Journal of CENTRUM Cathedra • Volume 5, Issue 1, 2012 • 70-90

JCC
Journal of
CENTRUM
Cathedra

Option Pricing Models with HF Data: An Application of the Black Model to the WIG20 Index

Ryszard Kokoszczyński
University of Warsaw, Warsaw, Poland

Paweł Sakowski
University of Warsaw, Warsaw, Poland

Robert Ślepaczuk *
University of Warsaw, Warsaw, Poland

Abstract

In this paper, we compared several Black option pricing models by applying different measures of volatility and examined the Black model with historical (BHV), implied (BIV), and several different types of realized (BRV) volatility. The main objective of the study was to find the best model; that is, the model that predicts the actual market price with the minimum error. The high frequency (HF) data and bid-ask quotes (instead of transactional data) for the Warsaw Stock Exchange (WSE) were used to omit the problem of non-synchronous trading and to increase the number of observations. Several error statistics and the percentage of price overpredictions (OP) showed the results that confirmed the initial intuition that the BIV model is the best model, the BHV model is the second best, and the BRV is the least efficient among the models studied.

Keywords: Option pricing models, financial market volatility, high frequency financial data, realized volatility, implied volatility, microstructure bias, emerging markets, Warsaw Stock Exchange

JEL Classification codes: G14, G15, C61, C22

Option trading dates back to the 17th century, when options were part of, and one of the reasons for, the South Sea bubble and the Amsterdam tulip mania. However, rapid growth in the options market came in the 1970s only. First, two seminal papers by Black and Scholes (1973) and Merton (1973) introduced the Black-Scholes-Merton (BSM) model, a formula for valuing European options. In 1973, the Chicago Board of Options Exchange (CBOE) was founded, which heralded the beginning of trading on standardized listed options; the CBOE adopted the BSM model for option pricing in 1975.

The rapid growth of option markets, due to the combination of a seemingly reliable pricing formula and a good exchange mechanism, brought a considerable amount of data and stimulated the intensive development of the option pricing research. Soon, empirical studies showed clearly that some theoretical assumptions of the BSM model are not fully supported by these data (Bates, 2003) and that the BS formula exhibits substantial pricing biases across both moneyness and maturity (Bakshi, Cao, & Chen, 1997). A number of new models were then proposed, each of them relaxing some of the restrictive assumptions of the BSM model (Broadie & Detemple, 2004; Garcia, Ghysels, & Renault, 2010; Han, 2008; Mitra, 2009). There is also a growing literature devoted to comparisons of their various features although even the best metric for model comparison is a controversial issue (Bams, Lehnert, & Wolff, 2009).

Notwithstanding the numerous criticisms, the BSM model is still widely used, both as some kind of benchmark in comparative studies mentioned earlier and among financial practitioners. A detailed analysis of the literature shows that the BIV model calculated on the basis of the last observation performs quite well even when compared with many different pricing models such as standard BSM model, BRV model, Generalized autoregressive conditional heteroskedasticity (GARCH) option pricing models, or various stochastic volatility (SV) models (An & Suo, 2009; Andersen, Frederiksen, & Staal, 2007; Bates, 2003; Brandt & Wu, 2002; Ferreira, Gago, Leon, & Rubio, 2005; Mixon, 2009; Raj & Thurston, 1998). However, practically all these studies used data from the mature capital markets of the United States of America, the United Kingdom, and Japan. Thus, the purpose of this study was to find the best option pricing model, but for a different market such as an emerging market. We used the WSE data for our study not only because the Polish market is best known to us but also because this market is the largest in the region[1] and can therefore lead to general conclusions.

The choice of the market for the present study also derived from the very limited knowledge available regarding option pricing in the Polish capital market. Most studies exist only in Polish or in the form of unpublished papers, which makes them practically inaccessible to a wider audience. Moreover, these studies are usually limited to GARCH option pricing models (Osiewalski & Pipień, 2003). The purpose of this paper was thus to close this gap, at least partially, as the only study covering similar issues is Fiszeder's (2010) although it is limited to daily data.

There are several ways of measuring and estimating volatility. A number of studies have indicated plausibly that it makes a difference what kind of volatility –historical, implied, or realized– has been applied (Ammann, Skovmand, & Verhofen, 2009; Berkowitz, 2010; Martens & Zein, 2004). That observation has been one of the major factors defining the scope of this study. Thus, in order to verify the initial hypothesis for this study, a few additional questions were as follows:

- What kind of volatility process should be used in the black model?
- What length of time period (parameter n – responsible for the memory of the process) should be used for averaging volatility in the estimation?
- What is the optimal interval (delta) for estimating volatility?
- Do errors depend on the option's time to maturity (TTM) and moneyness ratio (MR)?

These questions were the reason HF data (10s data interval, based on tick data) for WIG20 index[2] option quotes (bid and ask) were used in order to increase the observed liquidity of the market and to remove nonsynchronous bias.

The remaining sections of this paper are organized as follows. The next section indicates the methodology for option pricing with a special focus on various volatility measures and their estimators. Afterwards, there is a detailed section with the description of the data used and of the volatility processes that were studied. HF data brought a significant number of specific technical issues that constrained to some extent the whole research endeavor. The results section indicates the results of the study in detail. The last section shows the implications these results have for other financial models and for further research on option pricing.

Option Pricing Methodology

The main assumption was that one can price a European style option on the WIG20 index applying the Black model for futures option contract, where the WIG20 index futures contract is the basis instrument. This is possible for two reasons[3]:

- WIG20 index futures mature exactly the same day as WIG20 index options, and the expiration prices are set exactly in the same way; and,
- WIG20 index options are European-style options, so there is no need to worry about early expiration as in the case of American options.

The use of the Black model instead of the BSM model was motivated by the following facts:

- It was not necessary to calculate the dividend yield for the index when using the Black model, so we omited the problem of the BSM assumption that dividends are paid continuously; thus, we did not have to calculate the dividend yield nor set the exact time of dividend payment;
- It was possible to use the data from the period between 9:00 a.m. and 9:30 a.m. each day though index quotation starts only at 9:30 a.m., which gives a longer trading day.

The formulas for the Black model (Black, 1976) are presented below:

$$c = e^{-rT}[FN(d_1) - KN(d_2)] \quad (1)$$

$$p = e^{-rT}[KN(-d_2) - FN(-d_1)] \quad (2)$$

where:

$$d_1 = \frac{\ln(F/K) + \sigma^2 T/2}{\sigma\sqrt{T}} \quad (3)$$

$$d_2 = \frac{\ln(F/K) - \sigma^2 T/2}{\sigma\sqrt{T}} = d_1 - \sigma\sqrt{T} \quad (4)$$

where c and p are respectively valuation of a call and a put futures option, T is the expiration date, r is the risk-free rate, F – the futures price, K – the underlying strike, and *N(.)* is the cumulative standard normal distribution.

Some properties of the Black model were examined with three different types of volatility estimators: historical volatility, realized volatility, and implied volatility. We used high frequency data (10s data interval, based on the tick data) for WIG20 index option quotes (bid and ask) to calculate them. Formulas for all three estimators are presented below.

1. The historical volatility (HV) estimator (standard deviation for log returns based on the daily interval) was directly derived from

$$VAR_{\Delta}^{n} = \frac{1}{(N_{\Delta} n) - 1} \sum_{t=1}^{n} \sum_{i=1}^{N_{\Delta}} (r_{i,t} - \bar{r})^2 \quad (5)$$

where:

VAR_{Δ}^{n} – variance of log returns calculated with high frequency data on the basis of last n days;

$r_{i,t}$ – log return for i^{th} interval on day t with sampling frequency equal to Δ, which is calculated in the following way:

$$r_{i,t} = logC_{i,t} - logC_{i-1,t} \quad (6)$$

$C_{i,t}$ – close price for the i^{th} interval on the day t with the sampling frequency equal Δ;

N_{Δ} – number of Δ intervals during the stock market session;

n – memory of the process measured in days, used in the calculation of respective estimators and average measures;

$\bar{r}$ – average log return for the i^{th} interval on the basis of last n days with the sampling frequency Δ, which is calculated in the following way:

$$\bar{r} = \frac{1}{N_{\Delta}\, n} \sum_{t=1}^{n} \sum_{i=1}^{N_{\Delta}} r_{i,t} \quad (7)$$

In the case of the historical volatility estimator $i = 1$ and $N_{\Delta} = 1$ for every $r_{i,t}$ (daily log returns) and $C_{i,t}$ in formulas (5), (6), and (7). Moreover, we used the constant value of parameter $n = 21$, because we wanted to reflect the historical volatility from the last trading month.

2. The realized volatility (RV) estimator was based on the following formula:

$$RV_{\Delta,t} = \sum_{i=1}^{N_{\Delta}} {r_{i,t}}^2 \quad (8)$$

3. The implied volatility (IV) estimator was based on the most recent observation, therefore σ was derived from the Black formula with the assumption that other parameters and the valuation results were given. The IV for the previous observation was calculated separately for each class of the TTM and of the MR, that is, for 50 different classes; the details of this option classification are presented in the results section. This estimator was then treated as an input variable for the volatility parameter to calculate the theoretical value for the BIV model for the next observation.

In the next step, the HV was annualized and transformed into standard deviation because this is the parameter used in the Black model[4]:

$$HV={}^{annual\_std}SD_{\Delta}^{n}=\sqrt{252\,N_{\Delta}VAR_{\Delta}^{n}} \tag{9}$$

Contrary to HV which is based on information from many periods ($n > 1$), a RV estimator requires information only from one single period (interval Δ). Therefore, the procedure of averaging and annualizing a RV estimator is slightly different from that presented in formula (9)[5]:

$${}^{annual\_std}[RV]_{\Delta}^{n}=\sqrt{252}\sqrt{\frac{1}{n}\sum_{t=1}^{n}[RV]_{\Delta,t}} \tag{10}$$

With these volatility estimators, several types of option pricing models were examined:

- The Black model with historical volatility (sigma as standard deviation, $n = 21$) – BHV;
- The Black model with realized volatility (RV as an estimate of sigma; RV was calculated on the basis of observations with a different interval Δ and a different parameter n in the process of averaging) – BRV[6];
- The Black model with implied volatility (IV as an estimate of sigma; IV was calculated for the previous observation, separately for each TTM and MR - 50 different groups) – BIV.

