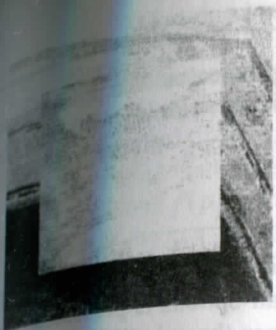
MACROMODELS' 2006
Zakopane 2006, pp.137–153



Piotr Płuciennik, Eliza Buszkowska
Adam Mickiewicz University, Poznań, Poland

Modeling and Forecasting the Implied Volatility of the Wig20 Index

Abstract. The implied volatility is one of the most important notions in the financial market. It informs about the volatility forecasted by the participants of the market. In this paper we calculate the daily implied volatility from options on the WIG20 index. First we test the long memory property of the time series obtained in such a way, and then we model and forecast it as an ARFIMA process.

Keywods: implied volatility, long memory, ARFIMA, GARCH, forecasting, stock index.

1. INTRODUCTION

For the last twenty years we could observe dynamic development of the theory and the practice of the financial markets. One of the most important notions, which was intensively developed is the market risk connected with the change of the financial instruments value. The basic measure of this risk is volatility. A special case of volatility is the implied volatility, a certain kind of volatility obtained from option values. Therefore, the implied volatility can be interpreted as the volatility forecasted by the participants of the market. One of the most important classes of the volatility forecasting models is the class of parametric volatility models. Options in the Polish market have been listed since September 22, 2003. Therefore, we can construct quite a long time series of implied volatility. In this paper we try to model and forecast implied volatility using parametric volatility models. We would like to prove that the implied volatility series have long memory dependence. It means that the correlogram of implied volatility decreases hyperbolically as the length of lag increases. The process can be well modeled with ARFIMA(p,d,q) model, in particular with $d \in (0,0.5)$.

2. IMPLIED VOLATILITY

In this section the method of calculation the implied volatility on the basis of the Black-Scholes pricing formula for the put and call options is presented. By the *implied volatility* of a financial instrument we mean the volatility calculated from the market price of a derivative written on this instrument, basing on some theoretical pricing model. In our case, put and call options are used. The implied volatility has usually a larger value than the realized volatility, whenever participants of market forecast the situation that the future volatility of stocks is bigger than volatility observed in the past.

In the Black-Scholes model (see Black and Scholes (1973)), the value of a European call option on a non-dividend paying stock is given by the formula

$$C_0^E = SN(d_1(S,T)) - K\exp(-rT)N(d_2(S,T)) \tag{2.1}$$

and the value of a European put option on a non-dividend paying stock we obtain from the formula

$$P_0^E = SN(d_1(S,T)) - K\exp(-rT)N(d_2(S,T)), \tag{2.2}$$

where

$$d_1(S,T) = \frac{\ln(S/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}}$$

and

$$d_2(S,T) = d_1(S,T) - \sigma\sqrt{T} = \frac{\ln(S/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}}$$

The stock price, the strike price, the interest rate, and the time to expiration are denoted by S, K, r, T, respectively, and $N(x)$ is the cumulative distribution function of the standard normal distribution evaluated at x. The annualized variance of continuously compounded return on the underlying stock is denoted by σ^2.

Since a closed-form solution of the implied standard deviation from the Equation (1.1) is not possible, the implied volatility must be calculated numerically. Manaster and Koehler (1982) discussed the Newton-Raphson algorithm for calculating the implied volatility. In the algorithm, we have the following formula

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})}$$

In our case for, European call options we have

$$\sigma_n = \sigma_{n-1} - \frac{SN(d_1(S,T)) - K\exp(-rT)N(d_2(S,T)) - C_0^E}{V_C}, \qquad (2.3)$$

where V_C is V-sensitivity coefficient, and it equals

$$V_C = \frac{\partial C_0^E}{\partial \sigma} = Sn(d_1(S,T))\sqrt{T}.$$

S. Manaster and G. Koehler proposed also the optimal start value for the algorithm

$$\sigma = \left[\left|\ln\frac{S}{K} + rT\right|\frac{2}{T}\right]^{\frac{1}{2}}.$$

The implied volatility from the put options can be calculated from the Black-Scholes formula for put options analogously.Unfortunately, all models basing on the Black-Scholes formula have introduced bias into the implied volatility measure. The reason of it is the assumption that volatility in the Black-Scholes model is constant. J. Hull and A. White (1987) found that the magnitude of the bias in models basing on the Black-Scholes formula is the smallest for near-at-the-money and close-to-maturity options. Therefore, only such options are taken into account in this paper.

