |            |                       | Generalized Method of Moments            |                                      |            |
|-------------------------------|---------------------------|------------|-----------------------|--|--------------------------------------|------------|
|                               | Vašíček                   | CIR        | CKLS                  | Vašíček                                  | CIR                                  | CKLS       |
| MSE                           | 6.9069e-5                 | 6.8895e-5  | 6.902e-5              | 6.8355e-5                                | 6.817e-5                             | 6.9836e-5  |
| MedE                          | 2.5575e-5                 | 2.6669e-5  | 2.6053e-5             | 2.6771e-5                                | 2.6696e-5                            | 2.7293e-5  |
| ME                            | 0.00077931                | 0.00077539 | 0.00078897            | 0.00032694                               | 0.00032105                           | 0.0012119  |
| MAE                           | 0.0062754                 | 0.0062588  | 0.0062567             | 0.0062394                                | 0.0062339                            | 0.0062989  |
| RMSE                          | 0.0083108                 | 0.0083003  | 0.0083078             | 0.0082677                                | 0.0082565                            | 0.0083568  |
| MAPE                          | 0.00072612                | 0.0007242  | 7.2396e-6             | 0.00072199                               | 0.00072135                           | 0.00072881 |
| AMAPE                         | 0.00036308                | 0.00036212 | 3.62e-6               | 0.000361                                 | 0.00036067                           | 0.00036444 |
| LL                            | 9.2627e-7                 | 9.2395e-7  | 9.2561e-7             | 9.1675e-7                                | 9.1422e-7                            | 9.3652e-7  |
| time series parametric models |                           |            |                       |  |                                      |            |
| Error                         | ARIMA(0,1,2)<br>(Student) |            | ARIMA(0,1,2)<br>(GED) | ARIMA(0,1,2)-<br>GARCH(1,1)<br>(Student) | ARIMA(1,1,1)-<br>GARCH(1,1)<br>(GED) |            |
|                               | 7.0223e-5                 | -          | -                     | 7.0214e-5                                | 7.0232e-5                            | 2.6895e-5  |
| MSE                           | 2.6884e-5                 | -          | -                     | 2.6873e-5                                | 0.00093054                           | 0.006338   |
| MedE                          | 0.00092873                | -          | -                     | 0.00092695                               | 0.0083805                            | 0.00073342 |
| ME                            | 0.0063371                 | -          | -                     | 0.0063367                                | 0.00036665                           | 0.00036673 |
| MAE                           | 0.0083799                 | -          | -                     | 0.0083794                                | -                                    | 9.42e-7    |
| RMSE                          | 0.00073332                | -          | -                     | 0.00073327                               | -                                    | -          |
| MAPE                          | 0.00036668                | -          | -                     | 0.00036665                               | -                                    | -          |
| AMAPE                         | 9.4187e-7                 | -          | -                     | 9.4175e-7                                | -                                    | -          |

Table 4. Values of the forecast errors obtained by using diffusion models. Logarithmic levels of DAX index

| Error                         | Phillips and Yu method    |            |                       | Generalized Method of Moments |  |                                      |
|-------------------------------|---------------------------|------------|-----------------------|-------------------------------|--|--------------------------------------|
|                               | Vašíček                   | CIR        | CKLS                  | Vašíček                       | CIR                                      | CKLS                                 |
| MSE                           | 7.7296e-5                 | 7.7358e-5  | 7.7691e-5             | 7.8574e-5                     | 7.8449e-5                                | 8.0539e-5                            |
| MedE                          | 2.6917e-5                 | 2.7475e-5  | 2.742e-5              | 2.706e-5                      | 2.9142e-5                                | 3.0036e-5                            |
| ME                            | 0.0014902                 | 0.0014927  | 0.0014958             | 0.0017948                     | 0.0017836                                | 0.0023101                            |
| MAE                           | 0.0066737                 | 0.0066797  | 0.0066842             | 0.0067379                     | 0.0067449                                | 0.0068758                            |
| RMSE                          | 0.0087918                 | 0.0087953  | 0.0088142             | 0.0088642                     | 0.0088571                                | 0.0089744                            |
| MAPE                          | 0.00075446                | 0.00075514 | 7.5564e-6             | 0.00076169                    | 0.00076248                               | 0.00077721                           |
| AMAPE                         | 0.00037726                | 0.0003776  | 3.7785e-6             | 0.00038089                    | 0.00038128                               | 0.00038867                           |
| LL                            | 9.8999e-7                 | 9.9078e-7  | 9.9505e-7             | 1.0063e-6                     | 1.0047e-6                                | 1.0312e-6                            |
| time series parametric models |                           |            |                       |                               |  |                                      |
| Error                         | ARIMA(0,1,2)<br>(Student) |            | ARIMA(0,1,2)<br>(GED) |                               | ARIMA(0,1,2)-<br>GARCH(1,1)<br>(Student) |                                      |
|                               |                           |            |                       |                               |  | ARIMA(1,1,1)-<br>GARCH(1,1)<br>(GED) |
| MSE                           | 7.8051e-5                 | 7.8615e-5  | 7.8058e-5             | 7.8058e-5                     | 7.8058e-5                                | 7.8058e-5                            |
| MedE                          | 2.7454e-5                 | 2.7438e-5  | 2.7546e-5             | 2.7546e-5                     | 2.7546e-5                                | 2.7546e-5                            |
| ME                            | 0.0015387                 | 0.0016234  | 0.0015409             | 0.0015409                     | 0.0015409                                | 0.0015409                            |
| MAE                           | 0.0067188                 | 0.0067622  | 0.0067197             | 0.0067197                     | 0.0067197                                | 0.0067197                            |
| RMSE                          | 0.0088346                 | 0.0088665  | 0.0088351             | 0.0088351                     | 0.0088351                                | 0.0088351                            |
| MAPE                          | 0.00075963                | 0.00076455 | 0.00075973            | 0.00075973                    | 0.00075973                               | 0.00075973                           |
| AMAPE                         | 0.00037985                | 0.0003823  | 0.0003799             | 0.0003799                     | 0.0003799                                | 0.0003799                            |
| LL                            | 9.9984e-7                 | 1.007e-6   | 9.9994e-7             | 9.9994e-7                     | 9.9994e-7                                | 9.9994e-7                            |

