MASTER IN
FINANCE

MASTER'S FINAL ASSIGNMENT
DISSERTATION

IMPLIED VOLATILITY AS A FORECAST FOR FUTURE VOLATILITY: EVIDENCE FROM EUROPEAN MARKET

DIOGO FRANCISCO FERREIRA BELCHIOR

SEPTEMBER - 2012

Instituto Superior de Economia e Gestão
UNIVERSIDADE TÉCNICA DE LISBOA
DESDE 1911

MASTER IN FINANCE

MASTER'S FINAL ASSIGNMENT

DISSERTATION

IMPLIED VOLATILITY AS A FORECAST FOR FUTURE VOLATILITY: EVIDENCE FROM EUROPEAN MARKET

DIOGO FRANCISCO FERREIRA BELCHIOR

ORIENTATION:

RAQUEL M. GASPAR

SEPTEMBER - 2012

Implied Volatility as a forecast for Future Volatility: Evidence from European Market

Diogo Francisco Ferreira Belchior

Technical University of Lisbon

September, 2012

Abstract

The main purpose of this master thesis is test whether implied volatility is an accurate estimator for future volatility. We collect data regarding to the Euro Stoxx 50 index, namely closing index prices and implied volatility from one-month ATM options, in order to conduct an analysis of the European Market. The Sample selected covers the period from January 2002 to April 2012. The tests conducted allow us to conclude that implied volatility can be considered an unbiased and efficient estimator for future volatility and also that has more predictive ability than historical volatility, which is an indication of market efficiency.

Keywords: Implied volatility; Index options; Volatility forecasting; Market efficiency.

JEL Classification: C32; C36; G13; G14;

I am grateful to my supervisor, Raquel M. Gaspar for all advices and recommendations that helped me to conduct this study and all my collegues that share this experience by my side.

I want to dedicate this thesis to my family, especially my little sister to whom I try to be an example of dedication and professionalism.

Implied Volatility as a forecast for Realized Volatility: Evidence from European Market

Diogo Francisco Belchior

2012

Contents

List of Figures

List of Tables

1. Introduction

In the last decade, and as a result of the technological boom of the late '90s we have seen a proliferation and massification of financial markets. This type of investment has become more accessible and more appreciated by the general public. However the 2008 crisis, known as subprime crisis which originated in the complex derivatives market for financial assets, led to a sharp drop in assets prices and bulky losses by investors that still plague consumers confidence. This situation alarmed for the risks associated with investments in financial assets and in financial slang risk is directly associated with volatility, a key word in this study.

Volatility can be defined as the degree of oscillations, no matter the direction, in the return of a financial asset for a certain period and is usually defined as the standard deviation of (future) returns continuously compounded of a financial instrument over a given period of time. An asset with returns presenting large oscillations is considered volatile and consequently risky since the probability of a significant drop in its value is high.

Therefore, it is easy to see that before making a financial investment it is necessary to understand the characteristics of the instrument one is willing to invest in, especially its expected volatility. The problem is that this variable, unlike others, is not a concrete fact and is not available in the market. One solution is to resort to historical volatility, which is not more than measure the volatility of the instrument based on past returns. However, this technique involves looking back and there is no guarantee that the behavior of the asset, specifically its volatility, is constant in a future period. As such, it is not given much relevance to this indicator and traders focus their attention on the information available in the market to make their own estimate of the level of volatility that the asset will have in a future period. Thus arises a new kind of volatility, implied volatility.

Implied volatility is one that is incorporated in asset prices, namely in derivatives - usually financial options. The calculation of implied volatility requires a model such as Black and Scholes (1973) (B&S), where volatility is one of the model's parameters, and implied volatility is that which results when the model is solved in order to volatility, using the current market price of financial instrument. So, when a trader establishes a price to a financial option, implicitly sets its estimate to the value of the volatility of the underlying during the life time of the option. When that period ends, it is possible to calculate the volatility that actually occurred, which is at that moment considered historical volatility and compare with the forecast made at the beginning of period. If implied volatility is equal or next to that actually occurred one can say that the trader (investor) had a correct perception of the future regarding to the asset based on information available in the market on the date the option was quoted. On the other hand, if the difference between these two variables is significant one can conclude that the trader was not able to get a clear vision of the future or that there was some event that has affected the volatility of the asset.

Which of these situations is the most common? Can implied volatility be a good estimator for future volatility?

The purpose of this master thesis is to test the predictive ability of implied volatility in what concerns to estimate correctly the volatility over the past ten years and simultaneously make a comparison between historical volatility and implied volatility in order to understand if some of these variables is superior in terms of estimation. In addition an analysis is made dividing the horizon into two distinct periods to test whether market conditions affect the predictive ability of implied volatility over future volatility. During the last ten years the market environment was not constant. In 2002 the markets were facing a crisis, which had in 9/11 one of the origins. The situation lasted until roughly the middle of 2003, when consumers regained confidence in markets and these return to their natural course. This stability period lasted about five years until that in September 2008 the Lehman Brothers collapsed as consequence of sub-prime crisis, already stated above. The instability in the markets keeps up nowadays, being this time the main cause the financial instability of European countries sovereign debt. Summing up during the last ten years we had about 5 years which markets were stable and others 5 years more troubled.

There are several previous studies that have addressed this issue, trying to prove the efficiency of implied volatility as unbiased and efficient estimator for future volatility. However there is no consensus and the findings are divergent, eg. Canina and Figlewski (1993) and Jorion (1995). As such, beyond draw our own conclusions, is also made a comparison with earlier literature.

This study differs from others by:

- rely on more recent data

- illustrate the reality of the markets by focusing on a representative index of European market - The index chosen was the Euro Stoxx 50.
- split the dataset in two distinct periods
- split the analysis of calls and puts

The Euro Stoxx 50 Index is a free-float market capitalization-weighted index of 50 European blue-chip stocks from those countries participating in the Eurozone. Each component's weight is capped at 10% of the index's total free float market capitalization. The index was developed with a base value of 1000 as of December 31, 1991. The composition of the index is reviewed annually in September (www.Bloomberg.com). Financial options on Euro Stoxx 50 are called OESX options and are traded on Eurex that is one of the world's largest options exchanges and the pricing of financial instruments traded should be on the best level achievable with current methods.

We found that, in European context, implied volatility can be considered an unbiased and efficient estimator for future volatility. With all the information available on market, traders are able to do accurate volatility forecasts. Moreover, we proved that there are significant differences between the two periods analyzed and that implied volatility from put options has important information, being a better forecast than implied volatility from call options.

This study is organized as follows: Chapter 2 presents the literature review. Chapter 3 describes the dataset. Chapter 4 sets out the methodology and Chapter 5 presents the empirical results. A comparison between volatility implied by call and volatility implied by put options as estimator for future volatility is made in Chapter 6. Final conclusions are displayed in Chapter 7.