2.1. Total implied volatility for class of options

There are often more than one option for some of the underlying instruments traded on the stock market. For this reason we must choose one of the methods to calculate the total implied volatility. Each of them gives a little different result because they have a different bias. Here we present the two following estimators.

The first one was proposed by R. Schmalensee and R. Trippi (1978). The authors calculated the total implied volatility as an arithmetic mean of the implied volatility obtained from all known options. This method has a large bias, so it is less useful in practice.

More precise value of the total implied volatility gives an estimator proposed by Chiras and Manaster (1978). It is the weighted mean with weights called flexibility coefficients of standard deviations. The estimator can be expressed by the following formula

$$\tilde{\sigma}_{kt} = \frac{\sum_{i=1}^{N_{kt}} \left(\frac{\partial C_{ikt}}{\partial \sigma_{ikt}} \frac{\sigma_{ikt}}{C_{ikt}} \sigma_{ikt} \right)}{\sum_{i=1}^{N_{kt}} \left(\frac{\partial C_{ikt}}{\partial \sigma_{ikt}} \frac{\sigma_{ikt}}{C_{ikt}} \right)}, \tag{2.4}$$

where $\frac{\partial C_{ikt}}{\partial \sigma_{ikt}} \frac{\sigma_{ikt}}{C_{ikt}}$ for $i = 1,\ldots,n$ are flexibility coefficients.

2.2. Algorithm to determine risk-free interest rate

One of the most important things in calculating implied volatility is a good selection of risk-free interest rate. In this paper we apply an algorithm presented in a booklet published by the Warsaw Stock Exchange. This algorithm requires current WIBOR and WIBID rates for periods of 1 week, 2 weeks, 1 month, 3 months and 6 months.

The algorithm consists of the following steps.

I. Convert all rates from annual capitalization to permanent capitalization.

II. Calculate WIMEAN rates as an arithmetical mean of converted WIBOR and converted WIBID for the all time periods.

III. Calculate the risk-free interest rate with the formula

$$\hat{r} = \frac{(r_1 - r_2)(t_s - t_1)}{(t_2 - t_1)} + r_1,$$

where

t_s is the time to expiration of the option taken to calculate of the implied volatility,

t_1 is the longest period shorter than the time to expiration of option taken to evaluate the implied volatility,

t_2 is the shortest period longer than the time to expiration of option taken to evaluate the implied volatility,

r_1 is the current WIMEAN rate for the period t_1,

r_2 is the current WIMEAN rate for the period t_2.

3. ARFIMA MODEL

The fact that time series exhibits long memory dependence suggests a possibility of describing this time series dynamics with ARFIMA models (Granger and Joyeux 1980, Hosking 1981). Let $\phi(L)$ and $\theta(L)$ denote lag-polynomials of the form

$$\eta(L) = 1 - \eta_1 L - \ldots - \eta_r L^r$$

and of the order p and q, respectively. The common ARFIMA(p,d,q) model is given by the formula

$$\phi(L)(1-L)^d (y_t - \gamma) = \theta(L)\varepsilon_t, \tag{3.1}$$

where ε_t is a sequence of independent and identically distributed random variables and d is the fractional integration parameter. The expression $(1-L)^d$ is called the fractional difference operator and equals

$$(1-L)^d = \sum_{i=0}^{\infty} b_i L^i, \tag{3.2}$$

where

$$b_0 = 1 \text{ and } b_i = \frac{-d\Gamma(i-d)}{\Gamma(1-d)\Gamma(i+1)} = \frac{i-d-1}{i} b_{i-1} \text{ for i} \geq 1.$$

Since all the time series are finite in practice, the fractional difference operator is approximated by replacing $\propto$ by $t-1$ in (3.2).

