COMPARISON BETWEEN IMPLIED AND HISTORICAL VOLATILITY FORECASTS: EVIDENCE FROM THE RUSSIAN STOCK MARKET

by

Denys Percheklii

A thesis submitted in partial fulfillment of
the requirements for the degree of

MA in Economic Analysis.

Kyiv School of Economics

2014

Thesis Supervisor: ______________________________ Olesia Verchenko

Approved by ____
Head of the KSE Defense Committee, Professor Irwin Collier

____

____

____

Date ______________________________

Kyiv School of Economics

Abstract

COMPARISON BETWEEN IMPLIED AND HISTORICAL VOLATILITY FORECASTS: EVIDENCE FROM THE RUSSIAN STOCK MARKET

by Denys Percheklii

Thesis Supervisor: Professor Olesia Verchenko

In this study we analyze which instrument provides a better volatility estimate on the Russian stock market: implied volatility or historical volatility. Using standard OLS regression we conclude that the results of early studies of developed markets can be extrapolated to the emerging markets like Russia. We find that implied volatility is an inefficient and biased predictor of realized volatility on the Russian stock market. Dividing out data set into three maturity buckets we found that historical volatility outperforms implied volatility in terms of predicting realized volatility for both call and put options and for all three groups of options. The analysis of three maturity buckets shows that for relatively longer-term options neither implied volatility nor historical volatility is useful for predicting realized volatility. This can be explained by two reasons. First, the longer-term options are traded more seldom. Second, when predicting on long horizon the errors of our prediction increase.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

The author wishes to his sincere gratitude to his thesis advisor, Professor Olesia Verchenko, for her irreplaceable guidance and encouragement during the process of working on this research, for giving him inspiration for the topic of his thesis and providing with relevant suggestions at every stage of my research.

The author wants to thank Owain Ap Gwilym and George Filis for sharing their articles with the author. These articles played very important role during this research.

The author is grateful to all KSE professors for their help, patient and inspiration over two years. Also the author wants to thank all his classmates for the incredible two years of his life.

Finally, the author wants to thank all members of his family for their faith, support and encouragement during two years at KSE. The author expresses personal thanks to his wife and his son for giving inspiration and stimulus for development.

GLOSSARY

IVC – implied volatility of call options.

HVC – historical volatility of call options.

RVC – realized volatility of call options.

IVP – implied volatility of put options.

HVP – historical volatility of put options.

RVP – realized volatility of put options.

B-S – Black-Scholes model.

C h a p t e r 1

INTRODUCTION

Over the last decades of the twentieth century we witnessed a rapid development of the financial markets of developed countries. The volume of trading rose significantly and various types of new complex tools were implemented on these markets. These new tools, generally called financial derivatives, are forwards, futures, swaps and options. Such instruments give to economic agents new means to hedge their portfolios and advance their income and utility. And such effects promote economic growth and raise living standards both in each separate country and in the world economy.

In 1973 the Chicago Board Options Exchange (CBOE) was founded, and, since then, options turned into one of the most actively traded type of financial derivatives. This is true not only for well-developed markets, but also for emerging ones. Loosely speaking, an option is a financial derivative, which gives the right but not the obligation to its holder to buy or sell an underlying asset (such as a stock, a bond, a commodity etc.) at a pre-specified price. Options generally require low initial investments and can generate large profits. That is why, option markets attract a lot of investors.

The theory developed by Markowitz (1952) and Sharp (1964) suggests that the main concern investors have while choosing their portfolio should be how return relates to risk. Risk can be measured as the volatility of an asset's price. An investor interested in the risk of an asset can get historical prices and calculate their standard deviation, which would provide a backward-looking estimate of volatility. However an investor is more interested in estimating the future volatility rather than the past one.

According to the classical option pricing theory (Black and Scholes, 1973) volatility is the main element that determines the fair price of an option or any other derivative. Accurate volatility forecast is not only important for investors, who use it, for example, for portfolio allocation and hedging decision, but also is needed for public policy decision. For example, how public expenditures influence consumption volatility (Herrera and Vincent (2008)) or what policy options can be used to decrease price volatility (The State Of Food Insecurity In The World (2011)).

The classical option pricing theory (Black and Scholes, 1973) provides the tool to estimate this future expected volatility given the observed option prices. This volatility is called the option implied volatility. But how good is this estimate? In particular, how tight is the relation between the expected volatility extracted from option prices and actual ex-post volatility calculated from the stock market?

To give better understanding of the volatility concept let's assume any point in time when all stock prices and option prices are observed for whole relevant time horizon. On Figure 1 this point is represented by point t. Then, as of this point in time, we can divide all information about prices into three blocks:

1. Historical stock prices. This information allows to calculate historical volatility using simple moving average or GARCH framework.
2. Future stock prices. This information allows to calculate realized volatility.
3. Actual option prices. This information allows to calculate implied volatility.

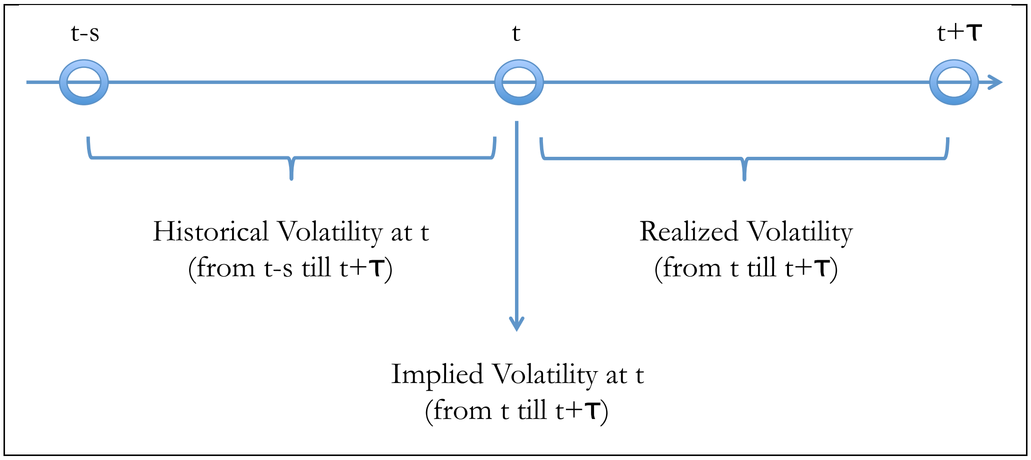

Figure 1. Visualization of data requirements to estimate volatilities

Talking about implied volatility, not much was done in this field referring to the emerging and developing markets. Most of studies analyze either liquid markets, like the US (Canina and Figlewski (1993), Christensen and Prabhala (1998), Ionesco (2011) etc.) and the U.K. (O.A. Gwilym and M. Buckle (1999)) or illiquid markets, but in developed countries, like Denmark (Hansen (2001)) or Australia (Li and Yang (2008)). However, the rapid development and growth of developing markets has attracted more and more investors, even though emerging markets are much more risky.