2. Literature Review

As stated in Chapter 1, to calculate implied volatility it is necessary to use an option pricing model, and the most popular in earlier literature and widely used by financial players is the B&S model. This is also the model used by Datastream to calculate the data used in our tests. This model, also known as Black-Scholes-Merton model was created in 1973 as a variant of the original model but taking into account the payment of dividends. The inputs of this model are the spot price, strike, interest rate, dividend yield, maturity and also volatility, the only variable not observable on market.

After the appearance of B&S formula some authors have addressed the question of implied volatility, trying to understand if this would be a good estimator of future volatility. This issue was highly discussed until the recent past and had its climax in the 90's.

The pioneers were Latane and Rendleman (1976) and Chiras and Manaster (1978). Using a cross-section analysis they were able to show that implied volatility is a more accurate estimator than historical volatility. Their studies were based on stock options traded on the Chicago Board Options Exchange (CBOE).

However, in the early 90s several studies raised doubts about the implied volatility efficiency as an estimator of future volatility, calling into question the conclusions obtained in earlier studies. Day and Lewis (1992) find out that implied volatility contains some information but is not a better forecast of future volatility than time series models. The study is done to the most active market of options, the options on the S&P100 index. Another study based on this market, and highly regarded, was conducted by Canina and Figlewski (1993). In their study they discovered that historical volatility is the best estimator for future volatility. Moreover, they argued that implied volatility has no correlation with future volatility and that implied volatility does not incorporate information contained in recent historical volatility. At a later stage, it is advocated by Christensen and Prabhala (1998) that the results obtained by Canina and Figlewski (1993) could be due to how the data was collected, since there is a large overlapping.

In the wake of two previous studies, Lamoureux and Lastrapes (1993) found that historical volatility contains more information than implied volatility, allowing to obtain more accurate estimates. Their analysis focused on individual stock options.

Duque (2000) conduct a study in what concerns to volatilities. They used several different methods to calculate implied volatility, historical volatility and also future volatility. Their analysis was based on a set of 30612 call options quotations written on 9 liquid stocks traded on the London Stock Exchange from April 1990 to December 1991. They found that the method used to calculate the variables has importance and more importantly that the best forecast for future volatility is achieved combining historical volatility and implied volatility. However, if it is necessary to choose only one variable to forecast future volatility, the best way to do it is using historical volatility because provides a more accurate estimation, confirming the results obtained one year earlier.

Jorion (1995) analyzed the market of currency options. This study was the first in a series of studies defending the predictive power of implied volatility. His analysis allowed him to conclude that implied volatility is an efficient estimator, although biased for future volatility.

Christensen and Prabhala (1998) also studied this issue. Like Day and Lewis (1992), its universe of study were options on the S&P100 Index, more specifically at-the-money one-month OEX call options. Their analysis covered the period from November 1983 until May 1995 and the data had monthly frequency. Unlike previous studies, the authors avoided overlapping data with the aim to reach more robust conclusions than those previously obtained. For the first time in a study of this nature, the authors were careful to examine possible measurement errors in the calculation of implied volatility. Using a Instrumental Variable (IV) technique they verified that there were in fact some sources of measurement errors. After correcting this fact through a two-stage least squares (2SLS), they found that implied volatility is in fact an unbiased and efficient forecast for future volatility, and a more accurate estimator for the future volatility than historical volatility. They also concluded that implied volatility incorporates all the information existing in historical volatility.

Two studies similar Christensen and Prabhala (1998) were conducted. The first was carried out by Hansen (2001). Its aim was to test whether implied volatility contains information about future volatility and hence whether the conclusions obtained by Christensen and Prabhala (1998) for the U.S. market could be extended to the Danish market in spite of low liquidity. The focus of the study were at-the-money European style call and put options with one month to maturity on KFX index, the representative index of Danish market, from September 1995 to December

1999. Using similar methodology to Christensen and Prabhala (1998), the author proved that the conclusions could actually be extended to the Danish market, since the results showed that implied volatility contains information about future volatility and moreover when compared to historical volatility, implied volatility provide a less biased forecast of future volatility. In fact, after correcting some measurement errors, using the IV technique the author could not reject the hypothesis that implied volatility is an unbiased estimator of future volatility. This study also made a comparison of volatility implied by call and volatility implied by put options. The conclusions were that both implied volatilities outweigh historical volatility but also that call implied volatility subsumes the information content in put implied volatility.

The second was made by Christensen and Hansen (2002). As Christensen and Prabhala (1998), the analysis was performed using options from the most active market's options, OEX options. The main difference is that instead of focusing in at-the-money options, the analysis is done using the implied volatility as an average of implied volatilities trade weighted from both in-the-money and out-of -the-money options and both calls and puts. Moreover, as Hansen (2001) they run a horse race between implied call, implied put and historical volatility. This study allowed to confirm the results obtained by Christensen and Prabhala (1998), namely that implied volatility is an unbiased and efficient estimator of future volatility and an estimator which is more accurate than historical volatility. In what concerns to the comparison of volatility implied by call and volatility implied by put, ordinary least squares (OLS) results indicate that call implied volatility is a better volatility forecast than put implied volatility, confirming what was discovered by Hansen (2001).

Finally, and most recently Pakarinen (2007) addressed this issue in his bachelor thesis. His analysis focused on 10 different call options on individual stock and covered the period from 2004 to 2006. Analyzed companies are listed either in Germany, Finland or in the U.S. This study supports the previous results by Canina and Figlewski (1993) and go against to those found by Christensen and Prabhala (1998). He found a positive relation between implied volatility and realized volatility, however the forecasting power of implied volatility does not outperform historical volatility.

When there seemed to be some consensus on this issue, this latest study raises doubts about the ability of implied volatility to forecast future volatility.

3. Data and sampling procedure

Data for this study span the period January 2002 through April 2012, and include closing index prices and option implied volatility for the Euro Stoxx 50. We begin with the study of the full period and then move on to a second analysis dividing the ten years in two sub-periods. The first sub-period runs from January 2002 to June 2003 and September 2008 to April 2012 and the second sub-period begins in June 2003 and ends in September 2008. The first sub-period represents the market on crisis situations while second sub-period represent a quiet period in financial markets. A similar approach is used by Christensen and Prabhala (1998), but referring to the stock market collapse in October 1987, analyzing the predictive power of implied volatility before and after this event.

3.1 Implied Volatility

Datastream produces a continuous series of implied volatility for at-the-money options with a constant time to maturity of 30 days. At-the-money prices are interpolated using two closest strikes available. This series is available for both calls and puts and for a number of equity index, including the Euro Stoxx 50.

The options on Euro Stoxx 50, traded on Eurex, are European style which means that only can be exercised at the maturity. These options expire, by convection, on the third Friday of each expiration month if this is an exchange day, otherwise the exchange day immediately preceding that day.