4. METHODS OF DETECTING THE EXISTENCE OF LONG MEMORY DEPENDENCE

The oldest and best-known method for detecting long memory is the R/S analysis. This method was introduced by Mandelbrot and Wallis (1968). It based on previous hydrological analysis of Hurst (1951). Suppose that x_t, is a time series where $t = 1,2,\ldots,n$. The rescaled range statistic is defined by the formula

$$(R/S)_n = \frac{1}{S}\left\{\max_{1\le i\le n}\sum_{t=1}^{i}(x_t-\bar{x}) - \min_{1\le i\le n}\sum_{t=1}^{i}(x_t-\bar{x})\right\},$$

where $\bar{x}$ is the sample mean of the time series and S is the sample standard deviation. The R/S statistics asymptotically follows the relation

$$(R/S)_n = Cn^H.$$

The value of H is obtained from the linear regression over a sample of growing temporal horizons $s = t_1, t_2, \ldots, n$. Hence we have

$$\ln(R/S)_s = \ln(C) + H\ln(s).$$

The estimated value of $H \le 0.5$ means that the time series has no memory. Other the value of H mean that the process has a long memory. Lo (1991) proposed modified R/S test. He introduced the statistic

$$Q_n = \frac{1}{s_n(q)}\left\{\max_{1\le i\le n}\sum_{t=1}^{i}(x_t-\bar{x}) - \min_{1\le i\le n}\sum_{t=1}^{i}(x_t-\bar{x})\right\},$$

where

$$s_n^2(q) = \frac{1}{n}\sum_{t=1}^{n}(x_t-\bar{x})^2 + \frac{2}{n}\sum_{t=1}^{q}\omega_t(q)\left(\sum_{i=t+1}^{n}(x_t-\bar{x})(x_{i-t}-\bar{x})\right) = \hat{s}_x^2 + 2\sum_{t=1}^{q}\omega_t(q)\hat{\gamma}_x,$$

and $\omega_t(q) = 1 - \dfrac{t}{q+1}$.

$\hat{s}_x^2$ and $\hat{\gamma}_x$ are the usual sample variance and autocovariance estimators of x_t. If the series has no long-range dependence, given the right choice of q, the distribution of $n^{-0.5}Q_n$ is asymptotically equal to that of

$$W = \max_{0 \le r \le 1} V(r) - \min_{0 \le r \le 1} V(r),$$

where V is the standard Brownian bridge. The distribution of the random variable W is expressed by the formula

$$P(W \le x) = 1 + 2\sum_{j=1}^{\infty}\left(1 - 4x^2 j^2\right)\exp\left(-2x^2 j^2\right)$$

Practical values of the distribution are tabulated in Lo(1991). We use the Q_n statistic for testing the null hypothesis that there is only a short-term memory in a given time series.

A semi-parametric approach to test long-memory was proposed by Geweke and Porter-Hudak (1983). Take a fractionally integrated process x_t. The spectral density of this process is given by following formula

$$f(\omega) = \left[2\sin(\omega/2)\right]^{-2d} f_u(\omega), \tag{4.1}$$

where ω is the Fourier frequency, $f_u(\omega)$ is the spectral density corresponding to u_t and u_t is a stationary short memory disturbance with mean zero. Taking the natural logarithm of each side and simplifying we get

$$\ln f(\omega_j) = \ln f_u(\omega_j) - d\ln\left[4\sin^2(\omega_j/2)\right], \tag{4.2}$$

where ω_j is the set of harmonic frequencies and $\omega_j = 2\pi j/n$ for $j = 0, 1, \ldots, n/2$, where n is the sample size. Geweke and Porter-Hudak (1983) showed that using a periodogram estimate of $f(\omega)$, if the number of frequencies used in the regression (4.2) is a positive integer function $g(n) = \left[n^{\alpha}\right]$ where $0 < \alpha < 1$, the least square estimate $\hat{d}$ obtained using the regression equation constructed from the (4.2) is asymptotically normaly distributed in large samples