Furthermore, newly emerging and developing countries have been playing a significant role in the global economy lately. For example, experts think about China as the main driving force for the future economic growth, while Russia strengthens its influence in Europe by enlarging the export of commodity goods.

In the given situation, the Russian government actively supports the development of financial markets, including of the derivative market. However, in such volatile markets as the Russian market, investors are looking for ways of hedging against the change in prices of raw materials, and, consequently, stock prices. Options' trading is one of the possible solutions to the problem. The largest options and futures exchange board in Russia is the Moscow Exchange, which was founded in 2011 by merging the Moscow Interbank Currency Exchange and the Russian Trading System (RTS).

Why the Russian market is interesting to study? First, export of natural resources generates large cash flows inside the country and this potentially can lead to further development of derivatives market in Russia. Second, Russian financial market is poorly researched and, maybe, due to youth of Russian derivatives market there is research about comparison of implied and historical volatilities.

Previous studies in the field can be divided into three parts: developed liquid markets, developed illiquid markets and emerging markets.

Developed liquid markets were the first to be researched. First scientists collected daily data and concluded that implied volatility is a weak predictor of implied volatility (Canina and Figlewski (1993), Day and Lewis (1992)). However, later on analysis of monthly data showed that implied volatility is a better predictor of realized volatility than historical volatility Christensen and Prabhala (1998), Ionesco (2011), Shu and Zhang (2003))

Even thought we can see less trading on developed illiquid it is still possible to collect and analyze monthly data on these markets. This analysis showed that on the developed illiquid markets results for monthly data are similar to developed liquid markets (Hansen (2001) Li and Yang (2008)).

In its term, trading data on the emerging markets doesn't let us to form monthly data. While daily data analysis shows that implied volatility is a poor predictor of realized volatility (Filis (2009)). Even thought, these results are in line with results for developed illiquid markets, this field is quite new and needs further research to be sure that emerging markets have the same patterns as developed markets. Thus, our research adds valuable study of one of the biggest emerging market.

In this research we will estimate the volatility extracted from option prices, then estimate the volatility calculated from historical data and find out which one gives better forecast for future volatility.

The contribution of this work is to fill the gap in the literature considering Russian derivatives markets and to test whether information on this market can be used to forecast volatility.

The remainder of the paper gives information about relevant literature, data set, methodology, primary results and further steps of the research.

Chapter 2

LITERATURE REVIEW

In this literature review we first focus on the earlier results for liquid markets and then move on to more recent results for liquid and less liquid markets.

A lot of studies were undertaken on indexes of developed markets, such as S&P and FTSE. Early papers in the field support the idea that implied volatility gives better volatility prediction than historical volatility (Latane and Rendleman (1976), Beckers (1981)). However, due to data limitation they used static cross-sectional regression approach.

Using time-series data later works reached opposite conclusions. Using daily data several authors concluded that implied volatility is poor estimator of realized volatility.

Day and Lewis (1992) formed their data set from daily closing prices of S&P 100 index and closing prices and contract volumes of call options (OEX options). Data sample covered the period of time from March 11, 1983 to December 31, 1989. They found that GARCH and EGARCH models had similar results to implied volatility. But, neither implied volatility nor conditional volatilities from GARCH and EGARCH fully explain realized volatility of stock. And the final conclusion was that GARCH model gave better forecast than EGARCH.

Canina and Figlewski (1993) showed that implied volatility is a very poor estimator of realized volatility. They covered the data of the period from March 15, 1983 to March 28, 1987 for S&P 100 index. They only took the options, traded 7-127 days and which were not more than 20 points out- or in-the-money. All of them were American options on dividend-paying stocks mostly. Their main conclusion was that the one must implied and historical volatility and make one's decision not just relying on one single estimator.

In the same year's research by Lamoureux and Lastrapes (1993) showed less sharp results, but, still, the conclusion was similar. They took 10 individual stocks, traded on the Chicago Board Options Exchange (CBOE) and the options written on those stocks for the period from April 19, 1982 to March 31, 1984. The stocks didn't pay any dividends for the period. Their data consisted of the inside spreads of both options' and stocks' prices on each day of the sample. They compared GARCH approach and implied volatility approach and concluded that GARCH framework gave better prediction than the B-S option pricing model.

Another argument "against" the efficiency of implied volatility was provided by Fleming (1998). He also took the data from Chicago Options Exchange, he observed the options with at least 15 days to expiration, and, given than S&P 100 options expire monthly, which gave the range of 15 to 47 days to expiration days. The sample covered the period from October 1985 to April 1992, which gave 1664 daily observations. The results showed that IV gave biased forecasts.

Further studies try to explain negative results of the former ones. The main finding of later papers is that using monthly data implied volatility becomes significant estimator and shows more predictive power of realized volatility than historical volatility

In 1998 Christensen and Prabhala presented their results of testing the relationship between realized and implied volatility for S&P 100 index options. They suggested that weak predictability of implied volatility was due to highly overlapping periods used to estimate volatility using daily data. Their research had two main differences, comparing to the ones done before. First, they took monthly data, rather than daily data, which was common for all previous studies.

Second, the data sample consisted of the time period from November 1983 to May 1995, which is longer than in previous researches. Monthly data let them avoid overlapping data problem and have only one implied and one realized volatility in the sample for each month. With such differences they found that implied volatility gave not worse, but in some cases even better results than Historical Volatility.