Finally, the following error statistics were calculated for all these models in order to verify the research hypothesis:

- Root mean squared error (RMSE):

$$RMSE=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(Black_i-MID_i)^2} \tag{11}$$

where:

MID_i – means the market price (midquote in this research study);

$Black_i$ – means the Black model price (BHV, BRV, or BIV).

- Mean absolute percentage error (MAPE):

$$MAPE=\frac{1}{n}\sum_{i=1}^{n}\left|\frac{Black_i-MID_i}{MID_i}\right| \tag{12}$$

- Percentage of price overpredictions (OP):

$$OP=\frac{1}{n}\sum_{i=1}^{n}OP_i \tag{13}$$

where:

$$OP_i = \begin{cases} 1, & if \quad Black_i > MID_i \\ 0, & if \quad Black_i < MID_i \end{cases}$$

The Data and the Description of Volatility Processes

Data

The empirical analysis was based on HF financial data for WIG20 index options and WIG20 futures[7], supplied by the Information Products Section of the WSE. These data covered the period from January 2, 2008, to June 20, 2008. Because of well-known statistical problems, tick data were aggregated to 10s quotes.

The number of 10s bid-ask quotes for a trading day depends on the trading hours for option and futures contracts. The trading takes place from 9:00 a.m. to 4:30 p.m. for the time period under consideration[8]. Only those quotes for which both bid and ask quotes were available simultaneously were taken into account, so that it was possible to calculate the mid quotes[9]. These quotes were later treated as the market consensus of option investors and were used for comparison with theoretical prices obtained from the option pricing models. Although these quotes do not represent actual prices at which transactions take place, most researchers who test alternative option pricing models and include the Black-Scholes model among models tested use bid-ask quotes (mid-quotes) as these enable them to avoid microstructural noise effects (Dennis & Mayhew, 2009). In addition, Ait-Sahalia and Mykland (2009) stated explicitly that quotes "contain substantially more information regarding the strategic behaviour of market makers" and that they "should be probably used at least for comparison purposes whenever possible" (p. 592).

There were no corrections for outliers because of the need to show fully the properties of models tested, even for options with low prices and short TTM, which are usually excluded from similar studies. Additionally, the Warsaw Interbank Offered Rate (WIBOR) interest rate (converted into 10s intervals) was used as the interest free rate in option pricing models, and TTM was calculated in seconds.

These operations led to complete data for 128 index options (65 call and 63 put options expiring in March, June, and September). Thus, the sample period (118 trading days with 2701 observations for each day) included 31 871 810s observations (mid quotes, WIBOR rates, TTM, and strike prices for each option). These data were then used in the process of calculation of volatility parameters (HV, RV, and IV) and later on for theoretical option valuation (BHV, BRV, and BIV models).

Descriptive statistics for WIG20 futures time series

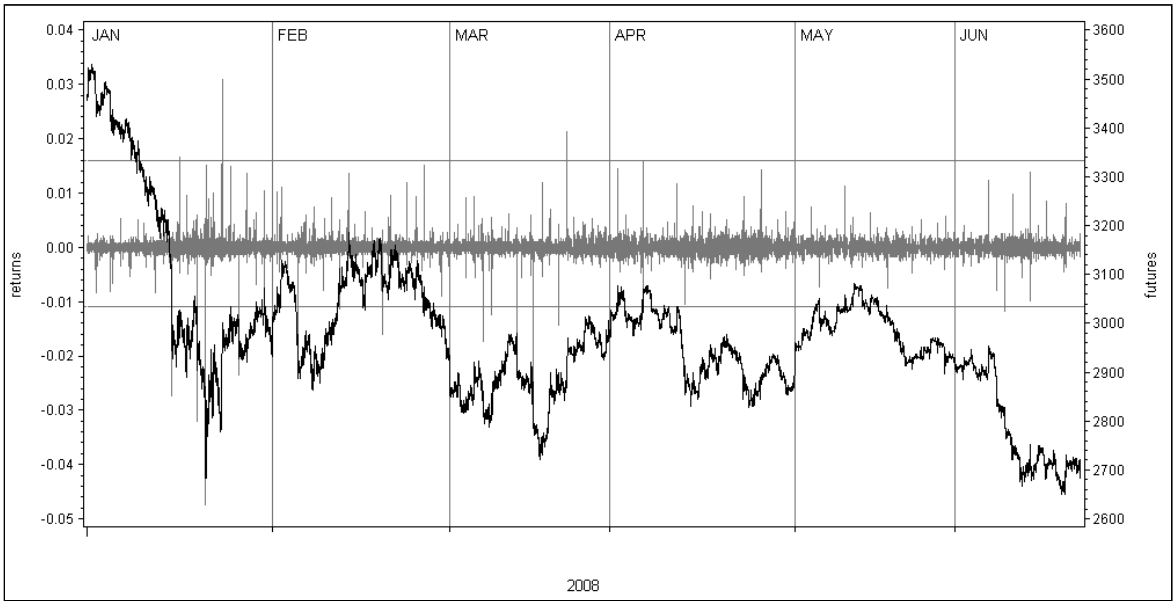
Table 1 summarizes the descriptive statistics for 10s interval data for continuous futures contract (with and without the opening jump effect representing the rate of return for the period between the closing of the market in the evening of the previous day and its opening in the morning of the current day – described respectively as R_f and $R_{f'}$) in order to show the distribution for the basis instrument. This distribution is shown in order to check the crucial assumption of option pricing models tested in this study; that is, the normality of returns. [10] The statistics presented below seem to confirm the belief that the distribution of HF data is not exactly normal.[11]

Table 1
Descriptive Statistics for Index Futures Returns (with and without Opening Jump Effect)

		R_f [a]	$R_{f'}$ [b]
	N	318717	318482
	Mean	-0.000000787	-0.000000624
	Median	0	0
	Std Deviation	0.0003985	0.0003358
	Range	0.07847	0.02551
	Minimum	-0.047473855	-0.010453057
	Maximum	0.030991753	0.015059446
	Kurtosis	1369.388606	79.746338
	Skewness	-7.31722	1.781351
	Test for Normality		
Kolmogorov-Smirnov	Statistic	0.3683	0.3684
	p-value	<0.01	<0.01
Jarque-Berra	Statistic	24 904 800 000	84 556 553.7
	p-value	<0.00001	<0.00001

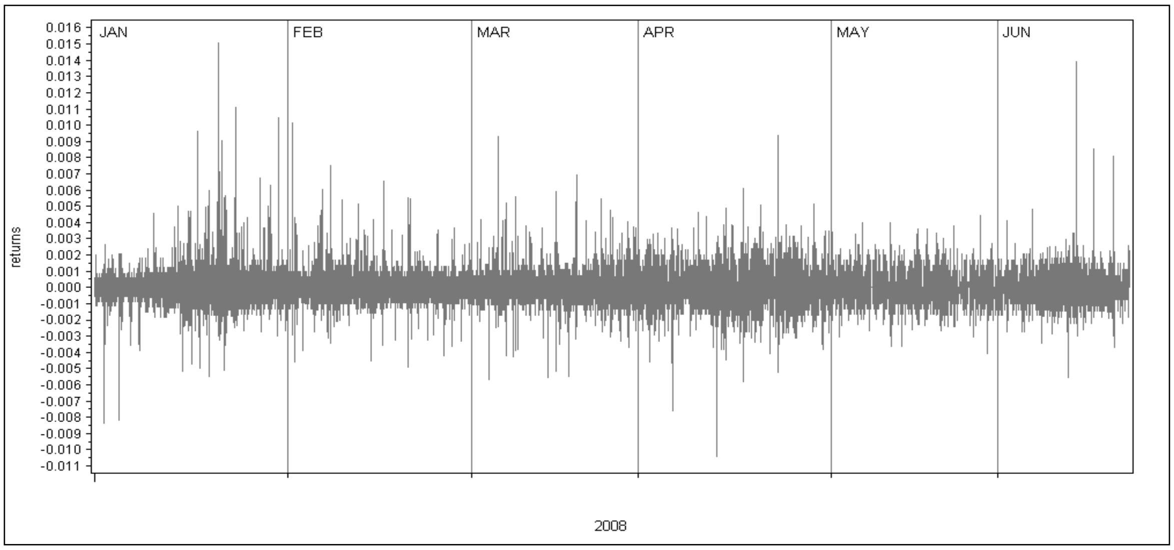
Note. [a] The original data. [b] The modified data, without opening jump effect.

Figures 1 and 2 show that the opening jump effect is responsible for the large fraction of the departure from normality.



Note. The returns cover the data span between January 2, 2008 and June 19, 2008.

Figure 1. Index futures returns with the opening jump effect.



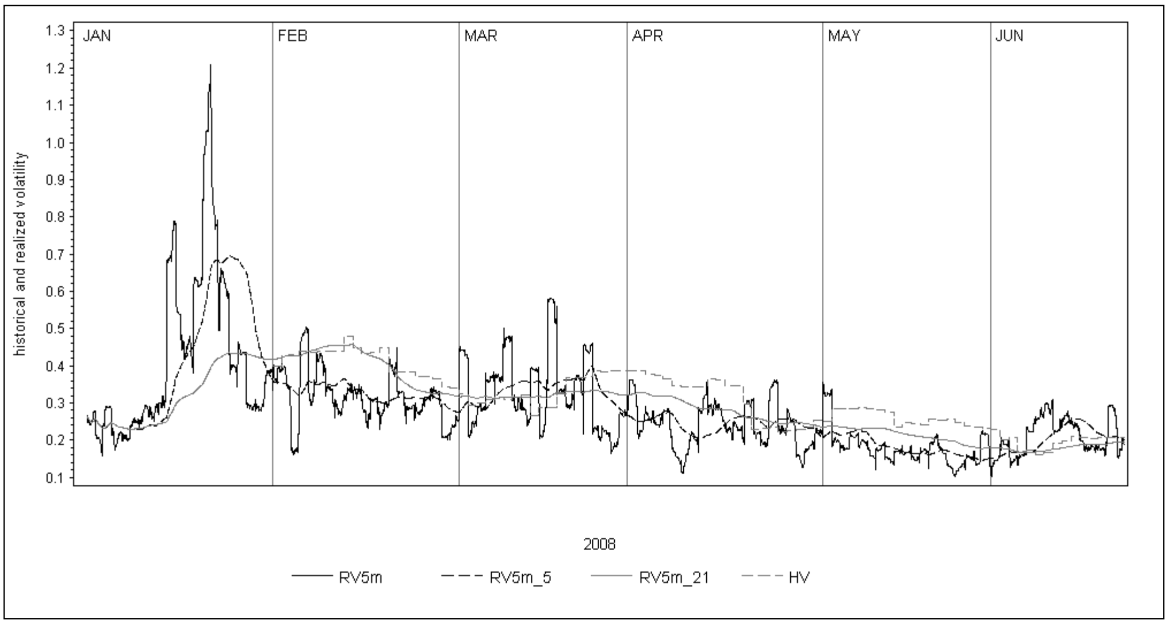
Note. The 10s returns between the closing price from each day and the opening price from the next day have been excluded. The returns cover the data span from January 2, 2008 and June 19, 2008.