$$\hat{d} \sim \mathrm{N}\left(d, \frac{\pi^2}{6\sum_{j=1}^{g(n)} \left(U_j - \overline{U}\right)^2} \right),$$

where

$$U_j = \ln\left[4\sin^2(\omega_j/2)\right]$$

and $\overline{U}$ is the sample mean of U_j for $j = 1,2,\ldots,g(n)$. Using this fact, we can test the null hypothesis that the time series has no long memory ($d = 0$) which is true if the t-statistic

$$t_{d=0} = \hat{d} \cdot \left(\frac{\pi^2}{6\sum_{j=1}^{g(n)} \left(U_j - \overline{U}\right)^2} \right)^{-0.5}$$

has a limiting standard normal distribution.

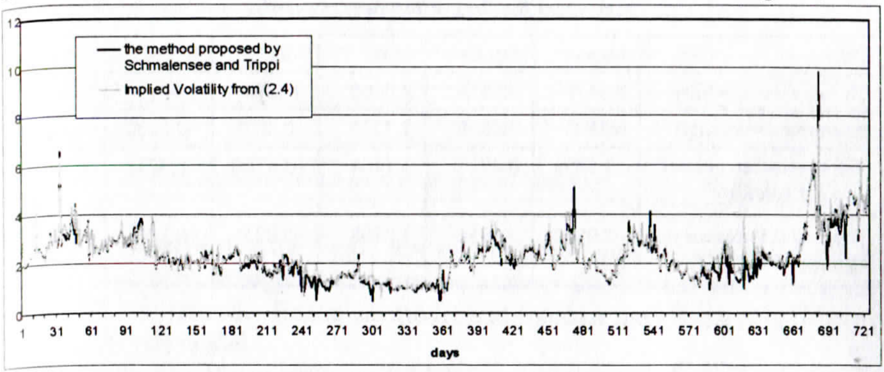
5. DATA

Our purpose is modeling and forecasting the implied volatility (or logarithm of the implied volatility) from option on the WIG20 index using the ARFIMA model. First we test the long memory dependence of the time series. Using the formula (2.4), we obtain the implied volatility from known option prices. We always take into account the nearest-at-the-money and the closest-to-maturity call option and put option. The risk-free interest rates are calculated on the basis of the algorithm presented in subsection 2.2. The Total Implied Volatility is calculated with the formula (2.4) and with the method proposed by R. Schmalensee and R. Trippi (1978).

Options on the WIG20 index have been listed since September 22, 2003. Therefore, till August 4, 2006 there are 725 quotations. Results concerning the implied volatility are presented in the Figure 1.

Figure 1.

Total daily implied volatility of WIG20 index since September 22, 2003 till August 4, 2006.



Before we modeled the series we tested stationarity of the series. We also took into account the logarithm of the implied volatility, and the processes of the first differences. We used the KPPS (Kwiatkowski et al. (1992)) test. The results are presented in Table 1.

The descriptive statistics of the implied volatility series and the logarithm of the implied volatility series are presented in Table 1.

Table 1.

The results of stationarity test for implied volatility, the logarithm of the implied volatility, and the processes of first difference

Data	KPPS
The implied volatility	2.1812 (<0.01)
ln(implied volatility)	2.1566 (<0.01)
The first differences of the implied volatility	0.0125 (<1)
The first differences of ln(the implied volatility)	0.0140 (<1)

The null hypothesis of stationarity is rejected for the implied volatility and the logarithm of the implied volatility series, but for the processes of the first difference we can't reject the H_0 hypothesis. It means that only this process can be modeled and forecasted.

Table 2.

The descriptive statistics of the implied volatility series and logarithm of implied volatility series and the first differences processes

Data	Mean	Std.dev.	Max.	Min.	Skew.	Kurtosis
The implied volatility	2.3479	0.9339	8.6066	0.6294	1.4093	7.2048
ln(implied volatility)	0.7803	0.3840	2.1525	–0.4629	–0.0830	3.2454
The first differences of implied volatility	–1.5893 e–5	0.4711	3.1464	–3.4740	–0.3253	12.2273
The first differences of ln(implied volatility)	–7.9812 e–6	0.1956	1.0708	–0.8356	–0.1403	6.0058

Table 3.