Table 5. Values of the forecast errors obtained by using diffusion models. Logarithmic levels of NASDAQ index

| Error                         | Phillips and Yu method    |            |                       | Generalized Method of Moments |  |                                      |
|-------------------------------|---------------------------|------------|-----------------------|-------------------------------|--|--------------------------------------|
|                               | Vašíček                   | CIR        | CKLS                  | Vašíček                       | CIR                                      | CKLS                                 |
| MSE                           | 7.3033e-5                 | 7.3078e-5  | 7.3089e-5             | 7.5079e-5                     | 7.4249e-5                                | 7.5569e-5                            |
| MedE                          | 1.8082e-5                 | 1.7335e-5  | 1.8443e-5             | 2.1237e-5                     | 1.9876e-5                                | 2.1606e-5                            |
| ME                            | 0.055699                  | 0.00064038 | 0.055435              | 0.0013926                     | 0.0013442                                | 0.0015984                            |
| MAE                           | 0.0062114                 | 0.0062146  | 0.0062015             | 0.0064573                     | 0.0064043                                | 0.0065172                            |
| RMSE                          | 0.0085459                 | 0.0085486  | 0.0085492             | 0.0086648                     | 0.0086168                                | 0.008693                             |
| MAPE                          | 0.00079535                | 0.00079575 | 7.9409e-6             | 0.00082676                    | 0.00081998                               | 0.00083443                           |
| AMAPE                         | 0.00039763                | 0.00039783 | 3.9699e-6             | 0.00041338                    | 0.00040998                               | 0.00041722                           |
| LL                            | 1.1983e-6                 | 1.199e-6   | 1.1992e-6             | 1.2319e-6                     | 1.2183e-6                                | 1.24e-6                              |
| time series parametric models |                           |            |                       |                               |  |                                      |
| Error                         | ARIMA(0,1,2)<br>(Student) |            | ARIMA(0,1,2)<br>(GED) |                               | ARIMA(0,1,2)-<br>GARCH(1,1)<br>(Student) |                                      |
|                               |                           |            |                       |                               |  | ARIMA(1,1,1)-<br>GARCH(1,1)<br>(GED) |
| MSE                           | 7.2567e-5                 | 7.2567e-5  | 7.2567e-5             | 7.282e-5                      | 7.282e-5                                 | 7.282e-5                             |
| MedE                          | 1.7764e-5                 | 1.7769e-5  | 1.7769e-5             | 1.8811e-5                     | 1.8819e-5                                | 1.8819e-5                            |
| ME                            | 0.056472                  | 0.056466   | 0.056466              | 0.055536                      | 0.055532                                 | 0.055532                             |
| MAE                           | 0.0062011                 | 0.006201   | 0.006201              | 0.0062328                     | 0.0062328                                | 0.0062328                            |
| RMSE                          | 0.0085187                 | 0.0085186  | 0.0085186             | 0.0085335                     | 0.0085334                                | 0.0085334                            |
| MAPE                          | 0.00079404                | 0.00079404 | 0.00079404            | 0.00079811                    | 0.00079811                               | 0.00079811                           |
| AMAPE                         | 0.00039697                | 0.00039697 | 0.00039697            | 0.00039901                    | 0.00039901                               | 0.00039901                           |
| LL                            | 1.1907e-6                 | 1.1907e-6  | 1.1907e-6             | 1.1949e-6                     | 1.1948e-6                                | 1.1948e-6                            |