Figure 2. Index futures returns without the opening jump effect.

Formally, this non-normality means that the standard BSM model should not be used to price an option on such a basis instrument. Therefore, we then decided to use the Black model with many different volatility measures.

Description of volatility processes

This section shows some properties of volatility parameters distributions, presented in Figures 3 and 4. We believe that they are the main reason for the differences between the option pricing models compared.



Note. The volatility time series cover the data period between January 2, 2008 and June 19, 2008.

Figure 3. HV and RV (5m, 5m_5, 5m_21).

First of all, RV time series that is not averaged (RV_f1_5m) exhibits substantial volatility of volatility (parameter kappa in stochastic volatility models), especially in comparison with averaged RV, which then strongly influences the results for BRV model. This feature of RV is responsible for the high errors of these models, especially in high volatility environment.

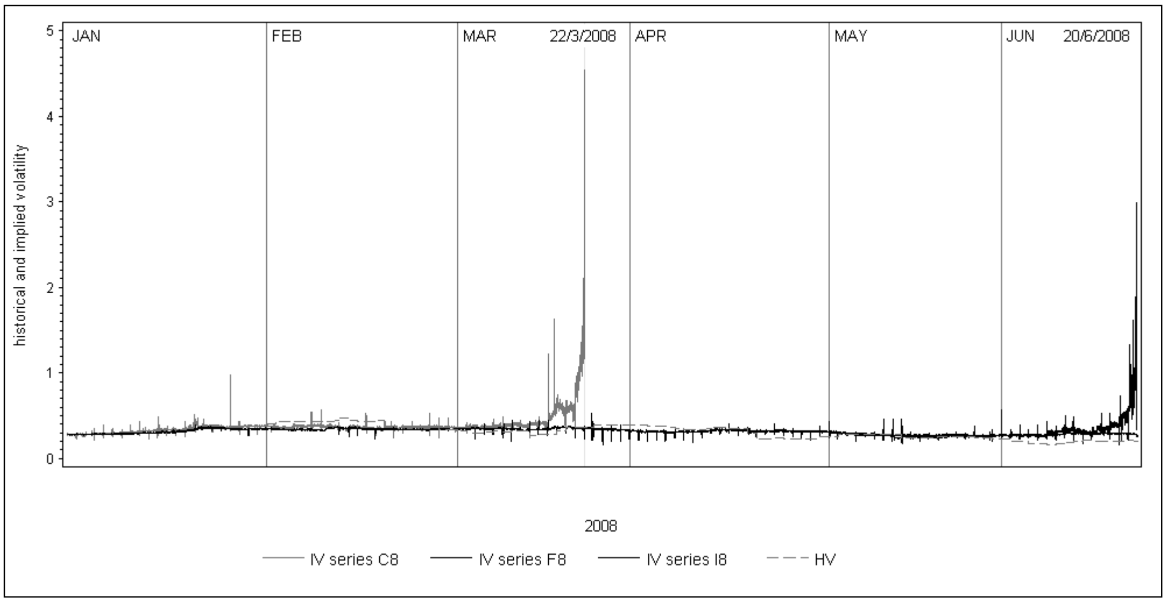
Table 2 also confirms the observation based on Figure 3 concerning the effect of averaging the RV estimator based on its efficiency; that is, a decreasing standard deviation and a narrower range of fluctuations of the RV while parameter *n* increases. Nevertheless, the mean value of RV is robust to the process of averaging, which implies that higher volatility for estimator with lower *n,* is responsible for the symmetrical departure from the mean value.

Table 2
Descriptive Statistics for Realized Volatility Estimators [a]

	RV5m	RV5m_2	RV5m_5	RV5m_10	RV5m_21
N	316 017	310 615	302 512	289 007	259 296
Mean	0.286	0.287	0.289	0.291	0.286
Std Deviation	0.137	0.123	0.112	0.102	0.082
Range	1.107	0.778	0.550	0.428	0.290
Minimum	0.101	0.124	0.146	0.156	0.166
Maximum	1.208	0.902	0.695	0.584	0.457

Note. [a] The different sample size is the result of different number of intervals which are necessary to compute the first value of averaged RV. The latter depends on parameter *n*.

Secondly, the process of averaging the RV estimator drives it closely to the classical volatility estimator, especially in cases with the same value of parameter *n* (responsible for the memory of the process).



Note. The volatility time series cover the data period between January 2, 2008 and June 19, 2008.

Figure 4. IV for at the money - ATM call option.

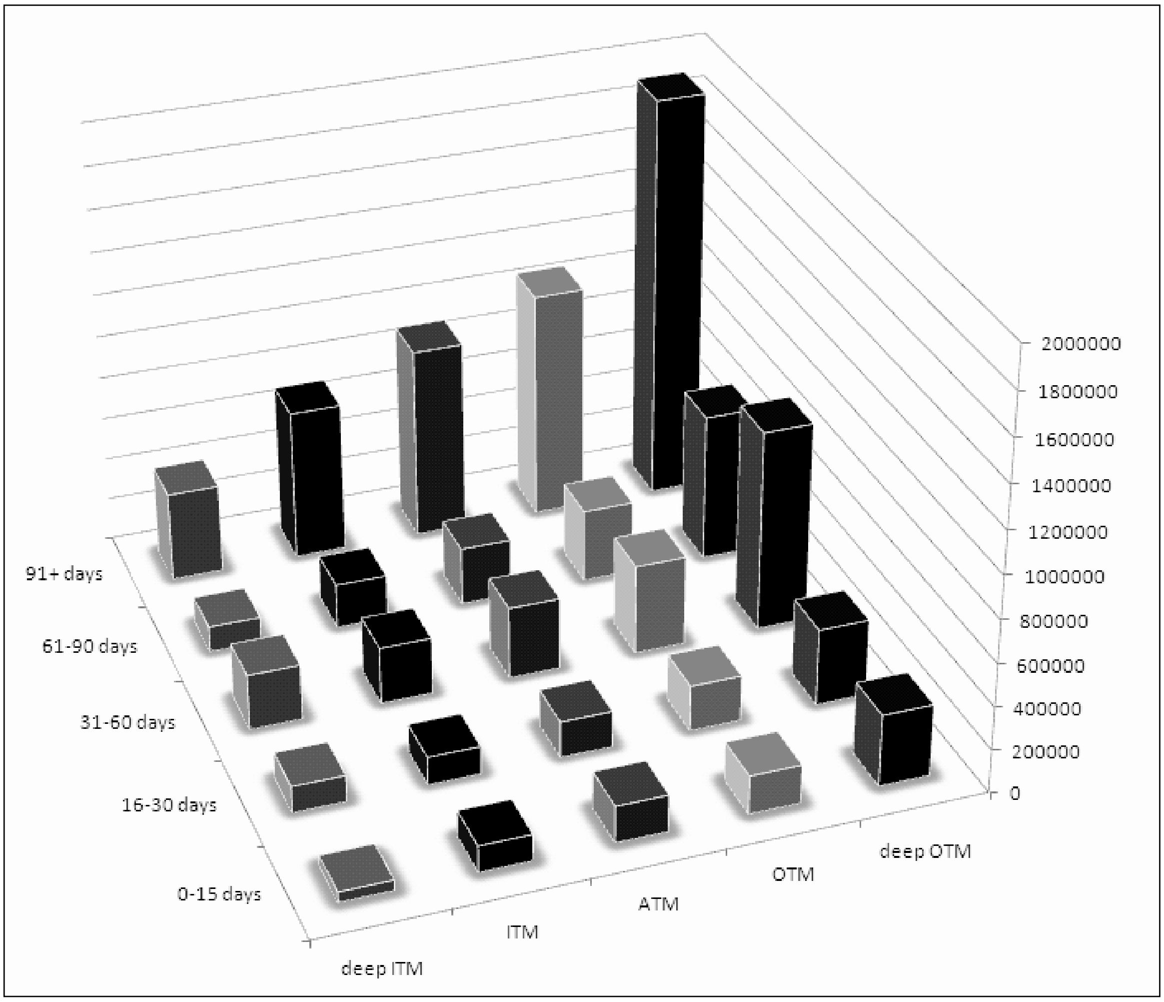
Thirdly, Figure 4 shows that IV for call option[12] explodes for the very short TTM (less than 5 days) and very low-priced options (i.e., deep out of the money - OTM). This result was probably the reason for excluding options with short TTM and market premium lower than 5 or 10 in most research studies comparing different volatility models. However, this exclusion was not done as the purpose of this study was to investigate comprehensively the volatility estimators.