In order to avoid overlapping we move four days ahead from the expiration date, normally to the following Tuesday[1], and collect the implied volatilities reported by Datastream, which are calculated as a trade-weighted averages of the implied volatilities of the near term options, that is they have between 17 and 23 trading days to expiry[2]. The sampling procedure ensures that each option in data set expires before the next option is sampled. Beyond implied call and implied put volatility we construct a third time series that represents the average from these

[1]Monday was excluded because, as Harvey and Whaley (1992) refer, Mondays tends to be a day in which traders open positions for the week, excess buying pressure may result in higher volatility.

[2]Due to mid-week holidays, in our sample, there are eight observations with less than 17 days to expiration.

two series and was calculated using

$$I_t = \frac{1}{2} I_{c,t} + \frac{1}{2} I_{p,t} \tag{3.1}$$

where $I_{c,t}$ represents the volatility implied in call option, $I_{p,t}$ represents the implied volatility from put option and I_t denotes the average from both implied volatilities.

3.2 Realized Volatility

There are several ways to calculate realized volatility. Some models, as the Parkinson (1980) estimator or the Yang and Zhang (2000) estimator are more complex and use intra day information. Although these are more accurate we will use the close-to-close volatility estimator. So, realized volatility is computed as the sample standard deviation of the daily index returns over the remaining life time of the option. The formula is

$$h_t = \sqrt{\frac{1}{\tau_t - 1} \sum_{k=1}^{\tau_t} (r_{t,k} - \bar{r}_t)^2} \tag{3.2}$$

where τ_t is the number of days to expiration, $r_{t,k}$ represents the index return on day K that runs from the Tuesday following to the third Friday in month t to the third Friday in month $t+1$ and $\bar{r}_t$ is the mean of daily returns.

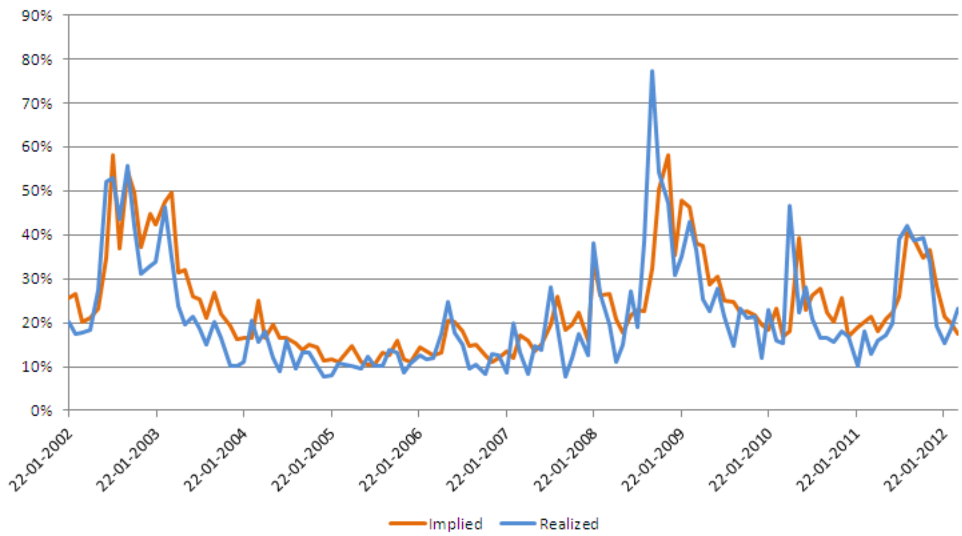
The index return is calculated through the formula $r_{t,k} = \ln(S_t/S_{t-1})$ where S_t is the index closing price on day t and S_{t-1} represents the index closing price on day $t-1$.

Finally, once the data relative to implied volatilities reported by Datastream are annualized, we should multiply by $\sqrt{252}$ to obtain the volatility in annual terms[3].

The main reason for use this method, instead others more accurate, e.g. Parkinson estimator is that it allow us to compare the results with previous studies that use this methodology to calculate realized volatility.

[3] 252 is the number of trading days considered in the analysis

Figure 3.1: Relation between implied and realized volatility on the full period



ATM implied volatility from one-month options and realized volatility from Euro Stoxx 50 index. Implied Volatility calculated as a simple average from implied call and implied put volatility as defined in Equation (3.1). Monthly data with 123 non-overlapping observations and covers the period from January 2002 to April 2012.

3.3 Descriptive statistics

Table 1 report the descriptive statistics, which originated in 123 observations for all time series:iImplied volatility (now on implied volatility represents the average from call and put implied volatility as defined in Equation (3.1)), realized volatility and their respective natural logarithmic transformations. Panel A present the entire period, while Panel B and Panel C represent both sub periods

There are two statistics that deserve further analysis which are the mean and standard deviation, because through these indicators we can understand the behavior of volatility and implied volatility over the last decade. Starting with the means and looking for the full period, we can see that both series have values slightly higher than 0.20 and implied volatility mean exceed realized volatility mean (0,2376 and 0,2144 for implied and realized volatility, respectively). Compared with Christensen and Prabhala (1998), which focused their analysis on $90'$s, its possible to conclude that financial markets have been more volatile in the beginning of this century, perhaps due to its growth and globalization. These results are similar to those found by Pakarinen (2007), where he examines several call options on stocks of European entities,

Table 3.1: Descriptive statistics: implied volatility and realized volatility

Statistic	Implied Vol.	Realized Vol.	Log Implied Vol.	Log Realized Vol.
Panel A: Full period - 01/2002 to 04/2012				
Mean	0,2376	0,2144	-1,5310	-1,6812
Std. Deviation	0,1106	0,1261	0,4250	0,5179
Skewness	1,2419	1,5799	0,4042	0,4599
Excess kurtosis	0,9894	2,7219	-0,5212	-0,4991
Jarque-Bera	36,6364	89,1376	4,7404	5,6118
Prob	0,0000	0,0000	0,0935	0,0605
Panel B: Subperiod 1 - 01/2002 to 06/2003 and 09/2008 to 04/2012				
Mean	0,3056	0,2858	-1,2487	-1,3550
Std. Deviation	0,1135	0,1366	0,3527	0,4511
Skewness	0,8114	1,1135	0,3669	0,2598
Excess kurtosis	-0,4000	1,1274	-1,0428	-0,7890
Jarque-Bera	7,0992	15,8363	4,1326	2,2688
Prob	0,0287	0,0004	0,1267	0,3216
Panel C: Subperiod 2 - 06/2003 to 08/2009				
Mean	0,1707	0,1442	-1,8089	-2,0022
Std. Deviation	0,0514	0,0580	0,2842	0,3537
Skewness	0,9637	1,6337	0,3987	0,6252
Excess kurtosis	0,5822	3,4038	-0,6339	-0,0190
Jarque-Bera	10,4722	57,5109	2,6802	4,0396
Prob	0,0053	0,0000	0,2618	0,1327

Descriptive statistic for implied volatility and realized volatility time series on Euro Stoxx 50 and also their natural logarithmic transformation. Implied Volatility is calculated as a simple average from implied call and implied put volatility as defined in Equation (3.1). Statistics are based on 123 monthly non-overlapping observations and covers the period from January 2002 to April 2012.

on the period from 2003 to 2006. However, this is not a concern because both series present the same movement, so that increase in the volatility values should not itself interfere with the predictive power of implied volatility.