The Results of the long memory dependence tests

Data	Lo	GPH
The first differences of implied volatility	0.8496 (<0.975)	–0.9015 (<0.2)
The first differences of ln(implied volatility)	0.7826 (<0.995)	–0.2278 (<0.425)

The Results of the long memory test of the first differences of the implied volatility series and the first differences of the logarithm of the implied volatility series are presented in Table 3. We have used the methods from previous section. We can't reject the null hypothesis of the Lo test for the first differences of implied volatility series as well as for the first differences of logarithm of the implied volatility series. We fail to reject the null hypothesis of the GPH test H_0 for both the time series.

6. EMPIRICAL RESULTS

We decided to use 645 elements of our data since September 22, 2003 to April 7, 2006 for model estimation. The next 80 data from April 10, 2006 to August 4, 2006 have been used for forecasting.

At first we tested the conditional heteroskedasticity of our series. We applied the Engle test. The results are presented in Table 4.

Table 4.

Results of conditional heteroskedasticity test

Series	Engle
Implied Volatility	36.4674 (0)
The Logarithm of implied Volatility	29.0194 (0.001)
The First difference of Implied Volatility	43.841 (0)
The First difference of Logarithm of Implied Volatility	31.2162 (0.001)

We observe that the Engle test rejects conditional homoskedasticity for all the series very strongly. Therefore we assume that all the series have varying conditional variance. Hence we model and forecast them with ARFIMA(p,d,q)-GARCH(r,s) model.

In the case of non-stationarity of the series, the stationarity of the process of the first differences can be defined from the following facts:

The fractional integration parameter d in some ARFIMA *model can be increased by* 1 *swapping* y_t *for the process of the first differences* $(y_t - y_{t-1})$.

We perform the estimation on the pre-differenced series, but the parametr estimates, plots and forecasts are all generated for the original series. The best fitted model for our series is ARFIMA(0,d,1)-GARCH(1,1). Our estimation results are presented in Table 5.

Table 5.

The parameters of ARFIMA(0,d,1)-GARCH(1,1) models for implied volatility and the logarithm of implied volatility. Estimation on pre-differenced series

Parameter	Implied volatility			Logarithm of implied volatility		
	Estimate	Std. Err.	p-Value	Estimate	Std. Err.	p-Value
Student's t d.f.	3.86089	0.6965	------	5.16993	1.0485	------
d	–0.18385	0.09377	0.05	–0.16691	0.07751	0.032
θ_1	0.352	0.12322	0.004	0.36373	0.12086	0.003
ω	0.00816	0.0069	------	0.01721	0.0031	------
$7\alpha_1$	0.1863	0.08613	0.031	0.28624	0.08526	0.001
β_1	0.79308	0.09455	0	0.1706	0.07936	0.032

The estimates of parameter d suggest that the both time series are stationary and invertible. It confirms the results of the KPPS test. These values suggest also that the processes have the so called average memory. Boutahar (2005) showed that an ARFIMA model for non-stationary series can give forecasts as good as the ones from an ARFIMA model for the pre-differenced series. Therefore, we also apply ARFIMA to modeling and forecasting non-pre-differenced series. Our estimation results are presented in Table 6.

Table 6.