Option Classification

After the calculation of theoretical prices for each model, there were more than 21 million theoretical premiums. Error statistics were then calculated for six pricing models (BHV, BRV10s, BRV5m, BRV5m_5, BRV5m_21, and BIV)[13] and ordered according to the following:

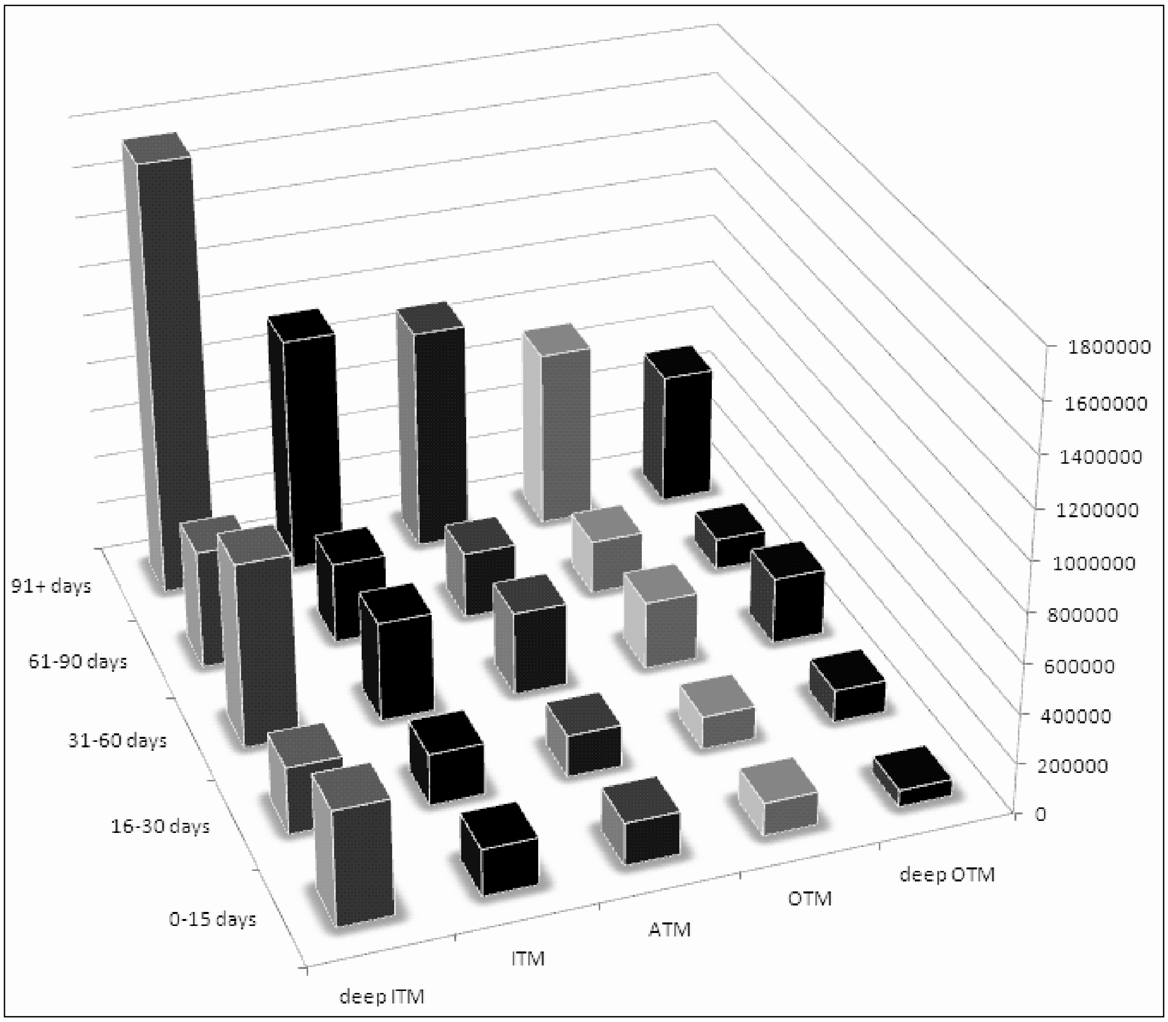
- two types of options (call and put);
- five classes of the MR [14]*:* deep OTM (0, 0.85), OTM (0.85, 0.95), ATM (0.95, 1.05), in the money - ITM (1.05, 1.15), and deep ITM (> 1.15) for call options and in the opposite order for put options;
- five classes for the TTM: (0, 15 days], [16, 30 days], [31, 60 days], [61, 90 days], and [91+ days).

This operation allowed a multidimensional comparison of the pricing models used in this study. Figures 5 and 6 show the sample size of each class for BRV models for call and put options separately. Based on mid-quotes data and 10s interval it was possible to calculate millions of theoretical prices for each of the models: 17 million observations for the BHV model because the first value of HV is for February 1, 2008, or 21 million observations for the BIV model.



Note. Active mid-quotes mean options that were quoted in the sample period.

Figure 5. Number of call options premiums to TTM and MR for active mid-quotes.



Note. Active mid-quotes mean options that were quoted in the sample period.

Figure 6. Number of put options premiums to TTM and MR for active mid-quotes.

The numbers in these two figures show that the activity of market participants, within the research period, was focused on deep ITM and ITM put options and deep OTM and OTM call options. However, this is not a general feature of the Warsaw market, but only the result of a sharp downward movement of WIG20 prices in the period under consideration and of the procedure of introducing new strike prices by the WSE.

Moreover, the analysis of information about available strike prices and active mid quotes[15] was performed separately for call and put options [16]. The histograms describing this issue indicate confirmation of the observation revealed by previous figures that there were a great number of ITM and deep ITM put options and OTM and deep OTM call options available on the WSE in the sample period. However, only part of the available strike prices were quoted by the market participants.

Technical and Statistical Issues when Dealing with HF Data

Before presenting a discussion of the results, it is necessary to outline briefly the main problems and obstacles encountered in the course of this research study. The first problem was the question of how to present the study results. Taking into account that the objective was to show detailed results (4-error statistics) for six models (call and put options presented separately) divided into five classes of TTM and five classes of MR, there were 1200 values of error statistics. Presenting these figures in a table or a number of tables did not appear practical. Therefore, we decided to use 3-D figures with boxes scaled with global and local minima and maxima in order to show in a transparent manner differences among our models along various dimensions. The results section includes a detailed description of the presentation method.

Second, the comparison of option pricing models is based on three types of error statistics: relative (MAPE), absolute (RMSE), and additionally OP. However, we believe that RMSE, the type of absolute statistics that is most often used in this kind of research, is not appropriate in some situations. It can lead to wrong conclusions (see Figures 8 or 11), especially when the aim is to find some patterns comparing models in the same TTM or MR class. In this situation, relative statistics are much better suited for evaluation. The next subsection includes a detailed discussion of this phenomenon.

Finally, the results analysis led to several untypical observations. After investigating their cause, we tried to answer the question whether they should be treated as outliers. These observations occured for the following models and error statistics:

- MAPE, Call, TTM = 3, and MR = 1 for BRV10s, BRV5m, BRV5m_5;
- MAPE, Call, TTM = 1, and MR = 1 for BIV;
- RMSE, Put, TTM = 1, 2, 3, and MR = 5 for all models;
- MAPE, Put, TTM = 1, and MR = 2, 3 for BIV models.

Results

It is first necessary to comment on the BRV model. In effect, several different BRV models were tested with different values of the Δ parameter: 10 seconds, 1 minute, 5 minutes, and 15 minutes. A number of averaging parameters were also considered: 1 day, 2 days, 3 days, 5 days, 10 days, and 21 days. In the end, the decision was to include in the comparison the BRV models with the Δ parameters of 10 seconds and 5 minutes only and for the latter, the averaging parameters of 5 days and 21 days. Therefore, the investigation of the impact of averaging parameters focused on the realized volatility computed in the interval of 5 minutes that is averaged across different periods.

The great number of values of 4-error statistics used led to a special presentation of results aimed at identifying any emerging patterns. Thus, three methods are used to show the results. In the first approach (Figures 7 - 12), we present values of 3-error statistics: OP, RMSE, and MAPE. They have been separately calculated for six different models. Each figure contains five boxes presented for five moneyness classes. Vertical axes in each box show the values of a given statistic for six models (first horizontal axis) and five TTM classes (second horizontal axis). The models are always presented in the same order: 1. BRV10s, 2. BRV5m, 3. BRV5m_5, 4. BRV5m_21, 5. BHV, and 6. BIV. Within each figure, each box has the same scale on the vertical axis.

The second approach is to consider the pricing errors in a different way. Figure 13 shows MAPE statistics for PUT options with respect to TTM and MR for six models separately. This method enabled the identification of the effect of TTM and MR on pricing error and how this effect differs for different models. The third approach is to compare the BRV models only. Figure 14 shows RMSE statistics for call options with respect to a given model and TTM or MR. Such an approach makes it easier to observe the effect of averaging RV across different time horizons and the way pricing errors are affected.

Figure 7 shows values of the OP for call options. There are no significant differences for models with RV: BRV10s, BRV5m, and BRV5m_5. The most expected value of OP (approximately 0.5, the same fraction of over- and underprediction) has the BIV model. The BHV model and the BRV5m_21 model are slightly worse. Most models show underpredicted premia for TTM between 0-15 days and 16-30 days. Exceptions are the BIV model and all models for deep ITM class.

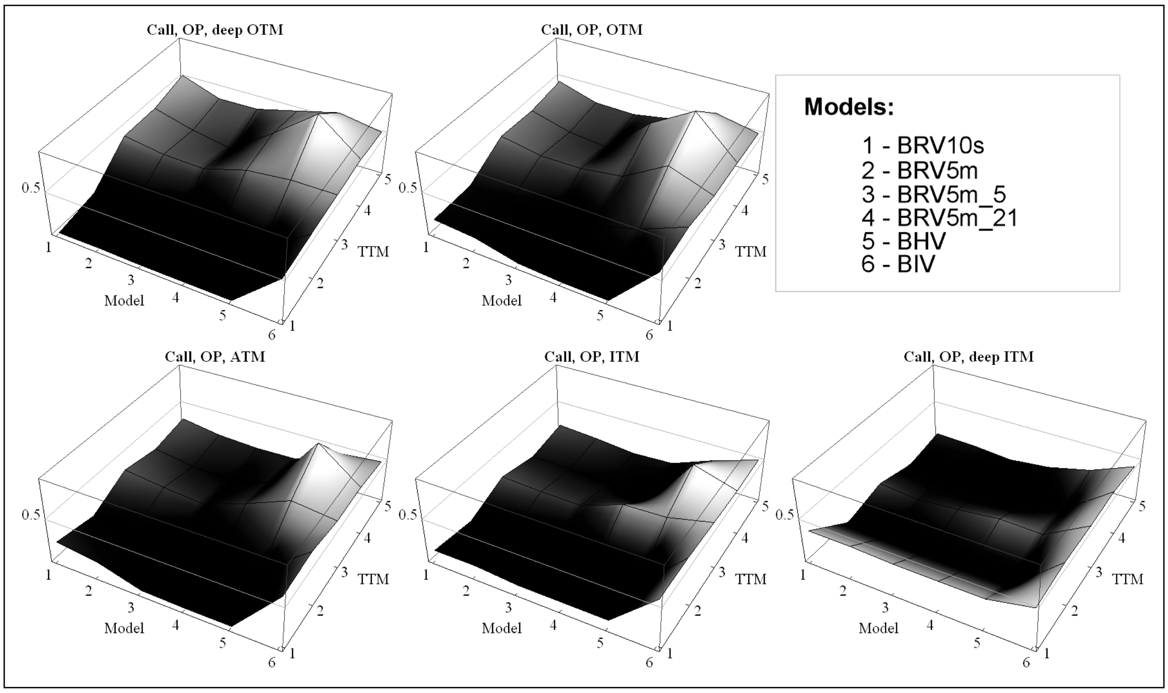


Figure 7. OP statistics for call options, all pricing models, TTM and MR.