Analyzing the standard deviations, and in good agreement with the notion that implied volatility should be an smooth expectation of realized volatility, we see that realized volatility has been more volatile than implied volatility, although the values are very close.

It is interesting to look at both periods separately because they have some significant differences. When markets goes through moments of crisis the values of implied and realized volatility are very high. On average during this period implied and realized volatility reach values close to 30%. On the other hand, during non-crisis period theses values drop significantly. As we can see, the realized volatility over the period 2003 to 2008 is only half of what occurred in moments of crisis. Moreover, the values of standard deviations are higher on crisis period, what means that, beyond present higher values, the volatility is more volatile (0,1366 in crisis period and only 0,0580 in non-crisis period for standard deviation of realized volatility).

Although original time series can not be considered normally distributed, we can see by Jarque-Bera statistic and also Probability that both log-transformed series are normality distributed. For this reason, from now on we will focus our analysis on the log-transformed series and let i_t denote the natural logarithm of the implied volatility and h_t denote the natural logarithm of the realized volatility. This allow us to have more confidence in the results because all realized tests fit better with normally distributed series and get a more accurate comparison with previous studies since the majority of these focused their analysis on the natural log-transformed series.

4. Methodology

To test whether implied volatility has any predictive ability over future volatility, we employ the following regression model used in Harvey and Whaley (1992), Canina and Figlewski (1993) and Christensen and Prabhala (1998).

$$h_t = \alpha_0 + \alpha_i i_t + \varepsilon_t, \tag{4.1}$$

where h_t denotes the natural logarithm of realized volatility for period t and i_t represents the natural logarithm of implied volatility at the beginning of period t.

Using this regression is possible to test three hypothesis. First, if implied volatility contains some information about future realized volatility, α_i should be different from zero. Second, $\alpha_0 = 0$ and $\alpha_i = 1$ means that implied volatility is a unbiased estimator for future volatility. Third, for implied volatility be efficient the residuals e_t should be a white noise.

Although Regression (4.1) allow us to understand the relation between implied volatility and future realized volatility its not enough to realize whether implied volatility is a more accurate estimator than historical volatility. To assess this, we must use a multiple regression adding a new variable

$$h_t = \alpha_0 + \alpha_h h_{t-1} + \varepsilon_t \tag{4.2}$$

$$h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + \varepsilon_t \tag{4.3}$$

Comparing these two regressions with Regression (4.1) it is possible make a comparison between implied volatility and historical volatility as predictor for future volatility. If implied volatility contains more information than historical volatility we would expect a higher R^2 from Regression (4.1) than from Regression (4.2) and this means that implied volatility has greater explanatory power than historical volatility. Moreover, if implied volatility incorporates all information regarding future volatility and historical volatility contains no additional information we should expect $\alpha_h = 0$ in Regression (4.3).

Beyond the conventional analysis, should be done an measurement errors analysis. There are several sources of measurement errors, related with features of the market or the model used to calculate implied volatility. These possibles errors-in-variables (EIV) are widely analyzed in earlier literature, since they may led to a lack of confidence in the results or even to fake conclusions.

Since B&S model is used to calculate implied volatility we must pay attention to its limitations. First, the model was created to apply on European options that does not pay dividends until maturity. Moreover, this formula is also based on assumption that the price of the underlying asset follows a log-normal distribution. If, for some reason, the value of the underlying vary dramatically the implied volatilities are calculated incorrectly. Another important aspect is the moneyness of the option. Options that are deep in or out of the money are very sensitive to volatility changes, so implied volatilities calculated from those options may not be reliable[1]. As we saw before, OESX options are European style and the use of ATM options eliminates the so called volatility smile effect.

Regarding the market there are two important aspects that deserve attention: (i) first, option price and index closing price may be non synchronous, because they are recorded at different times; (ii) Second, there may be an additional measurement error due to bid-ask spread.

To control for the possible EIV problem and assess whether these have some influence in our conclusions we will use an instrumental variable (IV), more specifically a regression of the form

$$i_t = \beta_0 + \beta_i i_{t-1} + \beta_h h_{t-1} + e_t \tag{4.4}$$

This regression, besides testing the presence of EIV, allows us to test whether implied volatility is predicted by past volatility. Is proved that past and future volatility are positively correlated, therefore implied volatility should not only predict future volatility but should also depend of past volatility.

Knowing that some of the problems are mitigated due to the use of ATM options with European Style, it is expected that the EIV has no influence, in other words, it is expected that the OLS estimate is consistent. To test this we use a Hausman (1975) test. In the presence of measurement error, the IV is consistent and provides a better estimator than OLS. However, whether implied volatility is correctly calculated, the IV outputs should not be used to any conclusion since it is less efficient than the OLS estimate.

[1] See Hull (2009), Chapter 13 and Chapter 16

5. Empirical results

This chapter reports the results of applying the methodology set out in Chapter 4. We start looking to the whole period and also with an analysis in order to check for possible measurement errors in variables. The last section of this chapter is devoted to analysis of the subperiods trying to understand whether the market conditions affect the predictive power of implied volatility.

5.1 Full Period analysis

OLS of Regression (4.1), that includes only intercept and implied volatility as explanatory variables, are reported in the first line of Table 5.1. We see that implied volatility contains information about future realized volatility, in accordance with Christensen and Prabhala (1998) and Hansen (2001), however there are some surprising aspects. The estimate of α_i is 1,005 and the intercept is not significantly different from 0 at a 95% level, what means that implied volatility appears to be a perfectly unbiased estimator of future volatility. In addition, the value of Adj.R^2 is twice that found in earlier studies, meaning that, in European market and in the last decade, a bigger percentage of future volatility is explained by implied volatility. The Durbin-Watson statistic is also not significantly different from two and so the implied volatility can be considered an efficient estimator.

In the second line, implied volatility is replaced by historical volatility as explanatory variable. Here we can see that historical volatility, when used as only explanatory variable, contains some information about future realized volatility. Nevertheless the value of Adj.R^2 drops more than 10% and, as we refer in Chapter 4, this means that implied volatility has better explanatory power than historical volatility.