The estimated parameters of ARFIMA(0,*d*,1)-GARCH(1,1) models for implied volatility and logarithm of implied volatility

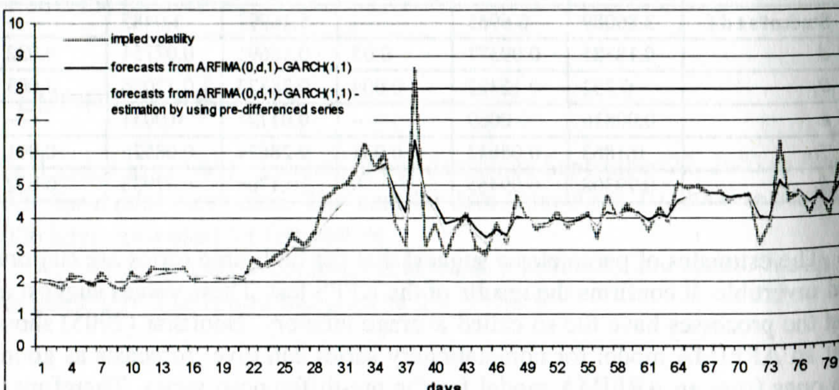
Parameter	Implied volatility			Logarithm of implied volatility		
	Estimate	Std. Err.	*p*-Value	Estimate	Std. Err.	*p*-Value
Student's t d.f.	3.54004	0.6264	------	5.06683	0.9962	------
γ	3.71358	0.56814	0	1.30205	0.26027	0
d	1.30854	0.17523	0	0.76166	0.06545	0
θ_1	0.84825	0.10579	0	0.29744	0.09681	0.002
ω	0.01122	0.0098	------	0.01685	0.0032	------
α_1	0.21153	0.09199	0.022	0.25718	0.0863	0.003
β_1	0.77068	0.10524	0	0.20544	0.08373	0.014

In both, implied volatility and logarithm of implied volatility series the hypothesis $d = 0$ is strongly rejected. The estimates of parameter d suggest that the implied volatility series and the logarithm of implied volatility series are invertible and not stationary. It also confirms results of the KPSS test.

The implied volatility forecasts obtained from ARFIMA(0,*d*,1)-GARCH(1,1), and ARFIMA(0,*d*,1)-GARCH(1,1) using estimation on the pre-differenced series are presented in Figure 2 together with the implied volatility calculated in the forecasting period.

Figure 2.

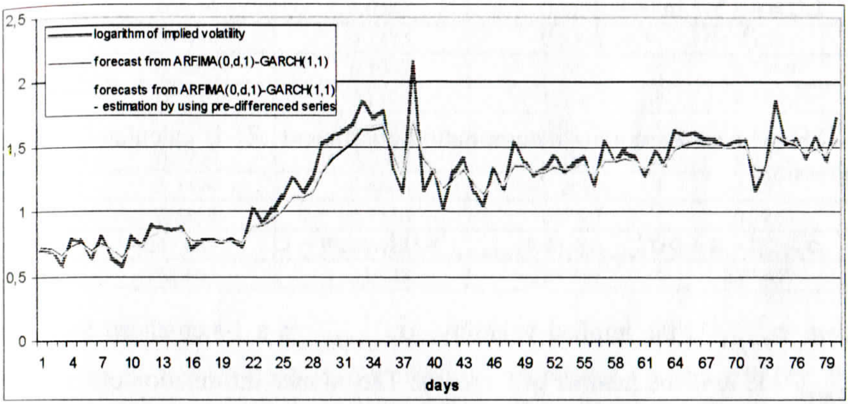
Implied volatility of the WIG20 index and the forecasts of implied volatility of the WIG20 obtained from ARFIMA(0,*d*,1)-GARCH(1,1), and ARFIMA(0,*d*,1)-GARCH(1,1) by using estimation on the pre-differenced series



The forecasts of logarithms of the implied volatility obtained from ARFIMA(0,d,1)-GARCH(1,1) and ARFIMA(0,d,1)-GARCH(1,1) using estimation on the pre-differenced series are presented in Figure 3 together with the logarithms of the implied volatility calculated in the forecasting period .

Figure 3.

Logarithms of the implied volatility of the WIG20 index and the forecasts of implied volatility of the WIG20 obtained from ARFIMA(0,*d*,1)-GARCH(1,1), and ARFIMA(0,*d*,1)-GARCH(1,1) by using estimation on the pre-differenced series



We observe that the variance of implied volatility is much smaller than the variance of realized volatility (Doman, 2006). Consequently, the forecasts from ARFIMA models can be better for implied volatility series than for realized volatility. From figures 2 and 3 we observe that ex-post 1-day forecasts have values similar to actual implied volatility in all cases.