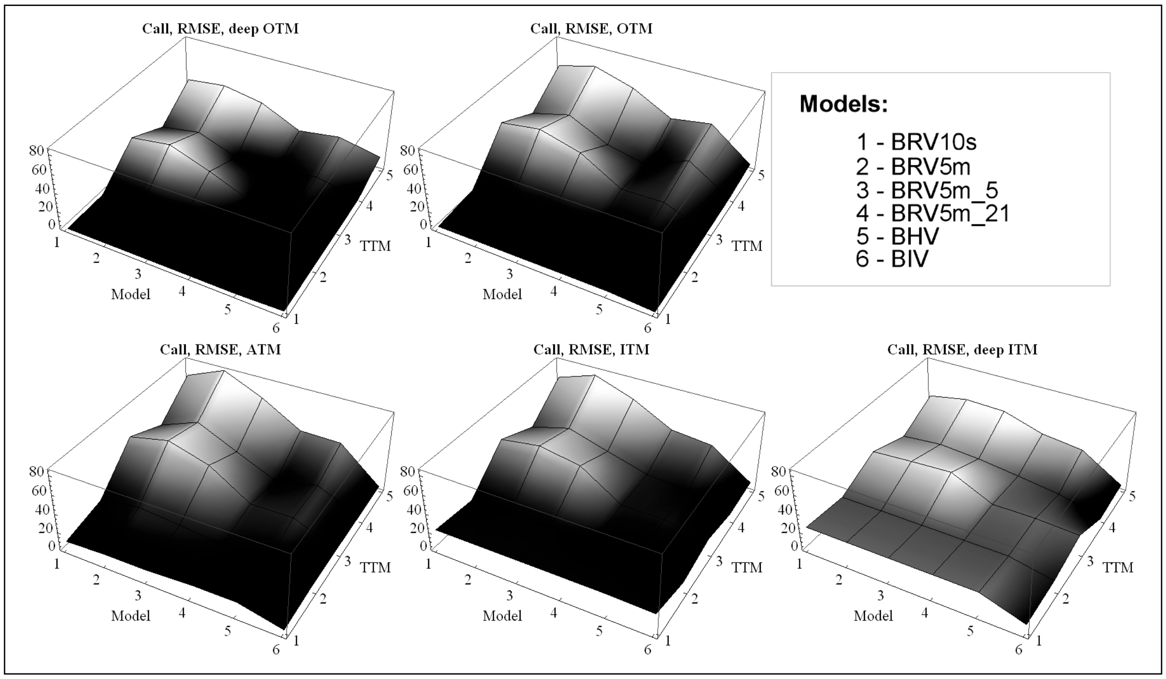


Figure 8. RMSE statistics for call options, all pricing models, TTM and MR.

Figure 8 shows the RMSE statistics for call options, which confirm the lack of significant differences for models with RV, for example, BRV10s, BRV5m, and BRV5m_5. Figures indicate the lowest values of RMSE for the BIV model and slightly higher values for the BHV and the BRV5m_21 models. Moreover, the pricing error increases with TTM though such an effect is not confirmed later on by MAPE statistics. Thus, this effect may arise only due to higher option values with relatively long TTM. Hence, we argue that the RMSE statistics used for comparing the pricing error for options with a different TTM is a misleading metric.

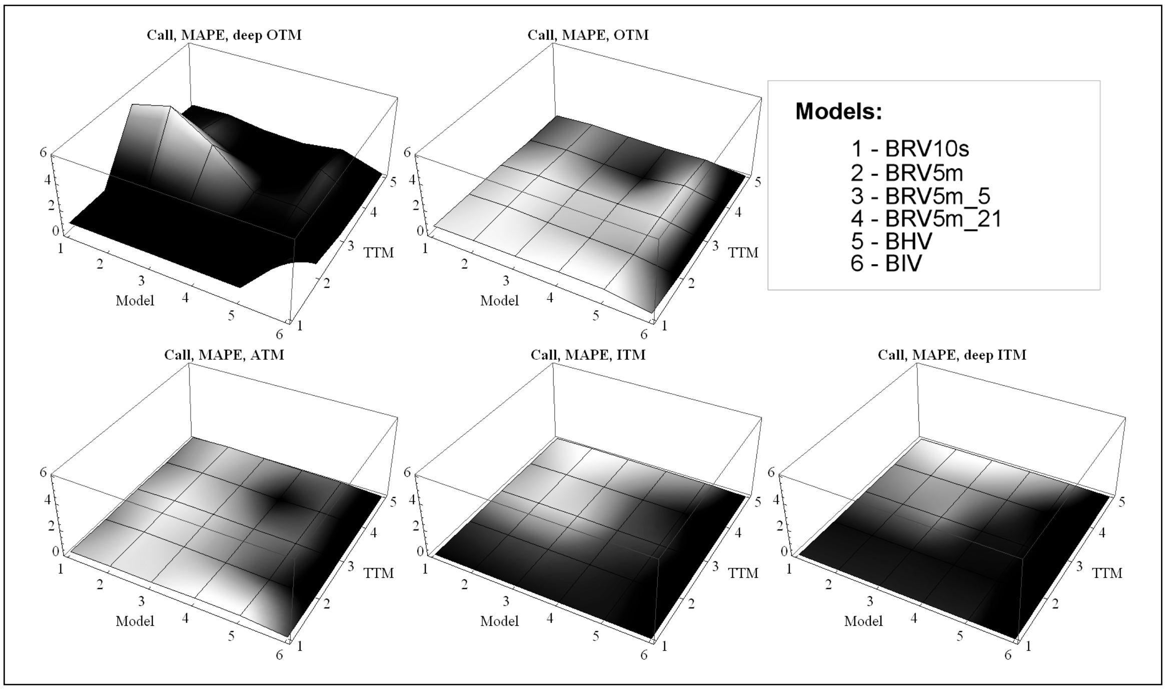


Figure 9. MAPE statistics for call options, all pricing models, TTM and MR.

Figure 9 shows values for the first relative statistic, MAPE. Again, no significant differences for models with RV (BRV10s, BRV5m, and BRV5m_5) can be found. Best results are oberved for the BIV model, and then for the BHV and the BRV5m_21 models. Very high MAPE values for the BRV models for deep OTM class appear to indicate the existence of two kinds of outliers in the initial data. We intentionally did not exclude them at this stage although they distort other results[17]. These outliers are the outcome of point estimates for volatility. That is why we call them *spurious outliers* as their cause is the specific nature of the BRV model and not the data pvalues themselves.

The first kind of spurious effect was observed for the BRV10s, BRV5m, and BRV5m_5 models and for a TTM between 31 and 60 days. For these classes, average pricing error values were up to 600%. The main reason for this error is that the BRV model is based on σ parameters (RV estimator) computed for the previous day (only one day, hence not averaged). During periods of high market volatility, differences between values of σ the parameter for two consecutive days are up to 50% and occasionally even higher. Enormous mispricings emerge as a result (even up to 40% in specific cases), which influence the average value of pricing error. It is worth investigating why such large errors appear when the TTM is between 31-60 days and only for deep OTM options. The reason is that during high market volatility periods in mid January 2008, options that matured in March (with a TTM between 31 and 60 days) were actually classified with TTM = 3 status. In addition, deep OTM options have the highest value of the relative Vega parameter, and hence their price is very sensitive with respect to changes of volatility of the futures prices of the market index.

The second kind of spurious effect is observed for the BIV model and for a TTM between 0 and 15 days. However, this effect is not visible for the RMSE statistic, hence it concerns low-priced options only. It is also not visible for the OP statistics, which means that the fraction of price underpredictions equals the OP. The pricing error is due to very high differences between the BIV model valuation and mid-quote when the former is lower than the latter. Such a situation appears mostly when the market-maker withdraws his bid offers. The new mid-quote is then calculated on the basis of an old ask offer and a new bid offer; in most cases, the new offer is significantly lower. As a result, mid-quote often changes by dozens of percentage points.

However, there is also an alternative explanation for the second kind of spurious effect, which is even more plausible when considering the properties of the BIV model. It may be partly explained by the characteristic path of the IV when the TTM is less than 10 or 15 days. In such periods, the IV simply explodes. As a result, for the low-priced deep OTM options, the pricing error related to the option price is so high that it could significantly alter an average value of MAPE statistics, even when the number of observations with extremely high volatility is relatively low.

The next four figures show four statistics for the put options in the same order.