Finally, and perhaps the most interesting analysis is presented in the last line (Table 5.1) where the results from Regression (4.3), which combine both variables as explanatory variables, are presented. Despite the results of second line, when implied volatility is added at the model, the value of α_h drops from 0,739 to 0,022 going to be not statistically significant, with a t-statistic of 0,198. In what concerns α_i, the value remains very close to what was found in

Table 5.1: Information content of implied and historical volatility: OLS estimates for full period

Dependent variable: realized volatility h_t

Independent variables			R^2	Adj.R^2	DW
Intercept	i_t	h_{t-1}			
-0,141 (-1,422)	$1,005^{***}$ (16,089)		68,15%	67,88%	1,811
$-0,438^{***}$ (-4,072)		$0,739^{***}$ (12,070)	54,63%	54,26%	2,190
-0,141 (-1,411)	$0,982^{***}$ (7,139)	0,022 (0,198)	68,16%	67,63%	1,813

*,**,*** significant at 10%, 5% and 1% respectively

Ordinary least squares estimates of the regression: $h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + \varepsilon_t$, where i_t denotes implied volatility for ATM options on Euro Stoxx 50 index measured at the beginning of month t and h_{t-1} denotes the historical volatility from the index. The data consist of monthly non-overlapping observations for the period from January 2002 to April 2012. DW is the Durbin-Watson statistic and the numbers in parentheses denote asymptotic t-statistics.

first line, what led us to conclude that implied volatility subsumes all information contained in historical volatility, reinforcing the idea that implied volatility is an unbiased and efficient estimator of future volatility. The Adj.R^2 of last Regression (4.3) does not suffer any significantly change comparing to the Adj.R^2 of the first line.

Results from measurement errors analysis are reported in Table 5.2 and the purpose of the tests related to these results is to test whether the possible errors mentioned in Chapter 4 have some influence on the results of the previous OLS estimates. The idea behind this method is to use values of implied volatility adjusted by the regression 4.4 instead of the implied volatility original values.

Looking at Panel A of Table 5.2, we see that implied volatility can be predicted by the first lag of implied volatility and historical volatility. In fact, the Adj.R^2 is very high (82,71%) which is an indication that these two variables explain almost total variation of implied volatility. The slope coefficient for historical volatility is about 0,50 and significantly different from zero. This evidence reinforces the idea that implied volatility not only predicts future volatility but also depends of past volatility.

Table 5.2: Information content of implied and historical volatility: 2SLS estimates for full period

Panel A: First Stage regression estimates

Dependent variable: implied volatility i_t

Independent variables			R^2	Adj.R^2	DW
Intercept	i_{t-1}	h_{t-1}			
$-0,169^{***}$ (-2,803)	$0,350^{***}$ (5,200)	$0,492^{***}$ (8,900)	83,00%	82,71%	2,466

Panel B: Second Stage regression estimates

Dependent variable: realized volatility h_t

Independent variables			R^2	Adj.R^2	DW
Intercept	i_t	h_{t-1}			
-0,144 (-1,313)	$1,004^{***}$ (14,490)		68,15%	67,88%	
-0,220 (-1,382)	$0,719^{*}$ (1,675)	0,214 (0,673)	67,30%	66,76%	

*,**,*** significant at 10%, 5% and 1% respectively

Panel A contains ordinary least squares estimates of the regression: $i_t = \beta_0 + \beta_i i_{t-1} + \beta_h h_{t-1} + e_t$, where i_t denotes implied volatility for one-month ATM options on Euro Stoxx 50 index measured at the beginning of month t and h_{t-1} denotes the historical volatility from the index. The data consist of 123 non-overlapping monthly observations covering the period from January 2002 to April 2012. Panel B reports the results from IV estimates for regression $h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + \varepsilon_t$ using implied volatility and historical volatility lagged one period as an instrument for implied volatility and using the same data as in Panel A. DW is the Durbin-Watson statistic and the numbers in parentheses denote asymptotic t-statistics.

Panel B reports the results from IV. The first line shows that, even using values for implied volatility fitted by Regression 4.4, the results does not differ from those found on OLS estimates. The slope coefficient for implied volatility remains very close to unity (drops from 1,005 to 1,004) and the Adj.R^2 , although smaller, keeps high.

Taking into account these results, we are led to believe that the explanatory variables are not measured with error. To confirm this we run a Hausman (1975) test. The idea of this test is to examine the difference $\tilde{\alpha} - \hat{\alpha}$ relative to the difference between the estimated variances, where $\tilde{\alpha}$ is the estimate for implied volatility from IV and $\hat{\alpha}$ is the estimate from OLS. A big difference between $\tilde{\alpha}$ and $\hat{\alpha}$ could be an indication of presence of measurement errors. Following the notation of Hansen (2001) we define

$$H = (\tilde{\alpha} - \hat{\alpha}) \left(\widehat{Var\,(\tilde{\alpha})} - \widehat{Var\,(\hat{\alpha})}\right)^{-1} (\tilde{\alpha} - \hat{\alpha}) = \frac{(\tilde{\alpha} - \hat{\alpha})^2}{\left(\widehat{Var\,(\tilde{\alpha})} - \widehat{Var\,(\hat{\alpha})}\right)} \tag{5.1}$$

and the statistic is asymptotically χ^2 distributed with one degree of freedom.

In our case H$=(1,004 - 1,005)^2/(0,069^2 - 0,063^2) = 0,0026$. Since the 95 per cent critical value for a χ^2 (one degree of freedom) distribution is 3,84, we cannot reject the hypothesis that implied volatility is calculated without measurement errors and we may argue that OLS estimates are consistent (p-value is almost 1). The Hausman test for when both variables are used as explanatory variables also indicate that OLS estimates are consistent.

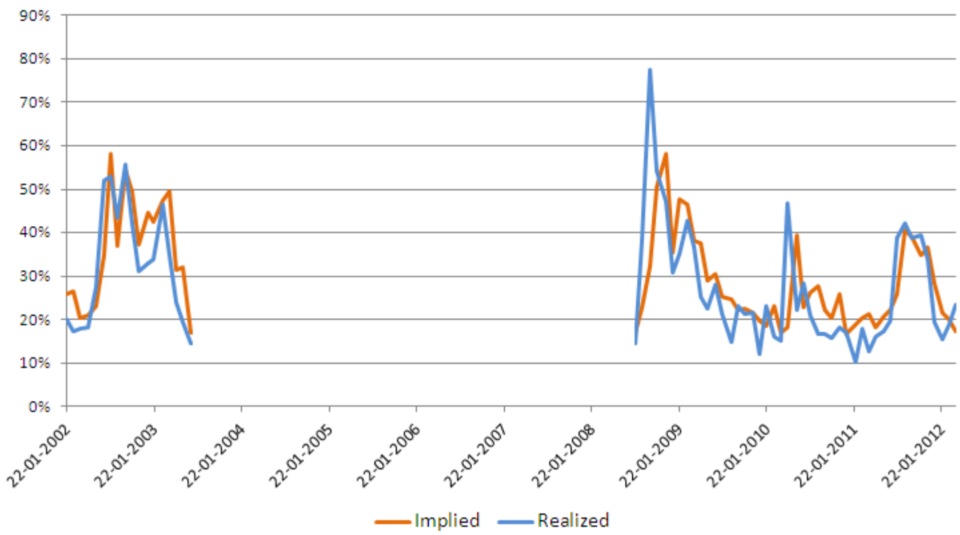
Hence, it is correct use implied volatility original values for subsequent analysis.