In Table 5, MSE, MedSE, ME, MAE, RMSE and MAPE denote Mean Squared Error, Median Squared Error, Mean Error, Mean Absolute Error, Root Mean Squared Error, Mean Absolute Percentage Error, respectively. Adjusted Mean Absolute Percentage Error is denoted as AMAPE and expressed by the formula

$$\mathrm{AMAPE} = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{y_{T+i} - \hat{y}_{T+i}}{y_{T+i} + \hat{y}_{T+i}}\right|,$$

where N is the number of forecasts, Theil Inequality Coefficient is denoted by TIC and given by the formula

$$\mathrm{TIC}=\frac{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_{T+i}-\hat{y}_{T+i})^2}}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}y_{T+i}{}^2}+\sqrt{\frac{1}{N}\sum_{i=1}^{N}\hat{y}_{T+i}{}^2}}$$

and the Logarithmic Loss Function is denoted as LL and defined by the formula

$$\mathrm{LL}=\frac{1}{N}\sum_{i=1}^{N}\left(\ln\frac{\hat{y}_{T+i}}{y_{T+i}}\right)^2.$$

The Mincer-Zarnowitz determination coefficient R^2 is calculated from the regression

$$\sigma_{T+j+1}^2=a+b\hat{\sigma}_{T+j+1|T+j}^2+e_{T+j+1},\ j=0,1,\ldots,m-1,$$

where σ_{T+j+1}^2 is the implied volatility, $\hat{\sigma}_{T+j+1|T+j}$ is a 1-step-ahead forecast of σ_{T+j+1} and m is the number of forecasts. The Akaike information criterion, and the Schwarz information criterion are given by the formulas

$$\mathrm{AIC}=L-k \quad \text{and} \quad \mathrm{BIC}=L-\frac{1}{2}k\log(T),$$

where L is the maximum of the log-likelihood function, k is the number of fitted parameters, and T is a sample size. The larger is the value of the criterion, the better the model is fitted.

From Table 5 we observe that the errors are smaller for logarithms of the implied volatility. The most essential is the fact that also the percentage errors and the TIC coefficient have smaller value for the logarithmized series. Values of R^2 coefficient also confirm that the forecasts are better for logarithms of the implied volatility. Values of the forecast errors suggest also that forecast obtained from the non-stationary series have a similar quality.

Table 7.

The values of some forecast errors, R^2 coefficient and information criteria for 1 day ahead volatility forecasts from ARFIMA(0,*d*,1) for forecast of implied volatility and logarithms of implied volatility series

Error or criterion	ARFIMA(0,*d*,1)-GARCH(1,1) (Estimation on pre-differenced series)		ARFIMA(0,*d*,1)- GARCH(1,1)	
	Implied volatility	Logarithms of implied volatility	Implied volatility	Logarithms of implied volatility
MSE	0.8260	0.0387	0.8153	0.0384
MedSE	0.0915	0.0117	0.0971	0.0107
ME	0.0619	0.0381	0.1152	0.0337
MAE	0.5494	0.1375	0.5398	0.1361
RMSE	0.9088	0.1967	0.9030	0.1960
MAPE	0.1415	0.1105	0.1364	0.1098
AMAPE	0.0685	0.0558	0.0675	0.0553
TIC	0.1900	0.1328	0.1890	0.1324
LL	−5.1499	−5.4868	−5.2620	−5.4411
R^2	0.6865	0.7536	0.6727	0.7461
AIC	−247.019	254.422	−246.465	252.9
BIC	−262.661	238.78	−259.868	239.497

One-step ex-post forecasts give a possibility of testing a model stability. We use a test of stability which is the usual difference-of-means test on the residual and forecast error variances. The test statistic is given by the formula

$$S = \frac{F^{-1}\sum_{t=T+1}^{T+F}\hat{u}_t^2 - T^{-1}\sum_{t=1}^{T}\hat{u}_t^2}{\sqrt{F^{-1}\mathrm{Var}_F\left(\hat{u}_t^2\right) - T^{-1}\mathrm{Var}_T\left(\hat{u}_t^2\right)}},$$

where F is the number of forecasts, and T is the number of data used to model the fitting. The statistics is asymptotically $N(0,1)$ under the stability hypothesis, assuming that the fourth moments exist (Davidson,2006). The results are presented in Table 8.