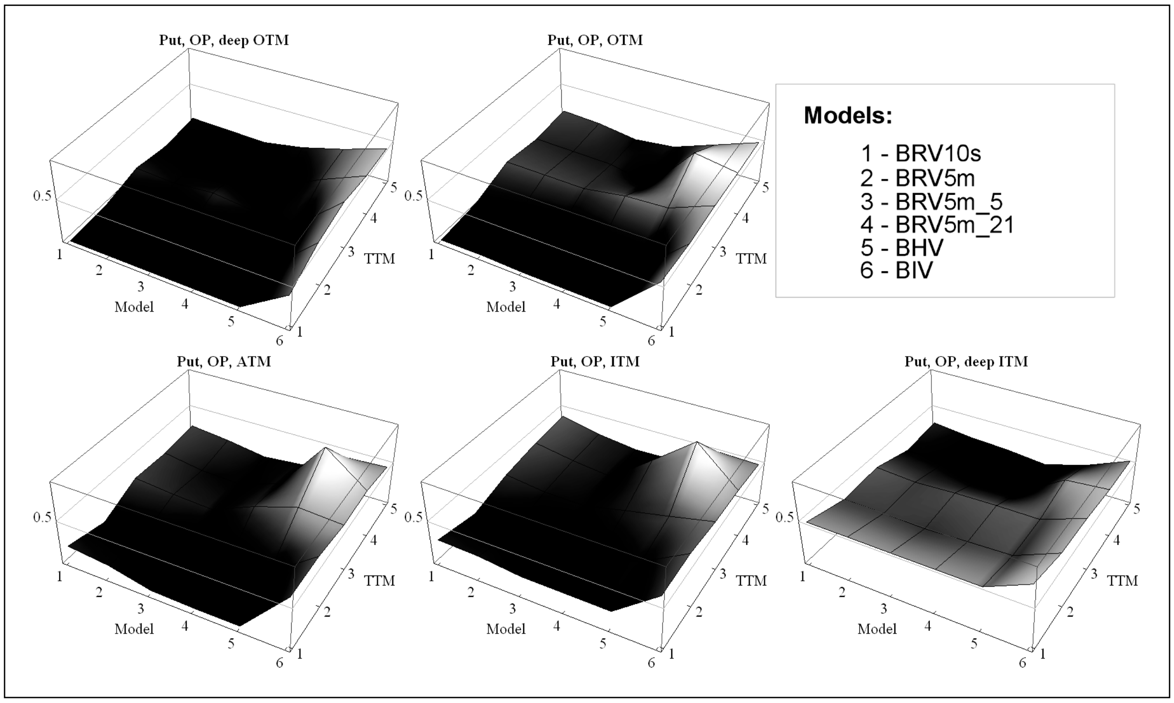


Figure 10. OP statistics for put options, all pricing models, TTM and moneyness classes.

Figure 10 does not show any significant differences between the OP statistics for all BRV models. Their theoretical premiums are on average underestimated when compared with the actual prices. Best results are observed for the BIV model, and then for the BHV and the BRV5m_21 models. On average, all models underestimate market prices; exceptions are options within the deep ITM class.

The reason for the high values of the OP statistics for the BHV model and a TTM between 61 and 90 days is that the prices of the BHV model are affected by the long-memory effect typical of the HV estimator. Opposed to that effect, market participants adjust to the new market volatility levels much more rapidly, which is reflected in mid-quotes. This phenomenon is actually related to one of the research questions regarding the optimal level of the *n* parameter, representing the long-memory effect of the volatility process.

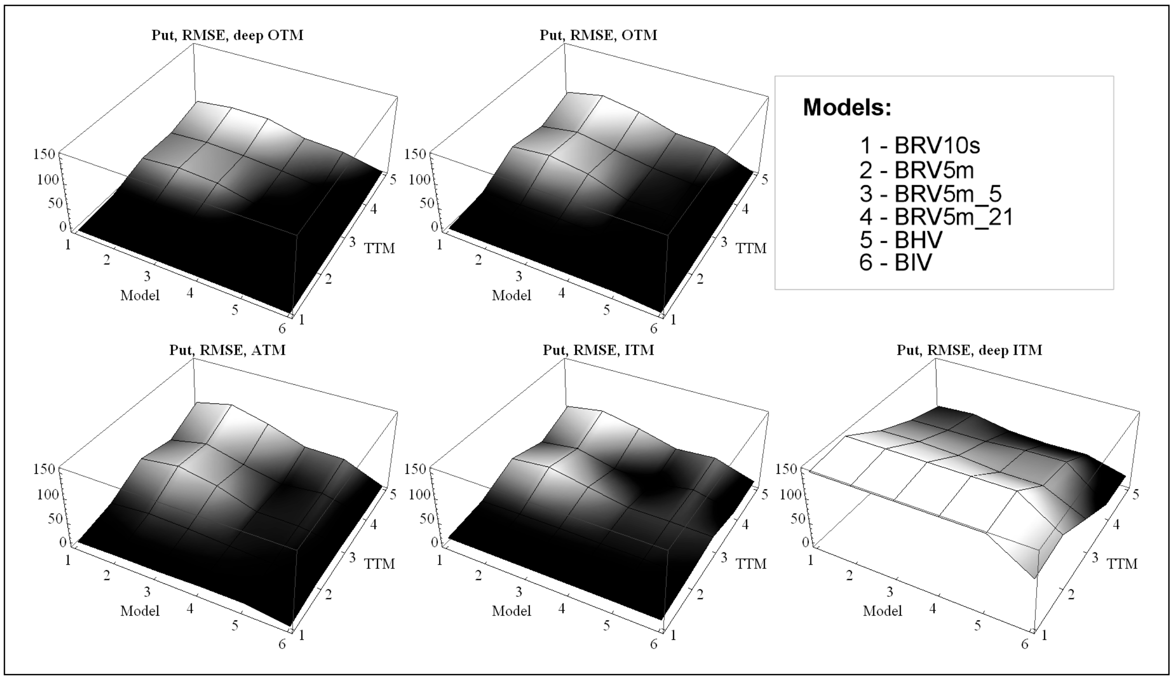


Figure 11. RMSE statistics for put options, all pricing models, TTM and MR.

RMSE values in Figure 11 show no significant differences between the BRV10s, the BRV5m and the BRV5m_5 models. The lowest values of the RMSE statistics have been obtained for the BIV model; they are slightly higher for the BHV model and even higher for the BRV5m_21 model. Similarly to the results for call options, the pricing error seems to increase with the TTM; exceptions are options within the deep ITM class. Again, the reason for that phenomenon could be much higher option prices when the TTM is relatively long. For that reason, we argue that a much better way to compare pricing errors for options with different TTM is to use relative statistics, like MAPE.

The last spurious effect is revealed through very high RMSE values for the BRV and the BHV models for the deep ITM class with a TTM between 0 and 15 days. Again, this phenomenon may be due to possible spurious outliers present in the data. This effect, however, is not visible for the MAPE statistics, which means that the effect concerns highly priced options only, where the pricing error related to the mid-quote is not as high as it is when computed in absolute values. Therefore, a possible explanation is a situation whereby the market maker withdraws his ask offers; hence, the mid-quotes can suddenly increase by a large amount. This effect is present not only for the TTM between 0 and 30 days, but also for other TTM classes, and is visible for the deep ITM options only because they have the highest prices; hence, their pricing errors in absolute values have the greatest effect on the RMSE values.

Two relative statistics once again confirm previous findings that there are no significant differences between the BRV10s, the BRV5m, and the BRV5m_5 models (see Figure 12). Similarly to the previous findings, the lowest values have been obtained for the BIV model, then for the BHV and BRV5m_21 models. This time, very high values are found for the deep OTM and OTM classes with a TTM between 0 and 60 days; exceptions are values for the BIV model. For these classes, the results show low OP values and low RMSE values, which means that model pricings were significantly smaller when compared with market prices. In contrast, the ITM and the deep ITM classes show values close to zero.

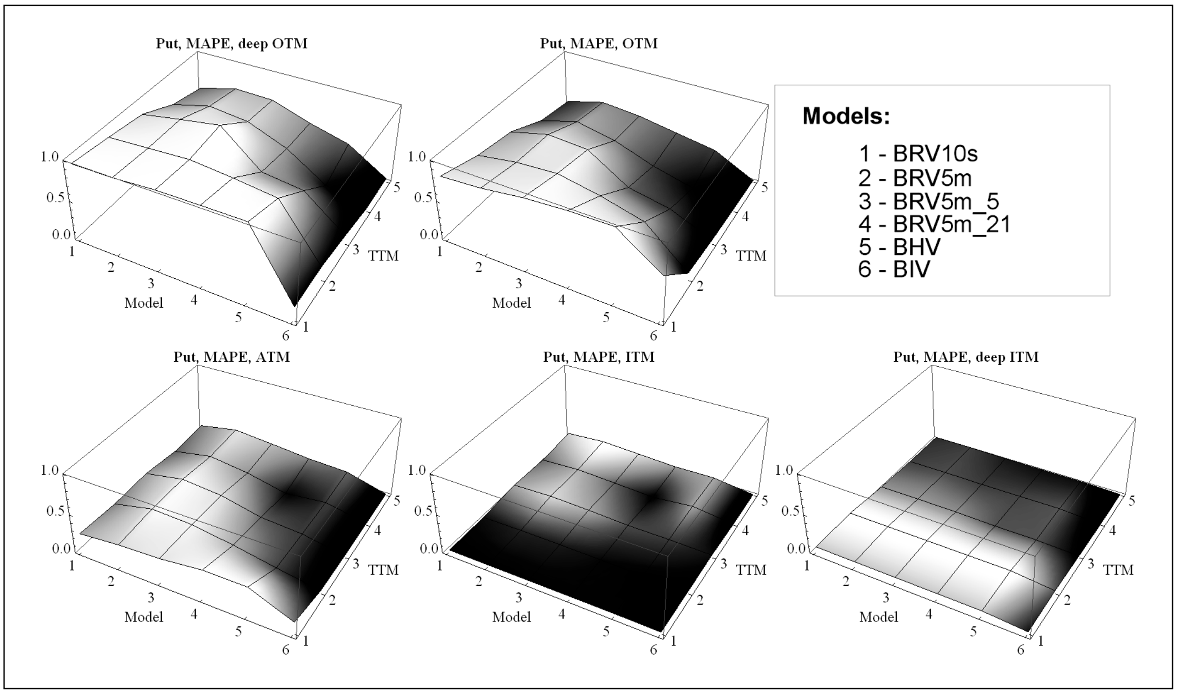


Figure 12. MAPE statistics for put options, all pricing models, TTM and MR.

Pricing errors were also analyzed according to other dimensions. Figure 13 shows MAPE statistics calculated for put options with respect to TTM and MR for six models separately. Once more, the BIV model appears to be the best model. The BHV model has slightly higher MAPE values. There are no significant differences among the first three BRV models.