5.2 Sub period analysis

Section 5.1 shows that during the last 10 years, implied volatility has been an unbiased and efficient estimator for future volatility. However, the financial and economic environment on the last decade was not constant, there being some quiet periods and others more troubled. So, it becomes interesting to analyse whether predictive ability of implied volatility is influenced by market instability or if, on the other hand, it remains unchanged. In this section this issue is studied. The methodology is similar to that used previously but for two distinct periods, already stated above.

There are some visible changes and in general the crisis period is most similar to the whole period than non crisis period.

Figure 5.1: Relation between implied and realized volatility in crisis period

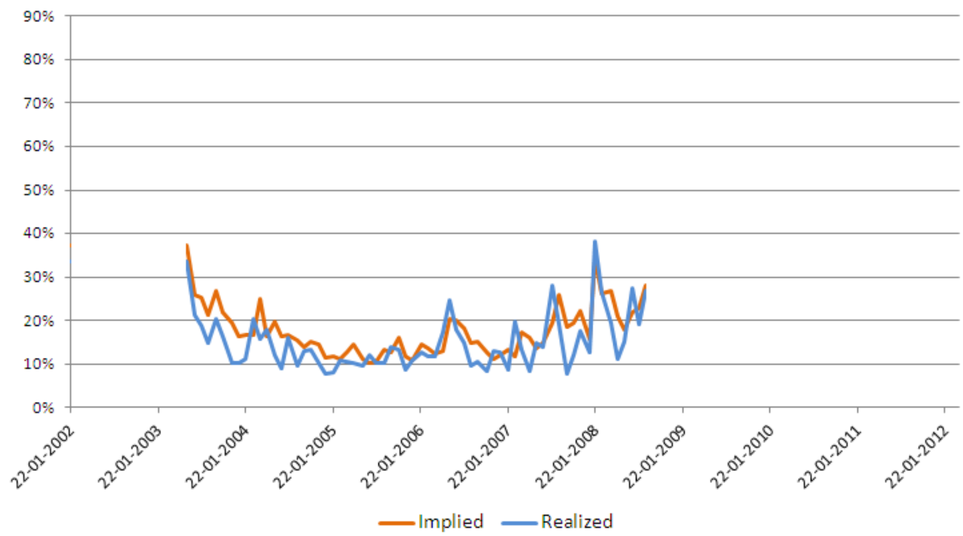


Implied volatility from one-month ATM options and realized volatility from Euro Stoxx 50 index. Implied Volatility is calculated as a simple average from implied call and implied put volatility as defined in equation 3.1. The data are monthly frequently with 61 non-overlapping observations and covers the period from January 2002 to June 2003 and September 2008 to April 2012.

It is possible see that the slope coefficient for implied volatility is closer to unity in moments of crisis than in non-crisis periods, 0,934 against 0,821 (first line of panel A and first line of panel B of Table 5.3, respectively). Besides in crisis periods implied volatility appears to be an unbiased estimator for future volatility contrary to what happens in stable moments. In non crisis periods we have α_0 different from zero, which is an indication that implied volatility is a biased estimator for future volatility, However, both cases present a DW significantly close to two which leads us to conclude that implied volatility is an efficient estimator. Regarding to the Adj.R^2, the value of the two periods has a significant difference. In times of crisis, the implied volatility explains about 53% of the total variation of future volatility, while in non-crisis periods the predictive power of implied volatility falls nearly 10%.

The second line of both panels report results when only historical volatility is used as explanatory variable. Here, the differences between the two periods are more pronounced. The slope coefficient for historical volatility in times of crisis is very similar to what was found in the whole period, contrary to what happens in times of non-crisis in which the slope coefficient decreases to almost half of the others. More interesting and more meaningful is the difference in the Adj.R^2. While the Adj.R^2 in times of crisis, although smaller, remains close to those

Figure 5.2: Relation between implied and realized volatility in non-crisis period



Implied volatility from one-month ATM options and realized volatility from Euro Stoxx 50 index. Implied Volatility is calculated as a simple average from implied call and implied put volatility as defined in equation 3.1. The data are monthly frequently with 62 non-overlapping observations and covers the period from June 2003 to September 2008.

found so far, in non-crisis moments the Adj.R^2 is much lower. Somehow this is surprising and means that in periods of financial stability, the historical volatility has almost no predictive power about future volatility.

Finally, the last line of both panels present the results when both explanatory variables are used. These are similar, and somehow in accordance with the results to the whole period. Once more, the model explains better the variation of future volatility in moments of crisis since has a higher Adj.R^2 when compared to non-crisis times, despite have a lower value of slope coefficient for implied volatility. More importantly is that in both cases, the slope coefficient for historical volatility is not significantly different from zero, with 95% confidence. This means that in none of the periods, the historical volatility contains information beyond what is already reflected in implied volatility.

Summing up the two periods present some differences. The results show that the market conditions affect implied volatility predictive ability. Implied volatility can only be considered as an unbiased estimator in crisis periods. However, in both periods historical volatility does not contain any information beyond what is already reflected in implied volatility, reinforcing what was seen in Section 5.1. Finally, and perhaps the most interesting conclusion is that the implied volatility and even historical volatility contains more predictive power in crisis times than when markets are stable.

However this results should be interpreted with careful. There is definitely a significant drop in Adj.R^2, nevertheless some percentage of this drop is a consequence of variable characteristics and OLS model. In non crisis periods, when markets are more stable and financial assets have smoother behaviors, the variables variation is lower, which itself decreases the Adj.R^2.

Table 5.3: Information content of implied and historical volatility: OLS estimates for sub periods

Dependent variable: realized volatility h_t					
Independent variables			R^2	Adj.R^2	DW
Intercept	i_t	h_{t-1}			
Panel A: Crisis period					
-0,188 (-1,277)	$0,934^{***}$ (8,215)		53,36%	52,56%	1,812
$-0,437^{***}$ (-3,200)		$0,677^{***}$ (7,078)	45,92%	45,00%	2,019
-0,197 (-1,326)	$0,771^{***}$ (3,148)	0,145 (0,756)	53,81%	52,22%	1,895
Panel B: Non crisis period					
$-0,516^{**}$ (-2,337)	$0,821^{***}$ (6,806)		43,57%	42,63%	1,855
$-1,206^{***}$ (-5,017)		$0,398^{***}$ (3,366)	15,88%	14,48%	1,972
$-0,570^{**}$ (-2,495)	$0,933^{***}$ (5,501)	-0,128 (-0,939)	44,40%	42,51%	1,714

*,**,*** significant at 10%, 5% and 1% respectively

Ordinary least squares estimates of the regression: $h_t = \alpha_0 + \alpha_i i_t + \alpha_h h_{t-1} + \varepsilon_t$, where i_t denotes implied volatility from one-month ATM options on Euro Stoxx 50 index measured at the beginning of month t and h_{t-1} denotes the historical volatility from the index. The results reported in Panel A are based on 61 monthly non-overlapping observations for the period from January 2002 to June 2003 and September 2008 to April 2012, while those reported in Panel B are based on 62 monthly non-overlapping observations for the period from June 2003 to September 2008 . DW is the Durbin-Watson statistic and the numbers in parentheses denote asymptotic t-statistics.