Table 6. Values of the forecast errors obtained by using diffusion models. Logarithmic levels of WIG20 index

| Error                         | Phillips and Yu method |            |            | Generalized Method of Moments |            |            |
|-------------------------------|------------------------|------------|------------|-------------------------------|------------|------------|
|                               | Vašíček                | CIR        | CKLS       | Vašíček                       | CIR        | CKLS       |
| MSE                           | 0.00018948             | 0.0001892  | 0.00018936 | 0.00018897                    | 0.00018938 | 0.00018973 |
| MedE                          | 6.1233e-5              | 6.3602e-5  | 6.5722e-5  | 6.4217e-5                     | 6.4506e-5  | 6.7774e-5  |
| ME                            | 0.00063017             | 0.00063866 | 0.058766   | 0.0001788                     | 0.00032633 | -0.0001169 |
| MAE                           | 0.010608               | 0.010619   | 0.010609   | 0.010582                      | 0.010595   | 0.010586   |
| RMSE                          | 0.013765               | 0.013755   | 0.013761   | 0.013747                      | 0.013762   | 0.013774   |
| MAPE                          | 0.0013032              | 0.0013046  | 1.3034e-5  | 0.0013001                     | 0.0013016  | 0.0013007  |
| AMAPE                         | 0.00065164             | 0.00065232 | 6.5172e-6  | 0.00065007                    | 0.00065083 | 0.00065032 |
| LL                            | 2.8646e-6              | 2.8603e-6  | 2.8627e-6  | 2.8568e-6                     | 2.8629e-6  | 2.8682e-6  |
| time series parametric models |                        |            |            |                               |            |            |
| Error                         | ARMA(0,2)<br>(GED)     |            |            | ARMA(1,1)<br>(Student)        |            |            |
|                               | MSE                    | 0.01904131 |            | 0.019061629                   |            |            |
| MedE                          | 6.6946E-05             |            |            | 6.87936e-5                    |            |            |
| ME                            | 0.00035112             |            |            | 0.000398688                   |            |            |
| MAE                           | 1.06889157             |            |            | 1.071935667                   |            |            |
| RMSE                          | 0.13799026             |            |            | 0.138063858                   |            |            |
| MAPE                          | 0.131329               |            |            | 0.1317012                     |            |            |
| AMAPE                         | 0.00065666             |            |            | 0.000658529                   |            |            |
| LL                            | 2.8789e-06             |            |            | 2.88198e-6                    |            |            |

Note: MSE – the mean squared error, MedE – the mean median error, ME – the mean error, MAE – the mean average error, RMSE – the root of the mean squared error, MAPE – the mean average percentage error, AMAPE – corrected average percentage error, and LL – logarithmic loss function (cf. Welfe 1998; Doman, Doman, 2004).

## 6. Conclusions

The high quality of the forecast obtained from the diffusion models is indicative of good fitting of the diffusion models to the studied time series. The values of the forecast errors are often smaller when diffusion models were used. It is notable that in diffusion models the volatility depends only on white noise and optionally on current value of the process. In ARIMA-GARCH models the volatility is described by the second parametric equation.

The most surprising fact is that the CKLS model leads to worse quality of the forecast than the Vašíček and CIR model. After all, both models are special cases of the CKLS model. The reason for that situation lies in very bad fitting of the CKLS model to the examined time series. The estimates of  $\beta$  are negative, while the model assumed that  $\beta \in [0,1]$ .

It is notable that determining the forecasts by using diffusion models is not laborious. The MATLAB procedures used in the conducted research need only a fraction of a second to estimate parameters by using two-stage Phillips-Yu method and a few seconds if we decide to use the GMM method. The most laborious part of the calculation is determining 10000 sample paths, but it takes up to two minutes to do this operation.

## References

Black, F., Scholes, M. (1973), The Pricing of Option and Corporate Liabilities, *Journal of Political Economy*, 81, 637–659.

Broadie, M., Jain, A. (2008), Pricing and Hedging Volatility Derivatives, *Journal of Derivatives*, 15, Issue 3, 7–24.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., Sanders, A. B. (1992), An Empirical Comparison of Alternative Models of Short Term Interest Rates, *Journal of Finance*, 47, 1209–1227.

Cliff, M. T. (2003) *GMM and MINZ Program Libraries for Matlab*, Krannert Graduate School of Management Purdue University.

Cox, J. C., Ingersoll, J., Ross, S. (1985), A Theory of the Term Structure of Interest Rates, *Econometrica*, 53, 385–407.