Another article, which also examined liquid market (UK FTSE), was written by Gwilym and Buckle (1999). The main contribution of this study is that the authors examine relative accuracy of several different time horizons. The conclusion is that the best forecasting method (either implied or historical volatility) depends on the time horizon and data frequency. The data consists of the American-style FTSE index options at LIFFE that expire on a monthly basis. With the exception of June and December, all contracts trade for four calendar months. They examined five forecast horizons: 5–20 trading days, 21–40 trading days, 41–60 trading days, 61–80 trading days and more than 80 trading days. They ignored horizons of less than 5 trading days because of possible distortions in option markets due to approaching maturity.

More recent studies also prove monthly data approach leads to implied volatility as significant estimator.

Szakmary et. al. (2003) studies thirty five futures option markets. They found that for most of commodities IV is a good estimator of realized volatility. Also, in their study implied volatility outperformed historical volatility. They took the data from eight separate exchanges, but among indexes only S&P 500 was included. Overall their data set covered the period from 1984 to 2001. For S&P 500 index they used quarterly data (from January 1983 to February 2001). They implemented ADF test to their results with a null hypothesis of a unit root. Both implied and historical volatility rejected the null hypothesis for most of the markets, though historical volatility showed weaker results and more potential of non-stationarity.

Shu and Zhang (2003) tested the relationship between implied and realized volatility. They took options' closing prices of S&P 500, one on a month. They choose S&P 500, but not S&P 100, because S&P 500 index options were European style. The data set included the period from January 1, 1995 to December 9, 1999. Those options expire on the third Friday of each month. They compute implied volatility using B-S model and Heston stochastic volatility option-pricing model. They found that both B-S and Heston models outperformed historical volatility. They proved this by using simple regression model.

One of the latest results was provided by Ionesco (2011). He used time period from January 2004 to December 2010 for the S&P 500 and the DAX, and for the FTSE 100 from November 2004 to December 2010. He took the monthly data of closing prices of call and put options traded on Wednesday right after the last trading day of each month or, if Wednesday was not a business day he took the next closest trading day. His results showed that implied volatility was as good estimator as GJR-GARCH model, or, sometimes, even better.

All previous papers considered liquid markets in developed countries. But, according to the analysis of Danish (Hansen (2001)) and Australian (Li and Yang (2008)) markets that the results are similar for liquid and illiquid developed markets.

Hansen (2001) tested the relationship between implied and realized volatility on the Danish option and equity market. Danish monthly-expired options on KFX index were first introduced in August 1995. Those options were traded infrequently and in low volumes compared to S&P and FTSE. However, she showed that the results of Cristensen and Prabhala (1998) and Gwilym and Buckle (1999) could be implemented not only to liquid option markets. Her data set contained monthly call and put options of KFX index, traded on the Copenhagen Stock Exchange in the period from September 1995 to December 1999. There were some missing variables because in some months some specific types of options were not traded. She showed that implied volatility was a very good and unbiased estimator even on the illiquid market.

Another research based on illiquid market was done by Li and Yang (2008). They considered the S&P/ASX 200 index options traded on the Australian Stock Exchange. The options traded in Australia are European style with a quarterly expiry cycle: March, June, September and December. Data set covered the time period from April 2, 2001 to March 16, 2006. The authors stressed that they used similar methodology to the one used by Christensen and Prabhala (1998) and to avoid the overlapping of the data option "traded on a business day close to but after an expiry date, and have expiration on the next expiry date". Thus, given the specific of an exchange, the authors used quarterly data, which gives 20 observations. They found out that implied volatility of both call and put options is the better estimator of realized volatility. However, implied volatility of a call option has no relation to realized volatility, while implied volatility of a put option can be used to forecast future volatility.

Even thought developed markets are well researched, not much was done considering emerging markets. One of the few articles was written by Filis (2009). He considered Greek derivative market from January 2000 to January 2003, when Greece was emerging market. He used daily data and computed implied volatility for call and put options. This approach let him to collect 749 observations of both implied and realized volatilities. He used standard OLS regression and concluded that implied volatility is inefficient predictor of realized volatility.

As we can see, there is a gap in a literature considering emerging markets. It is still not clear whether we can claim that the results are similar for developed and emerging markets. And our article will add evidence whether we can claim it not. Anyway, further researches are needed considering emerging markets,

The standard methodology used in most of the studies methodology used by is well accepted, that is why we will be using the same methodology to calculate implied, realized and historical volatilities. It might be interesting to see how different frequency influence. But in my case we have not so much data and this is not feasible at this point.

C h a p t e r 3

DATA AND METHODOLOGY

The goal of the thesis is to test whether index volatility forecasts contained in option prices are better predictors of actual index volatility than historical volatility values.

To calculate implied volatility we use two data sets: MICEX index futures prices and futures-style options prices (for both puts and calls) for the period from the foundation of Moscow Exchange in 2011 to the beginning of 2014. MICEX index is "a capital-weighted price index of the 50 major and most liquid Russian stocks traded at the MICEX Stock Exchange (MICEX SE), calculated in real time (dividend income is excluded from the index calculation)". Notice that excluding dividends is not an issue for us since the constituent stocks did not pay dividends during the period of interest. MICEX Index Futures expire every 3 months and are actively traded during the entire period since launch of the Moscow Exchange.

Our data covers the period from September, 30 2011 to March, 7 2014. During this period 2 to 10 futures contracts were traded simultaneously. At the same time number of options with different strikes written on these futures contracts were traded. These are the options of our interest.

Options were introduced together with the futures, but did not have high demand at the very beginning. The situation changed in 2012, when options trading volumes rose significantly. But in September 2013 the demand fell again and was close to zero. The pattern is the same for the period after September 2013, trading was very rare.

Among the array of options with different terms of maturities and strikes available on any particular day, we choose option contracts written on the shortest-term futures and which are the closest to being at-the-money. We consider at-the-money options because most of trading contracts involve at-the-money options rather than other types. Moreover, even if B-S model is wrong at-the-money options still give unbiased estimate (Knight and Satchell (1997)).

Table 1 shows the life cycle of options contracts. The numbers in shaded columns correspond to the first trading day (the top number) and the maturity day (the bottom number) of different contracts. For example, the very last shaded column represents options, which were traded since November 12, 2012 and matured on December 16, 2013.

We use the methodology proposed by Canina and Figlewski's (1993) : we record the last trade prices of the futures and options with the shortest maturity, but not less than 7 calendar days. At-the-money option is defined as the one whose strike is the closest to this day futures' contract last trade price. Strike price of an option is the fixed price at which the holder can buy or sell the underlying asset or commodity. Our data set consists of the futures-style options, which can be explained as futures contracts on the payoff of an option.