Table 8.

The results of the test of model stability

	Estimation using ARFIMA(0,*d*,1)		Estimation using ARFIMA(0,*d*,1) (Estimation on pre-differenced series)	
	Implied volatility	Logarithms of implied volatility	Implied Volatility	Logarithms of implied volatility
S	2.0452 (0.041)	0.8938 (0.371)	2.1108 (0.035)	0.9508 (0.342)

From the results of the test follows that the stability hypothesis is rejected for both models used to forecasting the implied volatility, but we failed to reject this hypothesis for both models used to forecasting logarithms of the implied volatility.

7. CONCLUSION

We proposed modeling and forecasting the implied volatility and its logarithm for the stock index WIG20 using ARFIMA models. We showed that these time series are not stationary and they haven't conditional variance. So we applied estimation on the pre-differenced series. According to Boutahar (2005), we try to model non-stationary implied volatility and its logarithm using ARFIMA directly. From Figures 2 and 3 and the comparison of values of forecast errors we observe that the forecasts obtained with this method have the quality which is closed to one observed for estimation on the pre-differenced series. It confirms the proposition submitted by Boutahar (2005). Moreover, we observe also that logarithms of the implied volatility are better forecasted with ARFIMA than the implied volatility.

Analyzing our results, we notice that in spite of non-stationarity of the implied volatility series, it can be well forecasted with ARFIMA models. This means that the expectations of the future volatility can be well forecasted.

REFERENCES

Black F., Scholes, M. (1973), *The Pricing of Options and Corporate Liabilities,* Journal of Political Economy 81, 637–654.

Boutahar M. (2005), *Optimal Prediction with Nonstationary ARFIMA Model,* Groupement de Recherche en Economie Quantitative d'Aix-Marseille – UMR-CNRS 6579 Ecole des Hautes Etudes en Sciences Sociales Universités d'Aix-Marseille II et III.

Chiras D., Manaster, S. (1978), *The Information Content of Option Prices and a Test of Market Efficiency,* Journal of Financial Economics 6, 213–234.

Davidson J. (2006), *Time Series Modelling Version 4.20*, Cardiff University.

Doman M. (2006), *Modeling the Realized Volatility with ARFIMA and Unobserved Component Models: Results from the Polish Financial Market.* Forthcoming.

Geweke J., Porter-Hudak S. (1983), *The estimation and application of long-memory time series models,* Journal of Time Series Analysis 4, 221–237.

Granger C. W. J., Joyeux R. (1980), *An Introduction to Long-Memory Time Series Models and Fractional integration,* Journal of Time Series Analysis 1 (1), 15–29.

Hull J., White A. (1987), *The Pricing of Options on Assets with Stochastic Volatilities.* Journal of Finance 42, 281–300.

Hurst H. E. (1951), *Long term storage capacity of reservoirs*, Transactions of the American Society of Civil Engineers 116, 770–199.

Kwiatkowski D., Phillips, P. C. B., Schmidt, P., Shin Y. (1992), *Testing the null hypothesis of stationarity against the alternative of a unit root,* Journal of Econometrics 54, 159–178.

Lo A. (1991), *Long term memory in stock market prices*, Econometrica 59, 451–474.

Mandelbrot, B.B., Wallis, J.R. (1968), *Noah, Joseph and operational hydrology*, Water Resources Research 4, 909–918.

Metodology of Calculating the „Greeks" (in Polish). www.gpw.com.pl/zrodla/gpw/Greki.pdf

Schmalensee, R., Trippi, R. (1978), *Common stock volatility expectations implied by option premia,* The Journal of Finance 33, 129–147.

Macromodels
International Conference