6. Call Vs Put

Chapter 5 shows that implied volatility is an unbiased and efficient estimator for future volatility and subsumes all information available including past volatility. There we used average implied volatility of both call and put options. However some earlier studies focus their analysis on implied volatility of call options as Lamoureux and Lastrapes (1993) and Christensen and Prabhala (1998). There are some empirical reason for that? Can we argue that implied volatility from call options have more explanatory power over future volatility than implied volatility from put options? There are studies to Danish and U.S markets - Hansen (2001) and Christensen and Hansen (2002) respectively - which indicate, in fact, that implied volatility from call options is a more accurate estimator for future volatility. Do these findings can be extended to the European market?

In this chapter we study this issue, trying to understand whether there is some significant difference from call and puts options in what concerns to predictive ability over future volatility.

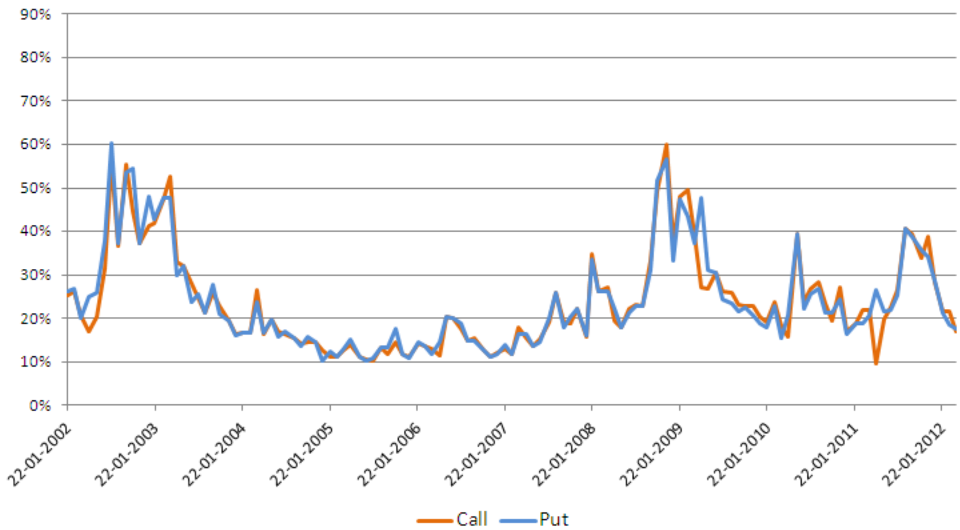
6.1 Descriptive statistics

Since both call and put implied volatility are related to the same underlying, would expect that the value were equal, once implied volatility should reflect the volatility of the underlying during the life time of option and this underlying is the same for calls and puts. However, as we can see in the Figure 6.1, although similar they are not exactly equal. This is not surprising because, as refer Harvey and Whaley (1992), buy index put options is a convenient and inexpensive way to implement portfolio insurance, what leads to an excess buying pressure on index puts and consequently to a higher implied volatility on put options.

We can see that both implied call and put volatilities are slightly higher than realized volatility reported on Table 3.1. Reinforcing what was said before, implied put volatility is higher and is also more volatile.

Analyzing the four last lines of the table it is possible to see that both log-transformed series are normally distributed. Thus, as in previous chapters, we focus our analysis on log-

Figure 6.1: Relation between implied call and implied put volatility



Implied in one-month ATM call and put options on Euro Stoxx 50 index. Monthly data with 123 non-overlapping observations and covers the period from January 2002 to April 2012.

transformed series and let $i_{c,t}$ denote the natural logarithm of implied call volatility and $i_{p,t}$ denote the natural logarithm of implied put volatility.

6.2 Empirical results

Once again, to examine the information content of put and call implied volatility is used a multiple regression, but this time implied call volatility and implied put volatility appear as separate variables.

$$h_t = \alpha_0 + \alpha_c i_{c,t} + \alpha_p i_{p,t} + \alpha_h h_{t-1} + e_t \tag{6.1}$$

The OLS estimates are presented in Table 6.2. Using the regression above, we conduct a series of combinations that allowed arrive at interesting conclusions.

The first line shows the results when all variables are included. The sum of all coefficients gives a value very close to unity (1,006). Despite all the slope coefficients are positive, only the coefficient of implied put volatility is significant at 95% and this is very interesting because is already a difference compared to previous studies. However, to support this evidence we have to look at the next lines. Finally, the Adj.R^2 is very high in accordance to what was seen in

Table 6.1: Descriptive statistics: implied call volatility and implied put volatility

Statistic	Implied Call	Implied Put	Log Implied Call	Log Implied Put
Mean	0,2363	0,2390	-1,5395	-1,5270
Std. Deviation	0,1113	0,1127	0,4331	0,4285
Skewness	1,2304	1,2695	0,4337	0,4285
Excess kurtosis	1,0361	1,0458	-0,5519	-0,4767
Jarque-Bera	36,5389	38,6454	4,1693	4,6845
Prob	0,0000	0,0000	0,1243	0,0961

Implied call and implied put volatility time series on Euro Stoxx 50 and their natural logarithmic transformation. Statistics are based on monthly data with 123 non-overlapping observations and covers the period from January 2002 to April 2012.

Chapter 5.

The next line shows that if implied put volatility is excluded from the regression, implied call volatility is significant and subsumes the information content of historical volatility. However the value for slope coefficient is not close to unity. Similar results are obtained in third line, where implied put volatility replaces implied call volatility in the regression. Implied put volatility also subsumes the information content of historical volatility and besides this, contrary to what happens with calls, the slope coefficient for implied put volatility is very close to unity.

The fourth line confirms that, even when historical volatility is excluded from the regression, implied put volatility dominates over implied call volatility. The slope coefficient for implied put volatility is three times bigger than slope coefficient for implied call volatility and this is not significant at 95%. The both coefficients add up almost to unity and the Adj.R^2 is also very high, which indicates that this combination of calls and puts can be a very accurate estimator for future volatility.

The last two lines allow us to conclude which of the two variables is the best estimator for future volatility. Here they are used as only explanatory variables and both have slope coefficients close to unity. This way both can be considered unbiased estimator for future volatility, however the Adj.R^2 is higher using implied put volatility which means that implied put volatility is a better forecast for future volatility than implied call volatility. These results go against what was found before by Hansen (2001) and Christensen and Hansen (2002) and shows that put options has more information and more explanatory power than that which has been assigned.