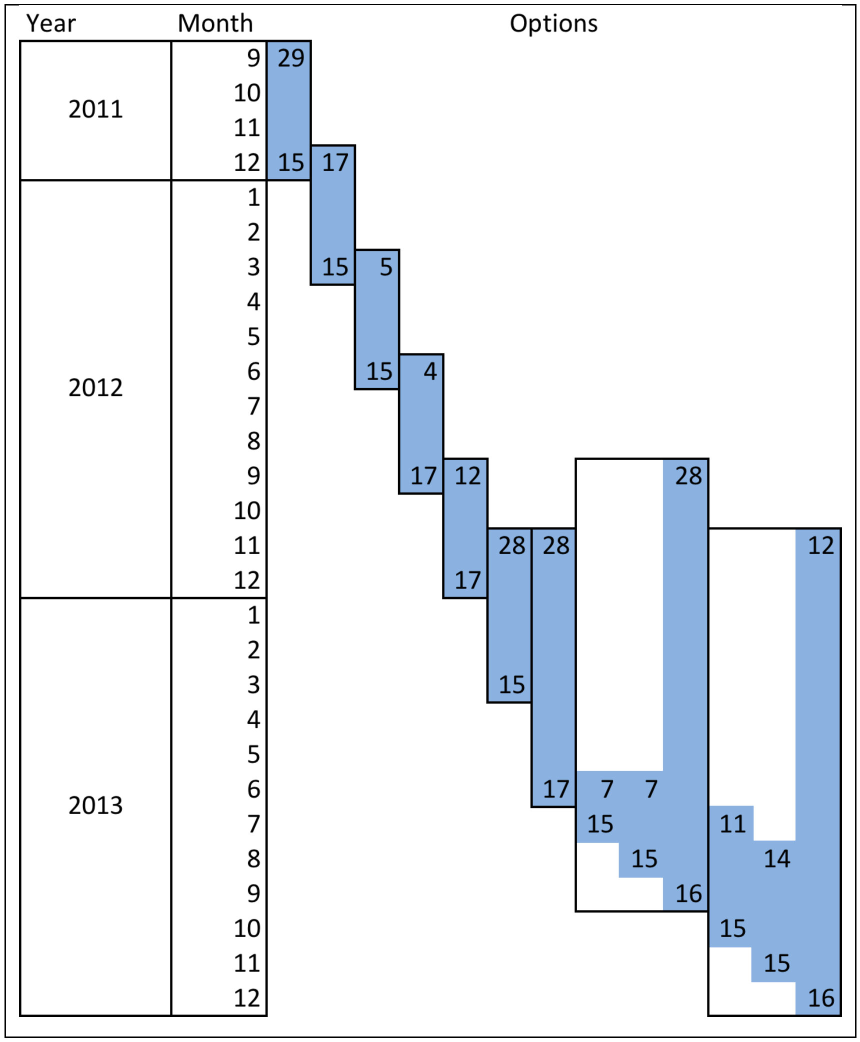

Figure 2. Options Maturities

We will used standard Black-Sholes call and put option pricing formulas for futures-style option provided by Hull (2012) in his classic textbook[1] and the Newton-Raphson algorithm to estimate their implied volatilities. Thus, futures-style call option price is

$$C = F_0 N(d_1) - KN(d_2) \tag{1}$$

[1] see p. 373

and futures-style put option price is

$$P = KN(-d_2) - F_0 N(-d_1) \quad (2)$$

where

$$d_1 = \frac{\ln(F_0/K) + 0{,}5\sigma^2 T}{\sigma\sqrt{T}} \quad (3)$$

$$d_2 = d_1 - \sigma\sqrt{T} \quad (4)$$

where

F_0 – futures price,
K – strike of an option,
$N(*)$ – cumulative normal density function,
T – years till maturity (trading days to maturity divided by number of trading days in a year, which is approximately equal to 249 in Russia),
σ – annualized expected volatility of an index over its remaining lifetime (from now till the maturity).

Notice that in these formulas all variables are observable except for sigma. Sigma is the expected volatility of an index over the time till maturity of the option. The assumption of Black-Sholes (BS) formula is that this volatility is a fixed parameter. We invert the BS formula and estimate the parameter sigma.

Following the studies, which analyze illiquid markets (Hansen, 2001; Li and Yang, 2008), historical and realized volatilities will be calculated as a standard deviation of index returns. We will use index returns instead of futures on index returns because according to Hull (2012) "the volatility of the futures price is the same as the volatility of the underlying asset".[2] Let's define logarithmic index returns as:

$$R_t \equiv \ln(S_t / S_{t-1}) \tag{5}$$

Then, realized volatility of the index prices at any time *t* is

$$\sigma_t = \sqrt{\frac{249}{\mathrm{T}} * \sum_{t+1}^{t+T} (R_t)^2} \tag{6}$$

where T is the number of trading days till maturity.

We calculate historical volatility of index returns similarly:

$$\sigma_{H,t-1} = \sqrt{\frac{249}{35} * \sum_{t-35}^{\mathrm{t}} (R_t)^2} \tag{7}$$

Following Canina and Figlewski (1993) we used 35 days as suggested forecasting horizon.

The final data pool consists of two subsets: for call and for put options. Due to low market liquidity we cannot calculate values for implied volatility for every day, but only for those days when trade actually occurred. As a result, for call options we have 301 values for each volatility (implied, realized and historical) and for put options – 241. There are more observations for calls since they are more actively traded

We will be using regression analysis at the end to find out which estimator, implied or historical volatility, gives better prediction of realized volatility.

Regression 1: $RV = \alpha + \beta * IV + u$

---

[2] see p. 371

Regression 2: $RV = \alpha + \beta * HV + u$

Regression 3: $RV = \alpha + \beta_1 * HV + \beta_2 * IV + u$

Implied volatility can be interpreted as the level of volatility that market agents expect when they set option prices. Thus, several hypotheses are of our interest for Regression 1. First, α=0 and β=1 imply that investors do not overestimate or underestimate implied volatility all the time. Second, the joint hypotheses that α=0 and β=1 which can be used to judge agents' rationality.