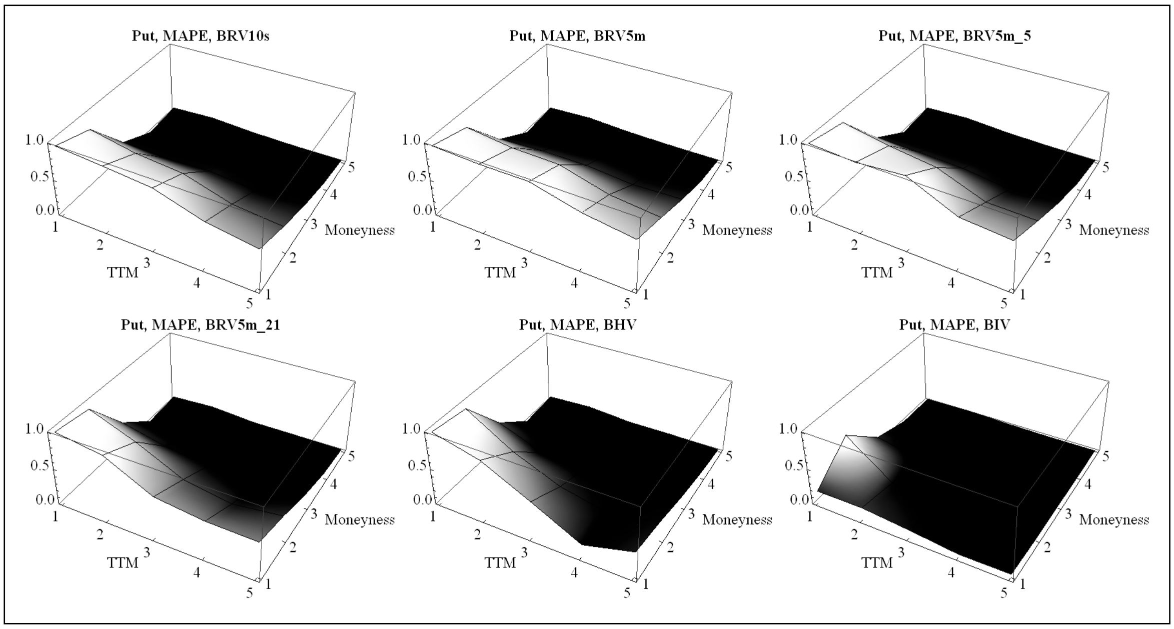


Figure 13. MAPE statistics for put options, all pricing models, TTM and MR.

A striking pattern is visible in this figure. The best model pricings are obtained for high TTM and MR while the highest error values are calculated for low TTM and MR classes.

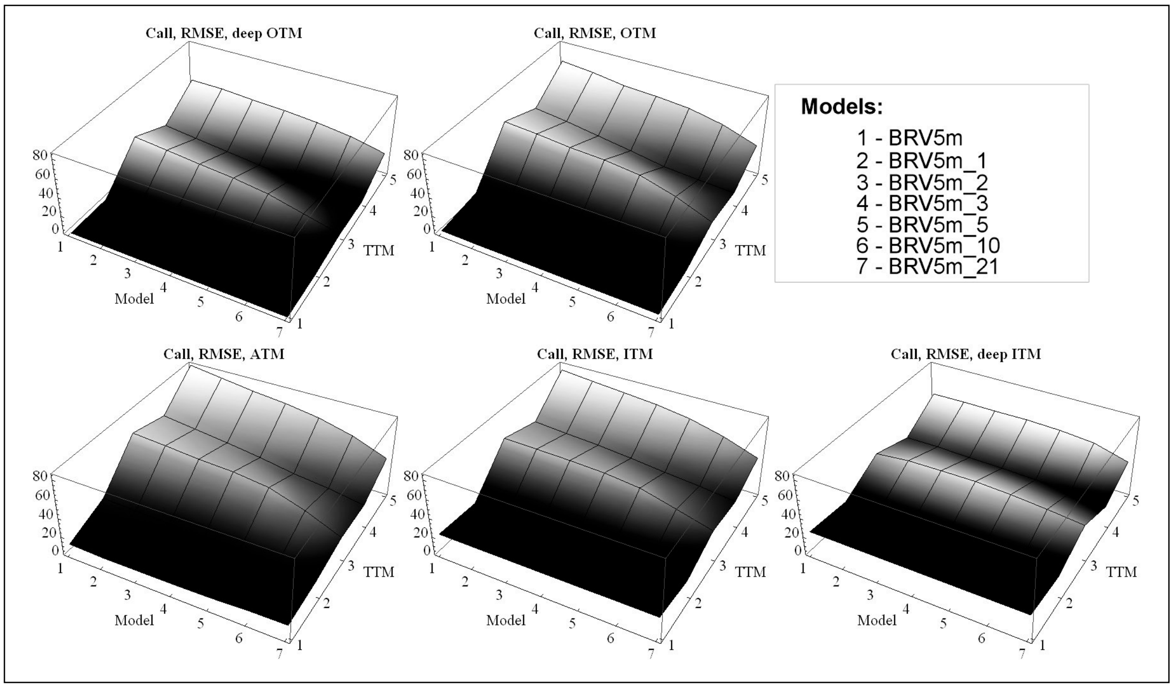


Figure 14. RMSE for call options, BRV model, different averaging parameters and MR.

Figure 14 shows a comparison of BRV models which differ by the averaging parameter *n* and are presented in the following order: 1. BRV5m, 2. BRV5m_1, 3. BRV5m_2, 4. BRV5m_3, 5. BRV5m_5, 6. BRV5m_10, and 7. BRV5m_21. The pricing error decreases as the averaging parameter increases: the smallest pricing error appears for the BRV5m_21 model; that is, the BRV model with 5-minute RV averaged across the last 21 days. In contrast, error values increase with higher TTM, but not in a stable way. Moreover, the MR does not influence the model quality.

The general conclusion is that, on average, the smallest pricing errors have been obtained for the BIV model. It has been the best model in most comparisons, regardless of the TTM and MR class and the type of option. The BHV model takes second place, and the third place goes to the BRV model averaged across the period of 21 days. The pricing errors for the latter model were only slightly worse than for the BHV model. The highest pricing error values were obtained for the Black model with non-averaged realized volatility. There are no significant differences between the Black models with non-averaged realized volatilities computed for 10-second and 5-minute intervals. This result implies that it is not necessary to calculate the RV for higher frequencies because the accuracy of volatility estimates can be offset by microstructure biases.

Another conclusion is that the averaging parameter used in calculating the RV has an important effect on pricing error. When RV is averaged across the period of 21 days, the error statistics are very similar to those obtained for the BHV model.

The investigation of values of error statistics with respect to different TTM and MR class showed that the pricing error is much smaller when the TTM is relatively long and the option belongs to the ITM and the deep ITM class. There were also spurious outliers; that is, observations with model valuations very distant from other theoretical prices and market mid-quotes within a given TTM and MR class. They are not excluded although, when present, they considerably complicate the analysis; when detected, their effect and the reasons for their occurrence have been noted.

Conclusions and Further Research

The application of HF (10s) data for WIG20 index options enabled the verification of the efficiency of different option pricing models. Various volatility processes (historical, realized, and implied) were applied to the Black model in order to check the research hypothesis. Volatility was also calculated in these models for different interval Δ and parameter *n* in order to study the influence of these parameters on final valuation and stability of volatility estimators. This paper also shows the analysis of results for five classes of the MR and the TTM in order to reveal some patterns in valuation and to explain the behavior of models used for options with different TTM and whole span of strike prices. Finally, the possibility of and the reason for outlier exclusion are discussed.

The findings support the initial hypothesis that the BIV model gives the best results, the BHV model is slightly worse, and BRV models give clearly the worst results, but the results for different concepts of BRV models differ significantly. These results are robust to changing TTM and MR and are confirmed by four different types of error statistics. The poor outcomes for the BRV model could result from the way the RV estimator is calculated as this estimator is characterized by very high volatility. RV can be described as a point estimate in comparison with HV which is rather a range estimate, and the analysis has shown that this characteristic of RV is responsible for the worst results. Focusing on the BRV models, the best results were obtained for averaged models with the largest parameter n tested for ($n = 21$). This value of n makes the result for the BRV5m_21 closer to the BHV model but does not clearly show the best value for the parameter describing the process memory in volatility estimation. This issue requires further detailed studies.

The results for different classes of the TTM and the MR show some patterns of valuation in the case of put options. There is a clear relation between the model error and the TTM, and between the model error and the MR, which can be described briefly as follows: high error values for low TTM and MR and best fit for high TTM and MR. The incidence of this pattern for put options only does not mean that call options are not characterized by similar behavior. However, the existence of possible outliers makes it impossible to reveal an analogous pattern for call options. The multidimensional presentation of raw data leads to the indication of some spurious outliers that in fact are not true outliers at all. They result from the model misspecification (e.g., not the appropriate volatility estimator) and can change the final evaluation of the specific model efficiency when excluded from further calculations.

The results presented in this paper and the significant lack of studies testing various option pricing models for emerging CEE markets data suggest several avenues for further research studies. First, researchers should test other option pricing models like GARCH, based on the methodology presented in Duan (1995), and SV models (Heston, 1993; Hull & White, 1987). Second, there is a space for models with different assumptions of volatility distributions taking into account not only the rigorous definitions of parameters and Δ interval in RV (Ślepaczuk & Zakrzewski, 2009) or IV estimator calculations but also the latest results concerning HF and model-free volatility indexes, based on VIX index methodology. Third, more effort should be applied to outlier identification because it is not clear whether one can exclude observations for which the TTM is lower than a specific number of days (e.g., 5 or 10 days), MR is not limited by the fixed interval (e.g., 0.8 to 1.2), and the market premium is lower than some established value (e.g., 5, 10, or 15 pts), as is assumed in many other research papers without any further investigations.

Fourth, the results for bid-ask quotes tested in this study should be compared with the transactional prices. This analysis will significantly decrease the number of observations but will enable the verification of the results for real market behavior in addition to the potential one. Time needed for calculations and disk space limitation made it impossible to conduct this study for the whole period of index options quotations on the WSE (starting from September, 2003), but we are aware that option pricing models can behave differently in high or low volatility environments or for upward or downward market trends. This issue requires further investigation. Results could significantly differ for markets with various degrees of efficiency, so conducting a similar study for other markets in different countries (emerging and developed markets with different depths in terms of liquidity) should be the subject of further analyses. Moreover, results obtained in terms of the efficiency of IV estimators in comparison with HV and RV could be additionally verified by simple econometric regression of future RV (as dependent variable) on HV, RV, or IV (as explanatory variables). The last issue which could be developed further is the verification of the statistical significance of our results, which can be done in several ways:

- estimating econometric models where we try to explain the magnitude of error statistics (RMSE, OP, or MAPE) with respect to TTM, MR, and dummy variables for each model separately; and
- the analysis of variance for each error statistic for call and put models separately, in order to additionally confirm the differences between models and patterns in valuation which is indicated in the figures in the results section.