Cziraky, D., Kucherenko, S. (2008), Monte Carlo Forecasting from CIR Square Root Diffusion Models, *BRODA Ltd.* <http://www.broda.co.uk> (8.01.2009).

Detemple, J., Osakwe, S. (1999), The Valuation of Volatility Options, *CIRANO Paper, Scientific Series*.

Doman, M., Doman, R. (2004), *Econometric Modeling of Polish Financial Market Dynamic* (In Polish), Poznań University of Economics, Poznań.

Engle, R. F. (1982), Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 50, 987–1007.

Hansen, L. P. (1982), Large Sample Properties of Generalized Method Of Moments Estimators, *Econometrica*, 50, 1029–1054.

Heston, S. L. (1993), A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options, *The Review of Financial Studies*, 6, Issue 2, 327–343.

Howison, S., Rafailidis, A., Rasmussen, H. (2004), On the Pricing and Hedging of Volatility Derivatives, *Applied Mathematical Finance*, 11, Issue 2, 317–346.

Javaheri, A., Wilmott, P., Haug, E. (2002), GARCH and Volatility Swaps, *Wilmott Magazine*, January, 1–17.

Jagannathan R., Kaplin A., Sun S. G. (2004), An Evaluation of Multi-Factor CIR Models Using Libor, Swap Rates, and Cap and Swaption Prices, *Journal of Econometrics*, 116, 113–146.

Ljung, G. M., Box, G. E. P. (1978), On a Measure of Lack of fit in Time Series Models, *Biometrika*, 65, 297–303.

Mannolini, A., Mari, C., Renò, R. (2008), Pricing Caps and Floors with the Extended CIR Model, *International Journal of Finance & Economics*, 13 (4) 386–400.

McLeod, A. I., Li, W. K. (1983), Diagnostic Checking ARMA Time Series Models Using Squared Residual Autocorrelations, *Journal of Time Series Analysis*, 4, 269–273.

Merton, R. C. (1973), Theory of Rational Option Pricing, *Bell Journal of Economics and Management Science*, 4 (1) 141–183.

Merton, R. C. (1974), On The Pricing of Corporate Debt: The Risk Structure of Interest Rates, *Journal of Finance*, 29, 449–470.

Nelson, D. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, *Econometrica*, 59, 347–370.

Newey, W. K., West, K. D. (1987), A Simple, Positive Semidefinite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica*, 59, 347–370.

Phillips, P. C. B., Perron, P. (1988), Testing for a Unit Root in Time Series Regressions, *Biometrika*, 75, 335–346.

Phillips, P. C. B., Yu, J. (2009), A Two-Stage Realized Volatility Approach to Estimation of Diffusion Processes with Discrete Data, *Journal of Econometrics*, 150, Issue 2, 139–150.

Psychoyios, D., Skiadopoulos, G. (2006), Volatility Options: Hedging Effectiveness, Pricing, and Model Error, *Journal of Futures Markets*, 26, 1–31.

Said, E., Dickey, D.A. (1984), Testing for Unit Roots in Autoregressive Moving Average Models of Unknown Order, *Biometrika*, 71, 599–607.

Sepp, A. (2008), VIX Option Pricing in a Jump-Diffusion Model, *Risk*, April, 84–89.

Vašiček, O. (1977), An Equilibrium Characterization of the Term Structure, *Journal of Financial Economics*, 5, 177–188.

Tamba, Y. (2006), Pricing the Bermudan Swaption with the Efficient Calibration, *NUCB Journal of Economics and Information Science*, 51, Issue 1, 17–31.

Welfe, A. (1998), *Econometrics* (In Polish), Polskie Wydawnictwo Ekonomiczne, Warszawa.

Yoshida, N. (1992), Estimation for Diffusion Processes From Discrete Observation, *Journal of Multivariate Analysis*, 41, 220–242.

### Prognozowanie procesów finansowych za pomocą modeli dyfuzji

Z a r y s t r e ś c i. Prognozowanie szeregów czasowych jest jednym z najważniejszych zagadnień współczesnej ekonometrii finansowej. W obliczu rosnącego zainteresowania modelami z czasem ciągłym i szybkiego rozwoju metod ich estymacji, podejmujemy w pracy próbę modelowania i prognozowania szeregów czasowych z różnych rynków finansowych za pomocą modeli dyfuzji. Stosujemy w tym celu bazującą na symulacjach Monte-Carlo metodę wprowadzoną przez Cziraky i Kucherenko (2008). Jakość otrzymanych prognoz zostaje skonfrontowana z jakością prognoz otrzymanych za pomocą powszechnie stosowanych parametrycznych modeli szeregów czasowych.

S ł o w a k l u c z o w e: model ARIMA, model GARCH, modele dyfuzji, pierwiastek jednostkowy, prognozy ex-post, symulacja Monte-Carlo.