Following the literature, for Regression 2 we will test the same hypotheses. However, it is important to notice that there are no reasons to treat historical volatility in the same way as implied volatility. In particular, historical volatility does not have to be an unbiased predictor of actual volatility. Therefore, these tests are done with the purpose of treating our two alternative predictors similarly, and no conclusions about agents' rationality can be drawn in the case of historical volatility.

For Regression 3 we test four hypotheses. First, α=0. Second, β_1=0. Third, β_2=0. Fourth, joint hypotheses that α=0, β_1=0 and β_2=1 which again can be used to judge agents' rationality.

We are mainly interested in the Regression 3, but we run Regression 1 and Regression 2 to compare with previous literature results.

C h a p t e r   4

DESCRIPTIVE STATISTICS

We start with graphic analysis of the data. According to the eye bowl test Figure 3 and Figure 3 show that implied volatility is more volatile than realized and historical volatilities for both options.

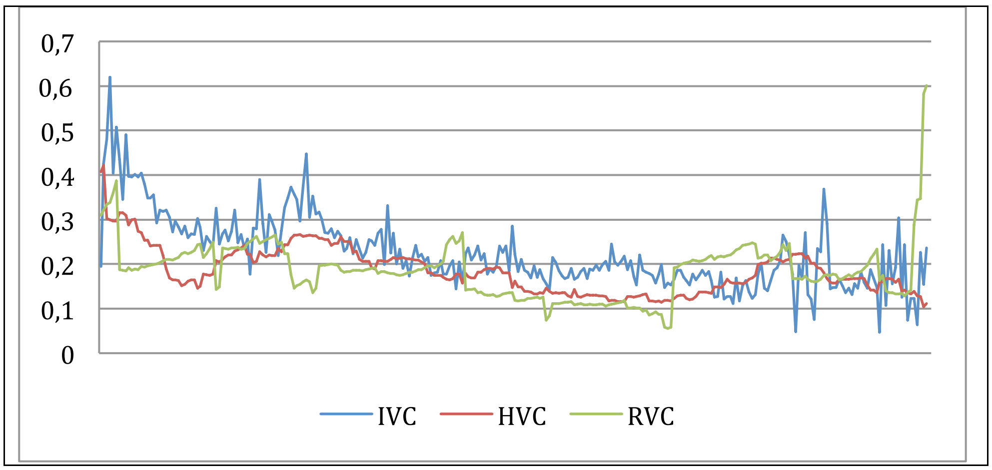

Figure 3. Call Options Volatilities

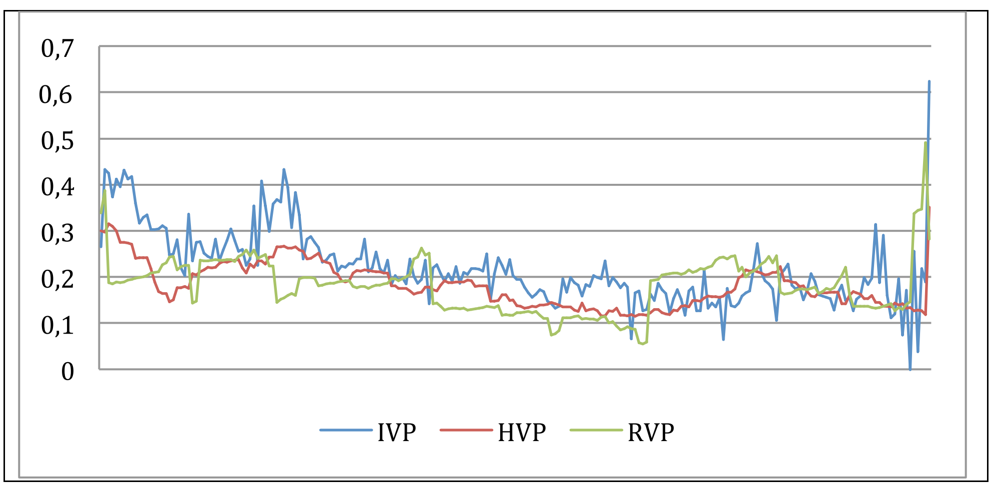

Figure 4. Put Options Volatilities

Table 1 provides further description of our option data sample. Maturity describes working days to maturity of an option. For both call and put it varies from 4 to 66 with average around 30. Moneyness statistics proves that when constructing our data set we observed only closest to being at the money options, meaning that the option's strike price it closest to the price of the underlying futures contract. Moneyness is calculated as a fraction of futures contract price divided by the option's strike. This fraction is very close to one with a small deviation.

Table 1. Implied volatility raw data description

	Maturity (working days)			S/K (moneyness)		
	Min	Average	Max	Min	Average	Max
Calls	4	31	66	0,9823	0,9979	1,0182
Puts	5	29	66	0,9825	0,9997	1,0182

Table 2 shows summary statistics of all volatility estimates. Here *IVC* stands for implied volatility for call options, *HVC* is historical volatility for call options, *RVC* is realized volatility for call options, *IVP* is implied volatility for put options, *HVP* is historical volatility for put options, *RVP* is realized volatility for put options.

We see that implied volatility data has the largest range of values, and the biggest standard deviation and the largest average values. Next comes historical volatility. This is in contrast to the results of Ionesco (2011) and Christensen and Prabhala (1998). In their subsamples IV has the same or smaller standard deviation than RV and HV. But, Filis (2009) obtained similar results to ours for the Greek market.

According to skewness, kurtosis statistics all variables show non-normality. Only *HVP* shows closest to normal distribution.[3]

As a consequence of non-normality we follow methodology of Filis (2009) and run Wilcoxon Signed Rank test to check whether the median of implied volatilities (IVC and IVP) are significantly different from the median of realised volatility (RVC, RVP). The results are shown in Table 3. As we can observe there is a significant difference between the implied volatilities and realised volatilities for both call and put options. Such results can be caused by two reasons. First, implied volatility is not an efficient predictor of realized volatility. Second, Russian market is not efficient.