This study is one of the first to compare the option pricing models on HF data for Eastern European emerging markets, so further research is needed.

Endnotes

1 Cf. the views of market participants quoted in Goebel (2010). Data also support this view, as the most recent report of the World Federation of Exchanges (2011) shows that the number and the notional turnover of contracts for stock index options and futures traded in Warsaw is much higher than in Budapest, Moscow, Vienna, and Athens.

2 The WIG20 is the index of the 20 largest companies on the WSE (further detailed information may be found at www.gpw.pl).

3 As Brenner, Courtadon, and Subrahmanyam (1985) showed, "the value of a European option on the cash instrument will always be equal to the value of a European option written on the futures contract based on the instrument, provided that the two options have the same maturity date and exercise price" (p. 1305). Cf. also a textbook approach to this question in Hull (2008).

4 We assume that we have 252 trading days in one calendar year.

5 In this study, RV is calculated for $\Delta = 10s, 1m, 5m, 15m$. However, averaging has been done for 5-minute intervals and n days only, where $n = 1, 2, 3, 5, 10$, and 21. It is widely accepted in financial literature that an interval between 5 and 15 minutes makes the consensus between the nonsynchronous bias and other microstructure biases.

6 Initially, we calculated the BRV model with different Δ: 10s, 1m, 5m, and 15m. We have checked the properties of average RVs with different values of parameter n in option pricing models. Therefore, we calculated BRV models based on 5m interval with different values of averaging parameter ($n = 1, 2, 3, 5, 10$, and 21). As a result, we obtained the following seven BRV models: BRV5m, BRV5m_1, BRV5m_2, BRV5m_3, BRV5m_5, BRV5m_10, and BRV5m_21. In Section 5, we present all these models in two set of comparisons:

1. BRV10s, BRV5m, BRV5m_5, BRV5m_21, BHV, and BIV – in order to choose the best model;
2. BRV5m, BRV5m_1, BRV5m_2, BRV5m_3, BRV5m_5, BRV5m_10, BRV5m_21 – in order to reveal the properties of averaging.

7 The study is based on the separate time series for futures contracts (F_1 – the expiration date is March 21, 2008, F_2 - the expiration date is June 20, 2008, and F_3 – the expiration date is September 19, 2008) where the choice of specific futures contract depends on the expiration date (the same date as for the options).

8 In practice, the continuous trading stops at 4:10 p.m., then the closing price is set between 4:10 p.m. and 4:20 p.m., and the next investors can trade until 4:30 p.m. only on the basis of the closing price.

9 Mid quote = (bid + ask) / 2.

10 The continuous time series for futures contracts was created based on the notion that the expiring futures contract was replaced by the next series.

11 Analyzing both return series, one can see high kurtosis, enormous Jarque-Berra statistics, and high negative (in case of R_f) or high positive (in case of R_{f^*}) skewness. Mean returns are small and are not significantly different from zero. Distributions of both time series are leptokurtic; that is, they have fat tails and a substantial peak at zero.

12 IV for put options reveals the same pattern. The additional figure is available upon request.

13 The first set of comparisons was prepared for BHV, BRV10s, BRV1m, BRV5m, BRV15m, and BIV, but the results were very similar to those presented in this paper, with BRV models as clearly the worst (the detailed results are available upon request). Then, we decided to present the results for models with averaged value of RV estimator. In the final part of the results section, we present the comparison only for the BRV models with RV5m with different values for parameter n. They are presented to show the properties of averaging the volatility parameter in the process of option pricing.

14 The MR is calculated according to the following formula, which was adjusted for the use of futures contract as the basis instrument:

$$MR = \frac{S}{K / e^{rT}} = \frac{F}{K} \qquad (15)$$

15 The detailed figures concerning the issue are available from the authors.

16 Taking into account that this research is based on bid-ask quotes, we wanted to have some reference to market liquidity through the presentation of the fraction of quoted options. Available strike prices mean the span of strike prices which were available to trade for market participants, whether they were quoted or not. Active mid-quotes stand for options with bid-ask quotes that were actually quoted.

17 These outliers can additionally be the reason that the patterns of pricing presented for put options in Figure 20 are not revealed for call options.

References

Ait-Sahalia, Y., & Mykland, P. A. (2009). Estimating volatility in the presence of market microstructure noise: a review of theory and practical considerations. In T. G. Andersen, R. A. Davis, J. P. Kreiss, & T. Mikosch (Eds.), *Handbook of Financial Time Series* (pp. 577-598). Berlin, Germany: Springer-Verlog.

Ammann, M., Skovmand, D., & Verhofen, M. (2009). Implied and realized volatility in the cross-section of equity options. *International Journal of Theoretical and Applied Finance, 12*(6), 745-765.

An, Y., & Suo, W. (2009). An empirical comparison of option pricing models in hedging exotic options. *Financial Management*, *38*(4), 889-914.

Andersen, T. G., Frederiksen, P., & Staal, A. D. (2007). *The information content of realized volatility forecasts*. Unpublished manuscript, Kellogg School of Management, Northwestern University.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *The Journal of Finance*, *52*(5), 2003-2049.

Bams, D., Lehnert, T., & Wolff, C. C. P. (2009). Loss functions in option valuation: a framework for selection. *Management Science*, *55*(5), 853-862.

Bates, D. S. (2003). Empirical option pricing: a retrospection. *Journal of Econometrics, 16*(1-2), 387-404.

Berkowitz, J. (2010). On justifications for the ad hoc Black-Scholes method of option pricing. *Studies in Nonlinear Dynamics and Econometrics*, *14*(1), 4. Retrieved from http://www.bepress.com/snde/vol14/iss1/art4

Black F. (1976). The pricing of commodity contracts. *Journal of Financial Economics, 3*(1-2), 167-179.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy*, *81*, 637-654.

Brandt, M. W., & Wu, T. (2002). Cross-sectional tests of deterministic volatility functions. *Journal of Empirical Finance, 9*(5), 525-550.

Broadie, M., & Detemple, J. B. (2004). Option pricing: valuation models and applications. *Management Science*, *50*(9), 1145-1177.

Brenner, M., Courtadon, G., & Subrahmanyam, M. (1985). Options on the spot and options on futures. *Journal of Finance*, *40*(december), 1303-1317.

Dennis, P., & Mayhew, S. (2009). Microstructural biases in empirical tests of option pricing models. *Review of Derivatives Research, 12*(3), 169-191.

Duan, J. (1995). The GARCH option pricing model. *Mathematical Finance, 5*(1), 13-32.

Ferreira, E., Gago, M., Leon, A., & Rubio, G. (2005). An empirical comparison of the performance of alternative option pricing models. *Investigaciones Economicas, 29*(3), 483-523.

Fiszeder, P. (2010). Pricing the WIG20 Index Options Using GARCH Models. In: Financial Markets. Principles of Modelling, Forecasting and Decision-Making, *FindEcon Monograph Series, Advances in Financial Market Analysis*, Number 8, (pp. 141-156). Uniwersytet Łódzki.

Garcia, R., Ghysels, E., & Renault, E. (2010). The econometrics of option pricing. In Y. Ait-Sahalia & L.Hansen (Eds.), *Handbook of Financial Econometrics* (pp. 479-552). Oxford, UK, & Amsterdam, the Netherlands: North Holland.

Goebel, M. (2010, November 30). CEE: rivals vie for the space between Frankfurt and Moscow. *FOW: The Global Derivatives Magazine*. Retrieved August 30, 2011, from http://www.fow.com/Article/2723404/Themes/26527/CEE-rivals-vie-for-the-space-between-Frankfurt-and-Moscow.html

Han, C. (2008). *The comparisons between three option pricing models and Black-Scholes formula in pricing performance and trading strategy: evidence from the FTSE 100 options* (Unpublished master's thesis). National Chung Cheng University, Minhsiung, Taiwan.

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*(2), 327-343.

Hull, J. C. (2008). *Options, futures, and other derivatives* (7th Ed.). Upper Saddle River, NJ: Prentice Hall.

Hull, J., & White, A. (1987). The pricing of options with stochastic volatilities. *Journal of Finance, 42*(2), 281-300.

Martens, M., & Zein, J. (2004). Predicting financial volatility: high-frequency time-series forecasts vis-à-vis implied volatility. *Journal of Futures Markets*, *24*(11), 1005-1028.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics and Management Science*, *4*(1), 141-183.

Mitra, S. (2009). *A review of volatility and option pricing*. Retrieved from http://arxiv.org/abs/0904.1292

Mixon, S. (2009). Option markets and implied volatility: past versus present. *Journal of Financial Economics, 94*(2), 171-191.

Osiewalski, J., & Pipień, M. (2003). Bayesian analysis and option pricing in univariate GARCH models with asymmetries and GARCH-in-mean effects. *Przegląd Statystyczny*, *50*(3), 5-29.

Raj, M., & Thurston, D. C. (1998). Transactions data examination of the effectiveness of the Black model for pricing options on Nikkei index futures. *Journal of Financial and Strategic Decisions*, *11*(1), 37-45.

Ślepaczuk, R., & Zakrzewski, G. (2009). *High-frequency and model-free volatility estimators* (Working Paper No. 09-13). Warsaw, Poland: University of Warsaw, Faculty of Economic Sciences. Retrieved from http://www.wne.uw.edu.pl/inf/wyd/WP/WNE_WP23.pdf

World Federation of Exchanges. (2011, April). *FOCUS No. 218, the monthly newsletter of regulated exchanges, with key market figures*. Retrieved from http://www.world-exchanges.org/files/focus/pdf/focus_april_white.pdf

Authors Note

Ryszard Kokoszczyński, Paweł Sakowski, and Robert Ślepaczuk, Faculty of Economic Sciences, University of Warsaw, Długa 44/50, Warsaw 00-241, Poland.

Correspondence concerning this article should be addressed to Robert Ślepaczuk, Email: rslepaczuk@wne.uw.edu.pl