Table 6.2: Information content of implied call, implied put and historical volatility: OLS estimates

Dependent variable: realized volatility h_t

Independent variables				R^2	Adj.R^2	DW
Intercept	i_{ct}	i_{pt}	h_{t-1}			
-0,140 (-1,402)	0,237 (1,150)	$0,745^{***}$ (3,585)	0,024 (0,217)	68,49%	67,70%	1,830
$-0,192^{*}$ (-1,862)	$0,813^{***}$ (5,997)		0,141 (1,246)	65,09%	64,51%	1,880
-0,155 (-1,564)		$0,930^{***}$ (7,135)	0,064 (0,592)	68,14%	67,61%	1,861
-0,140 (-1,412)	0,251 (1,280)	$0,757^{***}$ (3,825)		68,48%	67,96%	1,826
$-0,201^{*}$ (-1,947)	$0,961^{***}$ (14,873)			64,64%	64,35%	1,771
-0,159 (-1,611)		$0,997^{***}$ (16,054)		68,05%	67,79%	1,820

*,**,*** significant at 10%, 5% and 1% respectively

This table contains ordinary least squares estimates of the regression: $h_t = \alpha_0 + \alpha_c i_{c,t} + \alpha_p i_{p,t} + \alpha_h h_{t-1} + e_t$ where $i_{c,t}$ and $i_{p,t}$ denote implied volatility calculated by Datastream for call and put options on Euro Stoxx 50 index respectively, measured at the beginning of month t and h_{t-1} denotes the historical volatility from the index. The data consist of 123 monthly non-overlapping observations for the period from January 2002 to April 2012. DW is the Durbin-Watson statistic and the numbers in parentheses denote asymptotic t-statistics.

7. Conclusions

In this study we test whether implied volatility in OESX options prices can predict future index return volatility and how can be combined with historical volatility.

Our results confirm that those found by Christensen and Prabhala (1998) and Hansen (2001) can be extended to European market and go against what was concluded by Canina and Figlewski (1993) and more recently by Pakarinen (2007). We find that implied volatility is an unbiased and efficient forecast for future volatility. Moreover this variable subsumes the information content in historical volatility, since this last one is not significant when both variables are used to predict future volatility, which is an indication of market efficiency. So, the best way to forecast future volatility is use only implied volatility as estimator and this explains almost 70% of total variation of future volatility.

This study uses recent data and reflects today's market conditions. When compared with earlier studies, it seems that implied volatility has now more predictive power than in the 90's which can be an indication that market has become more efficient and the information available on market allows to predict future volatility accurately.

As in the previous literature we used an IV technique to test the existence of measurement errors. After applying this technique we conclude that the results found are reliable and robust and with no EIV that could affect our conclusions. Besides bring some robustness to the study, this IV technique also has great interest at economic level. Through this technique we show that implied volatility can also be estimated. Natural candidates have been implied volatility and historical volatility lagged one period and found that this simple model of two variables allows for a fairly accurate estimation of implied volatility.

Regarding to the sub periods analysis the results show that the market conditions affect implied volatility predictive ability. Implied volatility can only be considered as an unbiased estimator in crisis periods. However, in both periods historical volatility does not contain any additional information beyond what is already reflected in implied volatility, as happened with the whole period. Another important aspect is that implied volatility appears to have more predictive ability during crisis periods. However as indicated in the last paragraph of sub section 5.2 this conclusion should be interpreted carefully

Finally, as Christensen and Hansen (2002) we test implied volatility from call and put options separately to understand whether there is some significant difference regarding to predictive ability. Although Christensen and Hansen (2002) study conclude that implied volatility from call options provide a better forecast for future volatility and most of the previous studies focus their analysis on call options, our results show a different scenario.

Using the 10 years data as a whole, the OLS provides evidence that the volatility implied by put options allows a better forecast than volatility implied by call options and historical volatility. This way we show that put options has more predictive power than that which been assigned and for that seems there is no reason to conduct further analysis on call options instead of put options, at least on OESX options.

Bibliography

Black, F. and M. S. Scholes, 1973: The pricing of options and corporate liabilities. *Journal of Political Economy*, **81 (3)**, 637–54.

Canina, L. and S. Figlewski, 1993: The informational content of implied volatility. *Review of Financial Studies*, **6 (3)**, 659–681.

Chiras, D. P. and S. Manaster, 1978: The information content of option prices and a test of market efficiency. *Journal of Financial Economics*, **6 (2-3)**, 213–234.

Christensen, B. and C. Hansen, 2002: New evidence on the implied-realized volatility relation. *The European Journal of Finance*, 187–205.

Christensen, B. and N. Prabhala, 1998: The relation between implied and realized volatility. *Journal of Financial Economics*, **50 (2)**, 125 – 150.

Day, T. E. and C. M. Lewis, 1992: Stock market volatility and the information content of stock index options. *Journal of Econometrics*, **52 (1 - 2)**, 267 – 287.

Duque, J., 2000: Empirical evidence on volatility estimators, working paper (iseg).

Hansen, C. S., 2001: The relation between implied and realised volatility in the danish option and equity markets. *Accounting & Finance*, **41 (3)**, 197–228.

Harvey, C. R. and R. E. Whaley, 1992: Market volatility prediction and the efficiency of the s&p 100 index option. *Journal of Financial Economics*, **31 (1)**, 43–73.

Hausman, J. A., 1975: An instrumental variable approach to full information estimators for linear and certain nonlinear econometric models. *Econometrica*, **43 (4)**, 727–38.

Hull, J., 2009: *Options, Futures and Other Derivatives*. Prentice Hall Finance Series, Pearson/Prentice Hall.

Jorion, P., 1995: Predicting volatility in the foreign exchange market. *The Journal of Finance*, **50 (2)**, pp. 507–528.

Lamoureux, C. G. and W. D. Lastrapes, 1993: Forecasting stock-return variance: Toward an understanding of stochastic implied volatilities. *The Review of Financial Studies*, **6 (2)**, pp. 293–326.

Latane, H. A. and J. Rendleman, Richard J, 1976: Standard deviations of stock price ratios implied in option prices. *Journal of Finance*, **31 (2)**, 369–81.

Pakarinen, J., 2007: The relationship between implied and realized volatility, bachelor thesis.

Parkinson, M., 1980: The extreme value method for estimating the variance of the rate of return. *The Journal of Business*, **53 (1)**, 61–65.

Taylor, S. J., P. K. Yadav, and Y. Zhang, 2009: The information content of implied volatilities and model-free volatility expectations: Evidence from options written on individual stocks. *Journal of Banking & Finance*, **34 (4)**, 871 – 881.

Yang, D. and Q. Zhang, 2000: Drift-independent volatility estimation based on high, low, open, and close prices. *The Journal of Business*, **73 (3)**, 477–91.

Yu, W. W., E. C. Lui, and J. W. Wang, 2010: The predictive power of the implied volatility of options traded otc and on exchanges. *Journal of Banking & Finance*, **34 (1)**, 1 – 11.