Table 2. Summary statistics

	IVC	*HVC*	*RVC*	*IVP*	*HVP*	*RVP*
Mean	0.219	0.186	0.187	0.216	0.183	0.180
Maximum	0.619	0.421	0.601	0.624	0.351	0.491
Minimum	0.034	0.104	0.055	0.038	0.115	0.055
Std. Dev.	0.080	0.050	0.065	0.077	0.045	0.570
Skewness	1.198	0.930	1.831	1.366	0.673	0.963
Kurtoisis	5.686	4.656	12.027	6.396	3.151	6.644
Observations	301	301	301	241	241	241

[3] Skewness:
> 0 - Right skewed distribution - most values are concentrated on left of the mean, with extreme values to the right.
< 0 - Left skewed distribution - most values are concentrated on the right of the mean, with extreme values to the left.
= 0 - mean = median, the distribution is symmetrical around the mean.
Kurtosis:
> 3 - Leptokurtic distribution, sharper than a normal distribution, with values concentrated around the mean and thicker tails. This means high probability for extreme values.
< 3 - Platykurtic distribution, flatter than a normal distribution with a wider peak. The probability for extreme values is less than for a normal distribution, and the values are wider spread around the mean.
= 3 - Mesokurtic distribution - normal distribution for example.

Table 3. Wilcoxon Signed Rank test results

H0	Probability	H0	Probability
IVC=RVC	0.000[4]	IVP=RVC	0.000

Table 4 shows that in our data set realized volatility is more correlated with historical volatility than with implied volatility.

Table 4. Correlation matrix

	IVC	*HVC*	*RVC*		*IVP*	*HVP*	*RVP*
IVC	1.000			*IVP*	1.000		
HVC	0.636	1.000		*HVP*	0.736	1.000	
RVC	0.202	0.245	1.000	*RVP*	0.232	0.352	1.000

In our study we use relatively short-term options, which are defined as options with maturities with less than 67 working days. However, even within this maturity group there might be considerable differences in the behavior of very short-term options (generally defined as options with maturities up to one month) and relatively longer-term ones. Therefore, we follow Canina and Figlewski (1993) and divide our data into three maturity buckets:

- 1-month (4-23 working days)
- 2-month (24-46 working days)
- 3-month (47-66 working days)

Table 5 shows summary statistics of our maturity buckets. We can see that most observations belong to the 1-month bucket. This is not surprising because usually the most short-term options are the most actively traded ones.

[4] We reject H0 on 99% confidence interval

Table 5. Number of contracts in each maturity bucket

	Maturity			
	1-month	2-month	3-month	Total
Calls	124	111	66	301
Puts	111	81	49	241

On Table 6 and Table 7 we can see correlation matrixes adjusted for maturity buckets. Overall, we see the same patterns as on the Table 4 with two exceptions. In case of 1-month call options and 3-month put options realized volatility is more correlated with implied volatility than with historical volatility. As a preliminary result we can conclude that for these three maturity buckets implied volatility shows more predictive power than historical volatility.

Table 6. Correlation matrix of call options' volatilities for each maturity bucket

	1-month			2-month			3-month		
	IVC	*HVC*	*RVC*	*IVC*	*HVC*	*RVC*	*IVC*	*HVC*	*RVC*
IVC	1.000			1.000			1.000		
HVC	0.664	1.000		0.528	1.000		0.767	1.000	
RVC	0.243	0.237	1.000	0.159	0.389	1.000	0.198	0.220	1.000

Table 7. Correlation matrix of put options' volatilities for each maturity bucket

	1-month			2-month			3-month		
	IVP	*HVP*	*RVP*	*IVP*	*HVP*	*RVP*	*IVP*	*HVP*	*RVP*
IVP	1.000			1.000			1.000		
HVP	0.707	1.000		0.772	1.000		0.792	1.000	
RVP	0.252	0.416	1.000	0.223	0.389	1.000	0.149	0.139	1.000

Summary statistics, adjusted for maturity buckets is shown on the Table 8, Table 9 and Table 10. The results are similar to the Table 2. For all maturity buckets implied volatility has larges standard deviation for all maturity buckets. And, also, all variables show non-normality. However, distribution of the majority of variables is closer to normal than on Table 2.

After dividing data into three buckets we also run Wilcoxon Signed Rank test for them. Results are presented on Table 11. Even though the level of rejection droped a bit for 1-month bucket and for call options in 2-month bucket, we still reject the null hypotheses that the median of implied volatility equalt to the median of realized volatility.

Table 8. Summary statistics for 1-month maturity bucket

	IVC	*HVC*	*RVC*	*IVP*	*HVP*	*RVP*
Mean	0.213	0.176	0.190	0.206	0.174	0.180
Maximum	0.619	0.302	0.601	0.624	0.351	0.491
Minimum	0.047	0.104	0.055	0.038	0.115	0.055
Std. Dev.	0.089	0.046	0.084	0.086	0.045	0.071
Skewness	1.524	0.875	1.871	1.703	1.247	1.038
Kurtoisis	6.646	3.151	9.793	7.561	4.708	5.600

Table 9. Summary statistics for 2-month maturity bucket

	IVC	*HVC*	*RVC*	*IVP*	*HVP*	*RVP*
Mean	0.215	0.184	0.187	0.219	0.183	0.174
Maximum	0.426	0.422	0.344	0.432	0.274	0.259
Minimum	0.035	0.116	0.106	0.111	0.116	0.106
Std. Dev.	0.073	0.051	0.052	0.069	0.039	0.039
Skewness	0.689	1.647	0.404	1.246	0.415	0.019
Kurtoisis	3.429	8.664	3.053	4.459	2.655	2.184

Table 10. Summary statistics for 3-month maturity bucket

	IVC	*HVC*	*RVC*	*IVP*	*HVP*	*RVP*
Mean	0.238	0.209	0.180	0.232	0.205	0.189
Maximum	0.491	0.315	0.298	0.425	0.315	0.299
Minimum	0.074	0.119	0.111	0.064	0.119	0.111
Std. Dev.	0.072	0.052	0.039	0.067	0.051	0.042
Skewness	1.216	-0.073	0.504	0.644	-0.196	0.284
Kurtoisis	5.149	2.436	4.030	4.405	2.474	3.319

Table 11. Wilcoxon Signed Rank test results adjusted for maturity buckets

Bucket	H0	Probability	H0	Probability
1-month	IVC=RVC	0.0035	IVP=RVP	0.0025
2-month	IVC=RVC	0.0002	IVP=RVP	0.0000
3-month	IVC=RVC	0.0000	IVP=RVP	0.0000

Chapter 5

EMPIRICAL RESULTS

The results of our primary regressions are given in Table 12 and Table 13. These regressions' data set combines observations from all three maturity buckets. The first hypothesis that we test is whether $\alpha=0$ and $\beta=0$. If implied volatility is useful in predicting realized volatility, then its coefficient must be significantly different from zero. In our model this holds for both call options ($\beta=0.164$, se=0.05) and put options ($\beta=0.151$, se=0.01), but only in the specification where implied volatility is the only dependent variable (Regression 1). This suggests that implied volatility individually has some explanatory power for actual volatility. We also find that the intercept coefficient α is significantly different from zero for both call ($\alpha=0.151$, se=0.01) and put ($\alpha=0.143$, se=0.01) options.

Recall that implied volatility can be interpreted as the expected level of volatility that market agents bear in mind when they set option prices. Therefore, if agents are rational they should not consistently underestimate or overestimate this variable. This suggests that the joint hypothesis that $\alpha=0$ and $\beta=1$ can be used to judge agents' rationality. As can be seen from the next to last column in Tables 5 and 6, this hypothesis is easily rejected at any reasonable confidence level.

Regression 2 contains historical volatility as the only explanatory variable for actual (realized) volatility. This variable, as well as the intercept, are significant in both option samples. Therefore, historical volatility also has some explanatory power for actual volatility.

Next, we test the hypothesis that β1=1 for Regression 1 and Regression 2. We reject this hypothesis for all regressions and for all option contacts at 99% confidence level (see the last column of Tables 12 and Table 13).

Next, we include both implied and historical volatilities in one regression (Regression 3). Our goal is to check whether implied volatility contributes to predicting realized volatility once historical volatility is accounted for. There are several conclusions that we can make. First, the coefficient of implied volatility is insignificant for both call (β=0.063, se=0.06) and put (β=-0.04, se=0.06) options. Second, the coefficient of historical volatility remains significantly different from zero for both call (β=0.25, se=0.09) and put (β=0.488, se=0.11) options. Third, α remains significantly different from zero for both call (β=0.126, se=0.01) and put (β=0.099, se=0.01) options. Fourth, F-statistics null hypothesis (α=0, β_1=0 and β_2=1) is rejected on 99% confidence level.

As a primary result we can say that historical volatility is a better estimator of realized volatility than implied volatility.

Table 12. Regression results for Call Options

	Const (se)[5]	IVC (se)	HVC (se)	R2	F-stat (P>F)[6]	t-stat (P>|t|)
Reg 1	0.151*** (0.01)	0.164*** (0.05)		4.1%	12.73 (0.000)	-18.23 (0.000)
Reg 2	0.128*** (0.01)		0.314*** (0.07)	6.0%	19.02 (0.000)	-9.52 (0.000)
Reg 3	0.126*** (0.01)	0.063 (0.06)	0.250** (0.09)	6.3%	10.09 (0.000)	

[5] Standard error

[6] Probability

Table 13. Regression results for Put Options

	Const (se)	IVP (se)	HVP (se)	R2	F-stat (P>F)	t-stat (P>|t|)
Reg 1	0.143*** (0.01)	0.170*** (0.05)		5.5%	13.55 (0.000)	-17.96 (0.000)
Reg 2	0.100*** (0.01)		0.438*** (0.07)	12.5%	33.85 (0.000)	-7.49 (0.000)
Reg 3	0.099*** (0.01)	-0.040 (0.06)	0.488*** (0.11)	12.7%	17.11 (0.000)	

Our next step is to run regressions taking into account the buckets division to account for possible heterogeneity is the behavior of very short-term and relatively longer-term options. The results are slightly different for different maturity buckets.

First, we consider call options. In the 1-month maturity bucket's (Table 14) Regression 1's and Regression 2's betas remain significantly different from zero, even though the level of significance dropped from 99% to 95%. In Regression 3 both implied and historical volatility are insignificant. In the 2-month maturity bucket (Table 15) implied volatility is insignificant in both Regression 1 and Regression 3, while historical volatility is significantly different from zero in both Regression 2 and Regression 3. In 3-month maturity bucket (Table 16) implied and historical volatilities are insignificant in all Regressions. α coefficient is significantly different from zero for all maturity buckets.

Thus, for call options historical volatility remains better forecaster of historical volatility for 1-month and 2-month maturity buckets. In case of 3-month bucket neither implied nor historical volatility is an efficient predictor of realized volatility.

Table 14. Regression results for Call Options (1-month)

1-month	Const (se)	IVC (se)	HVC (se)	R2	F-stat (P>F)	t-stat (P>|t|)
Reg 1	0.142*** (0.02)	0.228** (0.08)		5.9%	7.66 (0.007)	-9.36 (0.000)
Reg 2	0.114*** (0.03)		0.430** (0.16)	5.6%	7.28 (0.008)	-3.58 (0.000)
Reg 3	0.116*** (0.03)	0.144 (0.11)	0.246 (0.21)	6.9%	4.51 (0.013)	

Table 15. Regression results for Call Options (2-month)

2-month	Const (se)	IVC (se)	HVC (se)	R2	F-stat (P>F)	t-stat (P>|t||)
Reg 1	0.162*** (0.02)	0.114 (0.07)		2.5%	2.82 (0.096)	-13.05 (0.000)
Reg 2	0.113*** (0.02)		0.400*** (0.09)	15.1%	19.38 (0.000)	-6.60 (0.000)
Reg 3	0.117*** (0.02)	-0.046 (0.07)	0.435*** (0.11)	15.4%	9.82 (0.000)	

Table 16. Regression results for Call Options (3-month)

3-month	Const (se)	IVC (se)	HVC (se)	R2	F-stat (P>F)	t-stat (P>|t||)
Reg 1	0.155*** (0.02)	0.107 (0.07)		3.9%	2.62 (0.110)	-13.57 (0.000)
Reg 2	0.146*** (0.02)		0.166 (0.09)	4.9%	3.27 (0.075)	-9.12 (0.000)
Reg 3	0.145*** (0.02)	0.038 (0.10)	0.125 (0.14)	5.1%	1.68 (0.194)	

Second, we consider put options. In the 1-month maturity bucket's (Table 17) Regression 1's and Regression 2's betas remain significantly different from zero, even though the level of significance in Regression 1 dropped from 99% to 95%. In Regression 3 implied volatility is significant, while historical volatility is insignificant. In 2-month maturity bucket (Table 18) implied volatility is insignificant in Regression 3, but remains significant in Regression 1 on 90% confidence interval. Historical volatility is significantly different from zero in both Regression 2 and Regression 3. In 3-month maturity bucket (Table 19) implied and historical volatilities are insignificant in all Regressions. α coefficient is significantly different from zero for all maturity buckets.

Overall, the results for put options are similar to those for call options: implied volatility has more predictive power in the case of 1-month and 2-month maturity buckets, while are insignificant in case of 3-month bucket.

Table 17. Regression results for Put Options (1-month)

1-month	Const (se)	IVC (se)	HVC (se)	R2	F-stat (P>F)	t-stat (P>|t|)
Reg 1	0.138*** (0.02)	0.205** (0.08)		6.5%	7.36 (0.008)	-10.40 (0.000)
Reg 2	0.066** (0.025)		0.656*** (0.137)	17.5%	22.85 (0.000)	-2.52 (0.013)
Reg 3	0.064* (0.02)	-0.064 (0.10)	0.744*** (0.19)	17.8%	7.73 (0.001)	

Table 18. Regression results for Put Options (2-month)

2-month	Const (se)	IVC (se)	HVC (se)	R2	F-stat (P>F)	t-stat (P>|t|)
Reg 1	0.145*** (0.01)	0.129* (0.06)		5.0%	4.14 (0.045)	-13.74 (0.000)
Reg 2	0.101*** (0.02)		0.398*** (0.11)	15.1%	14.04 (0.000)	-5.68 (0.000)
Reg 3	0.098*** (0.02)	-0.110 (0.09)	0.547** (0.17)	16.5%	7.73 (0.001)	

Table 19. Regression results for Put Options (3-month)

3-month	Const (se)	IVC (se)	HVC (se)	R2	F-stat (P>F)	t-stat (P>|t|)
Reg 1	0.168*** (0.02)	0.095 (0.09)		2.2%	1.07 (0.306)	-9.89 (0.000)
Reg 2	0.166*** (0.03)		0.116 (0.12)	1.9%	0.93 (0.339)	-7.36 (0.000)
Reg 3	0.165*** (0.03)	0.066 (0.15)	0.048 (0.20)	2.3%	0.55 (0.579)	

Taking into account all results we can conclude that in out data sample implied volatility is a weaker predictor of realized volatility than historical volatility, which is in line with the literature. The analysis of three maturity buckets shows that for relatively longer-term options neither implied volatility nor historical volatility is useful for predicting realized volatility. This can be explained by two reasons. First, the longer-term options are traded more seldom. Second, when predicting on long horizon the errors of our prediction increase.

As a final word we should mention that we also tested our regressions for serial correlation and heteroskedasticity. Using Durbin alternative test we found no serial correlation for all regressions in all maturity buckets. However, we see heteroskedasticity in almost all regressions, but robust regression show similar significance level in all regressions.

Chapter 6

CONCLUSIONS

Even thought in some cases implied volatility has some predictive power of realized volatility, overall we can conclude that implied volatility is not the best predictor of realized volatility on the Russian Stock Market. Historical volatility outperforms implied volatility in terms of predicting realized volatility for both call and put options. Robustness check of OLS regression gives similar results.

The analysis of three maturity buckets shows that for relatively longer-term options neither implied volatility nor historical volatility is useful for predicting realized volatility. This can be explained by two reasons. First, the longer-term options are traded more seldom. Second, when predicting on long horizon the errors of our prediction increase.

Our results support previous findings in the literature. However our results are important due to two reasons. First, we tested both call and put options' implied volatility and compared them to historical and realized volatility at the same time. Second, this is the first such study considering the Russian Derivative Market.

Biased and inefficiency of implied volatility can be either explained by some anomalies on the market or just data limitation. Data limitation is the more likely reason.

Further research is needed to test whether the changing of data frequency will lead to efficiency of implied volatility on the emerging market. However, there is no needed data to do this so far. At least, on the Russian Stock Market.

WORKS CITED

Black, Fischer and Myron Scholes. 1973. The valuation of options and corporate liabilities, *Journal of Political Economy*

Canina, Linda and Stephen Figlewski. 1993. The informational content of implied volatility, *Review of Financial Studies*

Christensen, B. J. and N. R Prabhala. 1998.: The relation between implied and realized volatility, *Journal of Financial Economics*

Day, Theodore. E. and Craig M. Lewis. 1992. Stock market volatility and the information content of stock index options, *Journal of Financial Economics*

Filis, George. 2009. An Analysis between Implied and Realised Volatility in the Greek Derivative Market, *Journal of Emerging Market Finance*

Fleming, Jeff. 1998. The quality of market volatility forecasts implied by S&P 100 index option prices, *Journal of Empirical Finance 5 (317-345)*

Gwilym, Owain Ap and Mike Buckle. 1999. Volatility forecasting in the framework of the option expiry circle, *The European Journal of Finance 5 (73-94)*

Hansen, Charlotte. S. 1999. The Relationship between implied and realized volatility in the Danish option and equity markets

Herrera, S. and Vincent, B: Public expenditure and consumption volatility, *Policy research working paper; 4633, (2008)*

Ionesco, Vladimir. 2011. The Performance of Implied Volatility in Forecasting future Volatility: an Analysis of Three Major Equity Indices from 2004 to 2010, *Thesis paper, MIT MA*

Lamoureux, Christopher G. and William D. Lastrapes. 1993. Forecasting stock return variance: understanding stochastic implied volatility, *Review of Financial Studies*

Li, Steven. and Qiangqiang Yang. 2009. The Relationship between implied and realized volatility: evidence from the Australian stock index option market, *Review of Quantitative Finance and Accounting*

Shu, Jinghong. And Zhang, Jin E. 2003 The Relationship Between Implied and Realized Volatility of S&P 500 Index, *Wilmott (83-91)*

Szakmary, Andrew, Evren Ors, Jin Kyoung Kim, Wallace N. Davidson III. 2003. The predictive power of implied volatility: evidence from 35 futures markets, *Journal of Banking and Finance*

Hull, John C. 2012. *Options, Futures and Other Derivatives. Eighth Edition.* London: Pearson Education

Knight, John N., Stephen E. Satchell 1997, Existence of Unbiased Estimators of the Black/Scholes Option Price, Other Derivatives, and Hedge Ratios, Econometric Theory